

Methods in  
Molecular Biology 1672

Springer Protocols



Marco Muzi-Falconi  
Grant W. Brown *Editors*

# Genome Instability

Methods and Protocols

 Humana Press

# METHODS IN MOLECULAR BIOLOGY

*Series Editor*

**John M. Walker**

**School of Life and Medical Sciences**

**University of Hertfordshire**

**Hatfield, Hertfordshire, AL10 9AB, UK**

For further volumes:

<http://www.springer.com/series/7651>

# Genome Instability

## Methods and Protocols

Edited by

**Marco Muzi-Falconi**

*Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Milano, Italy*

**Grant W. Brown**

*Donnelly Centre and Department of Biochemistry, University of Toronto, Toronto, ON, Canada*

 **Humana Press**

*Editors*

Marco Muzi-Falconi  
Dipartimento di Bioscienze  
Università degli Studi di Milano  
Milano, Milano, Italy

Grant W. Brown  
Donnelly Centre and Department of Biochemistry  
University of Toronto  
Toronto, ON, Canada

ISSN 1064-3745                      ISSN 1940-6029 (electronic)  
Methods in Molecular Biology  
ISBN 978-1-4939-7305-7            ISBN 978-1-4939-7306-4 (eBook)  
DOI 10.1007/978-1-4939-7306-4

Library of Congress Control Number: 2017953940

© Springer Science+Business Media LLC 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Humana Press imprint is published by Springer Nature  
The registered company is Springer Science+Business Media, LLC  
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.



---

## **Preface**

The integrity of the genome is a fundamental determinant of cellular identity, cellular fitness, and interactions between a cell and its environment. The study of genome integrity is now a mature field, but one marked by continuous innovations in techniques, technology, and systems, both in vitro and in vivo. We present 42 methods and protocols to analyze diverse aspects of genome instability.

Beginning in the realm of mutagenesis and repair, we present classic genetic assays to detect chromosome loss, mutation, and genome rearrangements, whole genome approaches to mapping base modifications and repair events, and a method for analyzing engineered base lesions in the genomic context. Methods to quantify and analyze the properties of DNA double-strand breaks include traditional and single-molecule approaches to measure double-strand break resection, and modern methods to map double-strand breaks at high resolution and high sensitivity. Given the importance of DNA replication errors as a source of genome instability, we include methods to profile replication, to probe replication and replication proteins strand specifically, to analyze replication intermediates at high resolution, and to specifically perturb DNA synthesis at specific sites. The increasing interest in the role of ribonucleotides and RNA–DNA hybrids in genome instability is reflected in methods to detect and map ribonucleotides and RNA–DNA hybrids. Techniques to study genome instability at specialized regions, in particular the telomeres and triplet nucleotide repeats, are presented and include molecular biological, genetic, and imaging-based methods. The application of imaging techniques to study genome instability has become common in the field. We present fluorescence microscopic techniques to detect and analyze genome instability, including single-molecule and single-cell analysis, as well as high-resolution methods to probe DNA structural properties. Finally, the contributions of genomic and proteomic approaches to identifying and defining genome instability pathways and networks are reflected in procedures for measuring cell fitness, protein interactions, gene and protein expression, protein–DNA interactions, and protein modifications, on a genome/proteome scale.

Together, the methods and protocols here form a comprehensive resource for the discovery and analysis of the proteins and pathways that are critical for stable maintenance of the genome.

*Milano, Italy*  
*Toronto, Canada*

*Marco Muzi-Falconi*  
*Grant W. Brown*

---

# Contents

<i>Preface</i> .....	<i>v</i>
<i>Contributors</i> .....	<i>xi</i>
1 The A-Like Faker Assay for Measuring Yeast Chromosome III Stability .....	1
<i>Carolina A. Novoa, J. Sidney Ang, and Peter C. Stirling</i>	
2 The Chromosome Transmission Fidelity Assay for Measuring Chromosome Loss in Yeast .....	11
<i>Supipi Duffy and Philip Hieter</i>	
3 Measuring Mutation Rates Using the Luria-Delbrück Fluctuation Assay .....	21
<i>Gregory I. Lang</i>	
4 Molecular Genetic Characterization of Mutagenesis Using a Highly Sensitive Single-Stranded DNA Reporter System in Budding Yeast .....	33
<i>Kin Chan</i>	
5 Analyzing Genome Rearrangements in <i>Saccharomyces cerevisiae</i> .....	43
<i>Anjana Srivatsan, Christopher D. Putnam, and Richard D. Kolodner</i>	
6 High-Resolution Mapping of Modified DNA Nucleobases Using Excision Repair Enzymes .....	63
<i>Monica Ransom, D. Suzi Bryan, and Jay R. Hesselberth</i>	
7 Integrated Microarray-based Tools for Detection of Genomic DNA Damage and Repair Mechanisms .....	77
<i>Patrick van Eijk, Yumin Teng, Mark R. Bennet, Katie E. Evans, James R. Powell, Richard M. Webster, and Simon H. Reed</i>	
8 Study of UV-induced DNA Repair Factor Recruitment: Kinetics and Dynamics .....	101
<i>Sarah Sertic, Stefania Roma, Paolo Plevani, Federico Lazzaro, and Marco Muzi-Falconi</i>	
9 Inserting Site-Specific DNA Lesions into Whole Genomes .....	107
<i>Vincent Pagès and Robert P. Fuchs</i>	
10 A qPCR-Based Protocol to Quantify DSB Resection .....	119
<i>Matteo Ferrari, Shyam Twayana, Federica Marini, and Achille Pellicioli</i>	
11 Alkaline Denaturing Southern Blot Analysis to Monitor Double-Strand Break Processing .....	131
<i>Chiara Vittoria Colombo, Luca Menin, and Michela Clerici</i>	
12 Single Molecule Analysis of Resection Tracks .....	147
<i>Pablo Huertas and Andrés Cruz-García</i>	
13 Mapping DNA Breaks by Next-Generation Sequencing .....	155
<i>Laura Baranello, Fedor Kouzine, Damian Wojtowicz, Kairong Cui, Keji Zhao, Teresa M. Przytycka, Giovanni Capranico, and David Levens</i>	

14	Genome-Wide Profiling of DNA Double-Strand Breaks by the BLESS and BLISS Methods. . . . .	167
	<i>Reza Mirzazadeh, Tomasz Kallas, Magda Bienko, and Nicola Crosetto</i>	
15	DNA Replication Profiling Using Deep Sequencing. . . . .	195
	<i>Xanita Saayman, Cristina Ramos-Pérez, and Grant W. Brown</i>	
16	Quantitative Bromodeoxyuridine Immunoprecipitation Analyzed by High-Throughput Sequencing (qBrdU-Seq or QBU) . . . . .	209
	<i>Joanna E. Haye-Bertolozzi and Oscar M. Aparicio</i>	
17	Strand-Specific Analysis of DNA Synthesis and Proteins Association with DNA Replication Forks in Budding Yeast . . . . .	227
	<i>Chuanhe Yu, Haiyun Gan, and Zhiguo Zhang</i>	
18	Analysis of Replicative Polymerase Usage by Ribonucleotide Incorporation . . . . .	239
	<i>Andrea Keszthelyi, Izumi Miyabe, Katie Ptasińska, Yasukazu Daigaku, Karel Naiman, and Antony M. Carr</i>	
19	Dynamic Architecture of Eukaryotic DNA Replication Forks In Vivo, Visualized by Electron Microscopy . . . . .	261
	<i>Ralph Zellweger and Massimo Lopes</i>	
20	A Molecular Toolbox to Engineer Site-Specific DNA Replication Perturbation . . . . .	295
	<i>Nicolai B. Larsen, Ian D. Hickson, and Hocine W. Mankouri</i>	
21	Single Cell Gel Electrophoresis for the Detection of Genomic Ribonucleotides . . . . .	311
	<i>Barbara Kind, Christine Wolf, Kerstin Engel, Alexander Rapp, M. Cristina Cardoso, and Min Ae Lee-Kirsch</i>	
22	Measuring the Levels of Ribonucleotides Embedded in Genomic DNA . . . . .	319
	<i>Alice Meroni, Giulia M. Nava, Sarah Sertic, Paolo Plevani, Marco Muzi-Falconi, and Federico Lazzaro</i>	
23	Mapping Ribonucleotides Incorporated into DNA by <i>Hydrolytic End-Sequencing</i> . . . . .	329
	<i>Clinton D. Orebaugh, Scott A. Lujan, Adam B. Burkholder, Anders R. Clausen, and Thomas A. Kunkel</i>	
24	Detection of DNA-RNA Hybrids In Vivo . . . . .	347
	<i>María García-Rubio, Sonia I. Barroso, and Andrés Aguilera</i>	
25	Analysis of De Novo Telomere Addition by Southern Blot. . . . .	363
	<i>Diego Bonetti and Maria Pia Longhese</i>	
26	Assays to Study Repair of Inducible DNA Double-Strand Breaks at Telomeres . . . . .	375
	<i>Roxanne Oshidari and Karim Mekhail</i>	

27	Telomerase RNA Imaging in Budding Yeast and Human Cells by Fluorescent In Situ Hybridization . . . . .	387
	<i>David Guérit, Maxime Lalonde, and Pascal Chartrand</i>	
28	Methods to Study Repeat Fragility and Instability in <i>Saccharomyces cerevisiae</i> . . . . .	403
	<i>Erica J. Polleys and Catherine H. Freudenreich</i>	
29	Quantitative Analysis of the Rates for Repeat-Mediated Genome Instability in a Yeast Experimental System. . . . .	421
	<i>Elina A. Radchenko, Ryan J. McGinty, Anna Y. Aksenova, Alexander J. Neil, and Sergei M. Mirkin</i>	
30	Measuring Dynamic Behavior of Trinucleotide Repeat Tracts In Vivo in <i>Saccharomyces cerevisiae</i> . . . . .	439
	<i>Gregory M. Williams and Jennifer A. Surtees</i>	
31	The Detection and Analysis of Chromosome Fragile Sites. . . . .	471
	<i>Victoria A. Bjerregaard, Özgün Özer, Ian D. Hickson, and Ying Liu</i>	
32	Imaging of DNA Ultrafine Bridges in Budding Yeast. . . . .	483
	<i>Oliver Quevedo and Michael Lisby</i>	
33	Detection of Ultrafine Anaphase Bridges. . . . .	495
	<i>Anna H. Bizard, Christian F. Nielsen, and Ian D. Hickson</i>	
34	A Chromatin Fiber Analysis Pipeline to Model DNA Synthesis and Structures in Fission Yeast. . . . .	509
	<i>Sarah A. Sabatinos and Marc D. Green</i>	
35	Long-Term Imaging of DNA Damage and Cell Cycle Progression in Budding Yeast Using Spinning Disk Confocal Microscopy. . . . .	527
	<i>Riccardo Montecchi and Etienne Schwob</i>	
36	The CellClamper: A Convenient Microfluidic Device for Time-Lapse Imaging of Yeast . . . . .	537
	<i>Gregor W. Schmidt, Olivier Frey, and Fabian Rudolf</i>	
37	Characterization of Structural and Configurational Properties of DNA by Atomic Force Microscopy . . . . .	557
	<i>Alice Meroni, Federico Lazzaro, Marco Muzi-Falconi, and Alessandro Podestà</i>	
38	Genome-Wide Quantitative Fitness Analysis (QFA) of Yeast Cultures . . . . .	575
	<i>Eva-Maria Holstein, Conor Lawless, Peter Banks, and David Lydall</i>	
39	Rewiring the Budding Yeast Proteome using Synthetic Physical Interactions . . . . .	599
	<i>Guðjón Ólafsson and Peter H. Thorpe</i>	
40	Reporter-Based Synthetic Genetic Array Analysis: A Functional Genomics Approach for Investigating Transcript or Protein Abundance Using Fluorescent Proteins in <i>Saccharomyces cerevisiae</i> . . . . .	613
	<i>Hendrikje Göttert, Mojca Mattiazzi Usaj, Adam P. Rosebrock, and Brenda J. Andrews</i>	

41	Statistical Analysis and Quality Assessment of ChIP-seq Data with DROMPA.....	631
	<i>Ryuichiro Nakato and Katsuhiko Shirahige</i>	
42	Quantitative Analysis of DNA Damage Signaling Responses to Chemical and Genetic Perturbations.....	645
	<i>Francisco M. Bastos de Oliveira, Dongsung Kim, Michael Lanz, and Marcus B. Smolka</i>	
	<i>Index</i> .....	661

---

## Contributors

- ANDRÉS AGUILERA • *Centro Andaluz de Biología Molecular y Medicina Regenerativa, CABIMER, Universidad de Sevilla, Seville, Spain*
- ANNA Y. AKSENOVA • *Amyloid Biology, Saint-Petersburg University, Saint Petersburg, Russia*
- BRENDA J. ANDREWS • *The Donnelly Centre, University of Toronto, Toronto, ON, Canada; Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada*
- J. SIDNEY ANG • *Michael Smith Laboratories, University of British Columbia, Vancouver, Canada*
- OSCAR M. APARICIO • *Molecular and Computational Biology Program, University of Southern California, Los Angeles, CA, USA*
- PETER BANKS • *Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Newcastle upon Tyne, UK*
- LAURA BARANELLO • *Laboratory of Pathology, NCI/NIH, Bethesda, MD, USA; Department of Cellular and Molecular Biology, Karolinska Institutet, Stockholm, Sweden*
- SONIA I. BARROSO • *Centro Andaluz de Biología Molecular y Medicina Regenerativa, CABIMER, Universidad de Sevilla, Seville, Spain*
- MARK R. BENNET • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- MAGDA BIENKO • *Science for Life Laboratory, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden*
- ANNA H. BIZARD • *Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Panum Institute, Copenhagen N, Denmark*
- VICTORIA A. BJERREGAARD • *Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Copenhagen N, Denmark*
- DIEGO BONETTI • *Institute for Molecular Biology (IMB) gGMBH, Mainz, Germany*
- GRANT W. BROWN • *Donnelly Centre and Department of Biochemistry, University of Toronto, Toronto, ON, Canada*
- D. SUZI BRYAN • *Department of Biochemistry and Molecular Genetics, Program in Molecular Biology, University of Colorado School of Medicine, Aurora, CO, USA*
- ADAM B. BURKHOLDER • *Integrative Bioinformatics, National Institute for Environmental Health Sciences, National Institute of Health (NIH), Research Triangle Park, NC, USA*
- GIOVANNI CAPRANICO • *Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy*
- M. CRISTINA CARDOSO • *Department of Biology, Technische Universität Darmstadt, Darmstadt, Germany*
- ANTONY M. CARR • *University of Sussex, Brighton, UK*
- KIN CHAN • *Department of Biochemistry, Microbiology, and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Canada*
- PASCAL CHARTRAND • *Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC, Canada*
- ANDERS R. CLAUSEN • *Department of Medical Biochemistry and Cell Biology, University of Gothenburg, Gothenburg, Sweden*
- MICHELA CLERICI • *Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy*

- CHIARA VITTORIA COLOMBO • *Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy*
- NICOLA CROSETTO • *Science for Life Laboratory, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden*
- ANDRÉS CRUZ-GARCÍA • *Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, Sevilla, Spain*
- KAIRONG CUI • *Systems Biology Center, NCI/NIH, Bethesda, MD, USA*
- YASUKAZU DAIGAKU • *Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai, Japan; Graduate School of Life Sciences, Tohoku University, Sendai, Japan*
- SUPIPI DUFFY • *Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada*
- PATRICK VAN EIJK • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- KERSTIN ENGEL • *Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany*
- KATIE E. EVANS • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- MATTEO FERRARI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- CATHERINE H. FREUDENREICH • *Department of Biology, Tufts University, Medford, MA, USA*
- OLIVIER FREY • *Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland*
- ROBERT P. FUCHS • *Cancer Research Center of Marseille, Team DNA Damage Tolerance, CNRS, UMR7258, Marseille, France; Inserm, U1068, Marseille, France; Institut Paoli-Calmettes, Marseille, France; Aix Marseille Univ., UM 105, Marseille, France*
- HAIYUN GAN • *Institute for Cancer Genetics, Columbia University, New York, NY, USA; Department of Pediatrics, Columbia University, New York, NY, USA; Department of Genetics and Development, Columbia University, New York, NY, USA; Irving Cancer Research Center New York, NY, USA*
- MARÍA GARCÍA-RUBIO • *Centro Andaluz de Biología Molecular y Medicina Regenerativa, CABIMER, Universidad de Sevilla, Seville, Spain*
- HENDRIKJE GÖTTERT • *The Donnelly Centre, University of Toronto, Toronto, ON, Canada; Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada*
- MARC D. GREEN • *Royal Ontario Museum, Toronto, ON, Canada*
- DAVID GUÉRIT • *Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC, Canada*
- JOANNA E. HAYE-BERTOLOZZI • *Molecular and Computational Biology Program, University of Southern California, Los Angeles, CA, USA*
- JAY R. HESSELBERTH • *Department of Biochemistry and Molecular Genetics, Program in Molecular Biology, University of Colorado School of Medicine, Aurora, CO, USA*
- IAN D. HICKSON • *Center for Chromosome Stability, Department of Cellular and Molecular Medicine, University of Copenhagen, Panum Institute, Copenhagen N, Denmark*
- PHILIP HIETER • *Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada*
- EVA-MARIA HOLSTEIN • *Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Newcastle upon Tyne, UK*

- PABLO HUERTAS • *Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, Sevilla, Spain*
- TOMASZ KALLAS • *Science for Life Laboratory, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden*
- ANDREA KESZTHELYI • *University of Sussex, Brighton, UK*
- DONGSUNG KIM • *Department of Molecular Biology and Genetics, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, USA*
- BARBARA KIND • *Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany*
- RICHARD D. KOLODNER • *Ludwig Institute for Cancer Research, University of California School of Medicine, San Diego, La Jolla, CA, USA; Department of Cellular and Molecular Medicine, University of California School of Medicine, San Diego, La Jolla, CA, USA; Moores-UCSD Cancer Center, University of California School of Medicine, San Diego, La Jolla, CA, USA; Institute of Genomic Medicine, University of California School of Medicine, San Diego, La Jolla, CA, USA*
- FEDOR KOUZINE • *Laboratory of Pathology, NCI/NIH, Bethesda, MD, USA*
- THOMAS A. KUNKEL • *Genome Integrity and Structural Biology Laboratory, National Institute for Environmental Health Sciences, National Institute of Health (NIH), Research Triangle Park, NC, USA*
- MAXIME LALONDE • *Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC, Canada*
- GREGORY I. LANG • *Department of Biological Sciences, Lehigh University, Bethlehem, PA, USA*
- MICHAEL LANZ • *Department of Molecular Biology and Genetics, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, USA*
- NICOLAI B. LARSEN • *The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark*
- CONOR LAWLESS • *Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Newcastle upon Tyne, UK*
- FEDERICO LAZZARO • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- MIN AE LEE-KIRSCH • *Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany*
- DAVID LEVENS • *Laboratory of Pathology, NCI/NIH, Bethesda, MD, USA*
- MICHAEL LISBY • *Department of Biology, University of Copenhagen, Copenhagen N, Denmark; Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Copenhagen N, Denmark*
- YING LIU • *Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Copenhagen N, Denmark*
- MARIA PIA LONGHESE • *Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy*
- MASSIMO LOPES • *Institute of Molecular Cancer Research, University of Zurich, Zurich, Switzerland*
- SCOTT A. LUJAN • *Genome Integrity and Structural Biology Laboratory, National Institute for Environmental Health Sciences, National Institute of Health (NIH), Research Triangle Park, NC, USA*
- DAVID LYDALL • *Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Newcastle upon Tyne, UK*



- HOCINE W. MANKOURI • *The Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark*
- FEDERICA MARINI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- RYAN J. MCGINTY • *Department of Biology, Tufts University, Medford, MA, USA*
- KARIM MEKHAIL • *Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada*
- LUCA MENIN • *Dipartimento di Biotechnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy*
- ALICE MERONI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- SERGEI M. MIRKIN • *Department of Biology, Tufts University, Medford, MA, USA*
- REZA MIRZAZADEH • *Science for Life Laboratory, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden*
- IZUMI MIYABE • *University of Sussex, Brighton, UK*
- RICCARDO MONTECCHI • *IGMM, CNRS, University of Montpellier, Montpellier, France*
- MARCO MUZI-FALCONI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- KAREL NAIMAN • *University of Sussex, Brighton, UK*
- RYUICHIRO NAKATO • *Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo, Japan*
- GIULIA M. NAVA • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- ALEXANDER J. NEIL • *Department of Biology, Tufts University, Medford, MA, USA*
- CHRISTIAN F. NIELSEN • *Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Panum Institute, Copenhagen N, Denmark*
- CAROLINA A. NOVOA • *Terry Fox Laboratory, BC Cancer Agency, Vancouver, Canada*
- GUÐJÓN ÓLAFSSON • *Mitotic Control Laboratory, The Francis Crick Institute, London, UK*
- FRANCISCO M. BASTOS DE OLIVEIRA • *Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Cidade Universitária, Rio de Janeiro, Brazil*
- CLINTON D. OREBAUGH • *Genome Integrity and Structural Biology Laboratory, National Institute for Environmental Health Sciences, National Institute of Health (NIH), Research Triangle Park, NC, USA*
- ROXANNE OSHIDARI • *Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada*
- ÖZGÜN ÖZER • *Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Copenhagen N, Denmark*
- VINCENT PAGÈS • *Cancer Research Center of Marseille, Team DNA Damage Tolerance, CNRS, UMR7258, Marseille, France; Inserm, U1068, Marseille, France; Institut Paoli-Calmettes, Marseille, France; Aix Marseille Univ., UM 105, Marseille, France*
- ACHILLE PELLICOLI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- PAOLO PLEVANI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- ALESSANDRO PODESTÀ • *Dipartimento di Fisica and C.I.Ma.I.Na, Università degli Studi di Milano, Milano, Italy*
- ERICA J. POLLEYS • *Department of Biology, Tufts University, Medford, MA, USA*

- JAMES R. POWELL • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- TERESA M. PRZYTYCKA • *Computational Biology Branch, NCI/NIH, Bethesda, MD, USA*
- KATIE PTASIŃSKA • *University of Sussex, Brighton, UK*
- CHRISTOPHER D. PUTNAM • *Ludwig Institute for Cancer Research, University of California School of Medicine, San Diego, La Jolla, CA, USA; Department of Medicine, University of California School of Medicine, San Diego, La Jolla, CA, USA*
- OLIVER QUEVEDO • *Department of Biology, University of Copenhagen, Copenhagen N, Denmark; Department of Cellular and Molecular Medicine, Center for Chromosome Stability, University of Copenhagen, Copenhagen N, Denmark*
- ELINA A. RADCHENKO • *Department of Biology, Tufts University, Medford, MA, USA*
- CRISTINA RAMOS-PÉREZ • *Donnelly Centre and Department of Biochemistry, University of Toronto, Toronto, ON, Canada*
- MONICA RANSOM • *Department of Biochemistry and Molecular Genetics, Program in Molecular Biology, University of Colorado School of Medicine, Aurora, CO, USA*
- ALEXANDER RAPP • *Department of Biology, Technische Universität Darmstadt, Darmstadt, Germany*
- SIMON H. REED • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- STEFANIA ROMA • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- ADAM P. ROSEBROCK • *The Donnelly Centre, University of Toronto, Toronto, ON, Canada; Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; Department of Pathology and University Cancer Center, Stony Brook Medicine, Stony Brook, NY, USA*
- FABIAN RUDOLF • *Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland*
- XANITA SAAYMAN • *Donnelly Centre and Department of Biochemistry, University of Toronto, Toronto, ON, Canada*
- SARAH A. SABATINOS • *Ryerson University, Toronto, ON, Canada*
- GREGOR W. SCHMIDT • *Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland*
- ETIENNE SCHWOB • *IGMM, CNRS, University of Montpellier, Montpellier, France*
- SARAH SERTIC • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- KATSUHIKO SHIRAHIGE • *Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo, Japan*
- MARCUS B. SMOLKA • *Department of Molecular Biology and Genetics, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, USA*
- ANJANA SRIVATSAN • *Ludwig Institute for Cancer Research, University of California School of Medicine, San Diego, La Jolla, CA, USA*
- PETER C. STIRLING • *Terry Fox Laboratory, BC Cancer Agency, Vancouver, Canada; Department of Medical Genetics, University of British Columbia, Vancouver, Canada*
- JENNIFER A. SURTEES • *Department of Biochemistry, School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA; Genetics, Genomics and Bioinformatics Program, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA*
- YUMIN TENG • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*

- PETER H. THORPE • *Mitotic Control Laboratory, The Francis Crick Institute, London, UK*
- SHYAM TWAYANA • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy*
- MOJCA MATTIAZZI USAJ • *The Donnelly Centre, University of Toronto, Toronto, ON, Canada*
- RICHARD M. WEBSTER • *Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK*
- GREGORY M. WILLIAMS • *Department of Biochemistry, School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA*
- DAMIAN WOJTOWICZ • *Laboratory of Pathology, NCI/NIH, Bethesda, MD, USA*
- CHRISTINE WOLF • *Department of Pediatrics, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany*
- CHUANHE YU • *Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, USA*
- RALPH ZELLWEGER • *Institute of Molecular Cancer Research, University of Zurich, Zurich, Switzerland*
- ZHIGUO ZHANG • *Institute for Cancer Genetics, Columbia University, New York, NY, USA; Department of Pediatrics, Columbia University, New York, NY, USA; Department of Genetics and Development, Columbia University, New York, NY, USA; Irving Cancer Research Center, New York, NY, USA*
- KEJI ZHAO • *Systems Biology Center, Bethesda, MD, USA*

# Chapter 1

## The A-Like Faker Assay for Measuring Yeast Chromosome III Stability

Carolina A. Novoa, J. Sidney Ang, and Peter C. Stirling

### Abstract

The ability to rapidly assess chromosome instability (CIN) has enabled profiling of most yeast genes for potential effects on genome stability. The A-like faker (ALF) assay is one of several qualitative and quantitative marker loss assays that indirectly measure loss or conversion of genetic material using a counterselection step. The ALF assay relies on the ability to count spurious mating events that occur upon loss of the *MAT $\alpha$*  locus of haploid *Saccharomyces cerevisiae* strains. Here, we describe the deployment of the ALF assay for both rapid and simple qualitative, and more in-depth quantitative analysis allowing determination of absolute ALF frequencies.

**Key words** Chromosome instability, Genome instability, Aneuploidy, Gene conversion, Gross chromosomal rearrangement, A-like faker, Marker loss

---

### 1 Introduction

Baker's Yeast (*Saccharomyces cerevisiae*) can divide stably as haploid *MATa* or *MAT $\alpha$*  cells. This phenotype is determined by the composition of the active mating locus *MAT* which carries one of two mating type alleles. Silent copies of *MATa* and *MAT $\alpha$*  also exist at distal positions on the right and left arms of chromosome III and are called *HMR* and *HML*, respectively [1]. In 1981, it was recognized that the transcriptional program driven by *MATa* is constitutive, such that deletions causing loss of the *MAT $\alpha$*  locus lead to a *MATa* phenotype [2]. While the mutant cells were able to mate, they usually did not sporulate normally and thus were not truly switched to normal *MATa* cells. As a result, these *MAT $\alpha$*  mutants were called A-Like Fakers (ALF). This model helped explain how deletion alleles could lead to a mating type switch, including the first of such observations, the so-called Hawthorne Deletion [3].

Subsequently, it was realized that the frequency of ALF in a cell population reflected instability in yeast Chromosome III [4–6]. The outcomes of the ALF phenotype were finally characterized

and applied in large-scale screening less than a decade ago [7]. Events leading to ALF cells include chromosome loss, MAT allele disruption by deletions or unbalanced translocations, and gene conversion from the silent mating type locus HMRA. The latter is likely to be a product of inappropriate homologous recombination between *MAT $\alpha$*  and silent locus HMRA, and the subsequent generation of an active *MAT $\alpha$*  locus [7].

The measurement of relative ALF frequency can be applied to large collections of mutants using a qualitative patch-based assay [7–9], and absolute frequencies can also be quantified using a more demanding assay. The presence of ALF cells is detected by a mating test involving selection of prototrophic mated products after exposure to a *MAT $\alpha$*  tester strain (YPH316, *See* Subheading 2.3). In the protocol described here, *MAT $\alpha$*  locus loss can be followed because *MAT $\alpha$*  strains are deficient in *HIS3* (and carry other auxotrophies) while the *MAT $\alpha$*  tester strain is deficient in *HIS1*; therefore, only diploid cells can synthesize histidine and grow normally on media lacking amino acids (Fig. 1). Virtually any auxotrophic *MAT $\alpha$*  strain can be used for ALF, so long as it contains a functional *HIS1* gene to complement the mating tester. As the mating tester strain is a constant across all control and experimental conditions, any change in the frequency of ALF can be ascribed to the experimental strains under study and compared with an otherwise isogenic control. While here we primarily focus on assessing genomically integrated mutant alleles, ALF can also be conducted on strains bearing plasmids or in the presence of alternative media which we highlight in our notes (e.g., to induce expression, such as with galactose).

---

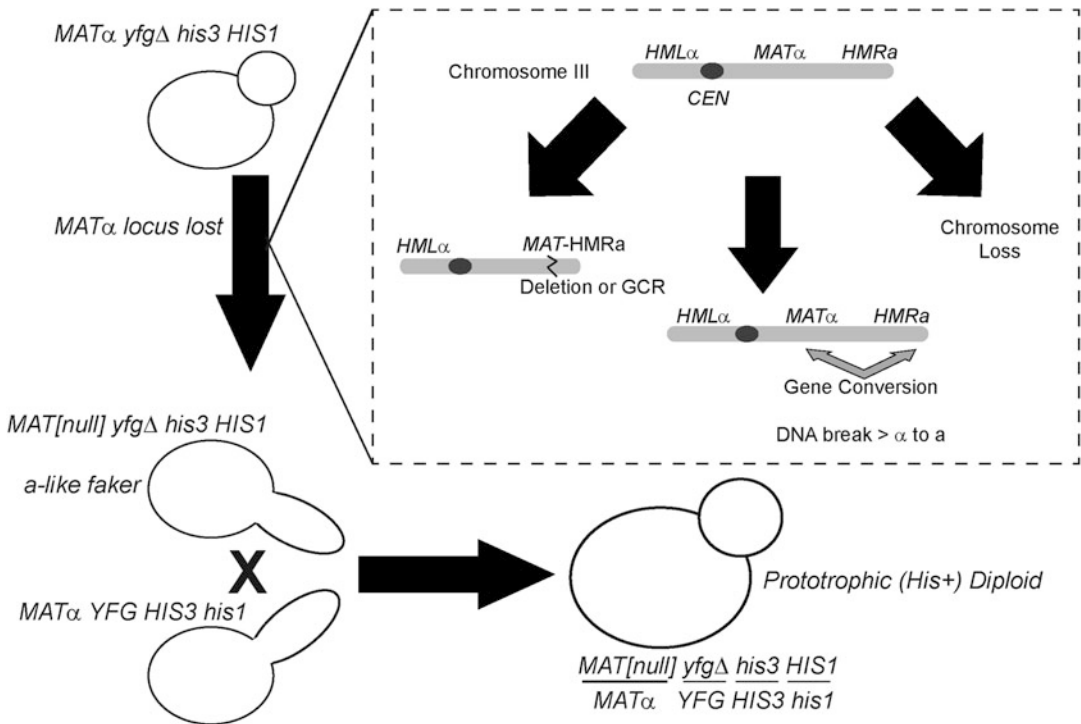
## 2 Materials

### 2.1 Lab Ware

1. Disposable 10 cm plastic petri dishes.
2. Autoclaved toothpicks or sticks.
3. Sterile 1.5 mL microcentrifuge tubes.
4. Sterile velvet squares (~15 × 15 cm).
5. Replica plating block.
6. 10 cm round filter paper discs.

### 2.2 Media Preparation

1. Yeast Peptone Dextrose (YPD) agar plates: 10 g/L yeast extract, 20 g/L peptone, 0.33 g/L tryptophan, 20 g/L dextrose, 20 g/L agar. All solutions dissolved in distilled H<sub>2</sub>O (*see* **Notes 1** and **2**).
  - (a) In separate autoclavable flasks, prepare yeast extract-peptone-tryptophan broth, 4% agar, and 20% dextrose as indicated above. Make sure to use at least one flask with enough volume carrying capacity to hold the desired final volume.



**Fig. 1** Schematic of the ALF assay genetics. A *MAT $\alpha$*  query strain deleted for any gene (*yfg $\Delta$*  = your favorite gene) is grown and at some frequency loses the *MAT $\alpha$*  locus (black arrow at left). A dotted box expands on possible mechanisms of this event which are, from left to right, focal deletion or GCRs that disrupt *MAT*; inappropriate recombination with *HMRa*, leading to gene conversion, and whole chromosome loss. That *MAT[null]* strain is rescued by mating with a tester strain lacking only the *HIS1* gene. Fully prototrophic diploids are selected on minimal media and scored

- (b) Autoclave at 121 °C for at least 20 min, then cool to ~60 °C.
  - (c) While stirring the autoclaved mixture contained in the largest flask, mix in the other two solutions.
  - (d) While still warm, pour 20–25 mL of media into 10 cm plastic petri dishes.
  - (e) Allow agar mixture to solidify and dry overnight on the bench top before using.
  - (f) Unused YPD plates can be stored at 4 °C for several months.
2. YPD liquid media: 10 g/L yeast extract, 20 g/L peptone, 0.32 g/L tryptophan, 20 g/L dextrose. All the solutions dissolved in distilled H<sub>2</sub>O.
    - (a) As for agar plates, prepare and autoclave dextrose solution separately from yeast extract-peptone-tryptophan broth.
    - (b) Combine sterilized mixtures and stir well.
    - (c) Cool prior to use.

3. Synthetic Dextrose agar plates (SD): 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, 20 g/L dextrose, 20 g/L agar. All the solutions dissolved in distilled H<sub>2</sub>O.
  - (a) In separate autoclavable flasks prepare SD broth, 4% agar, and 20% dextrose as indicated above. Make sure to use at least one flask with enough volume carrying capacity to hold the desired final volume.
  - (b) Autoclave each mixture.
  - (c) While stirring the autoclaved mixture contained in the largest flask, mix in the other solutions.
  - (d) While still warm, pour media into 10 cm diameter plastic petri dishes.
  - (e) Allow agar mixture to cool and solidify before using.
  - (f) Unused SD plates can be stored at 4 °C for several months.

### 2.3 Yeast Strains

1. Mating tester strain YPH316 which is a *MAT $\alpha$*  strain bearing an inactivating mutation in *HIS1* but otherwise prototrophic.
2. Positive control for high levels of ALF. There are many possible controls among the hundreds of known ALF-inducing mutants [7, 9]. A BY4742-derived haploid deletion of *BIMI* from the yeast knockout collection [10] serves well as a strong ALF mutant.
3. Negative control for ALF. A wild-type strain representing your strain background of study is the best control. Here, we use BY4742 (*MAT $\alpha$  his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 ura3 $\Delta$ 0*) as our wild-type control.
4. Positive control for mating. A *MAT $\alpha$*  strain can provide an additional control for mating. Here, we use BY4741 (*MAT $\alpha$  his3 $\Delta$ 1 leu2 $\Delta$ 0 met15 $\Delta$ 0 ura3 $\Delta$ 0*).

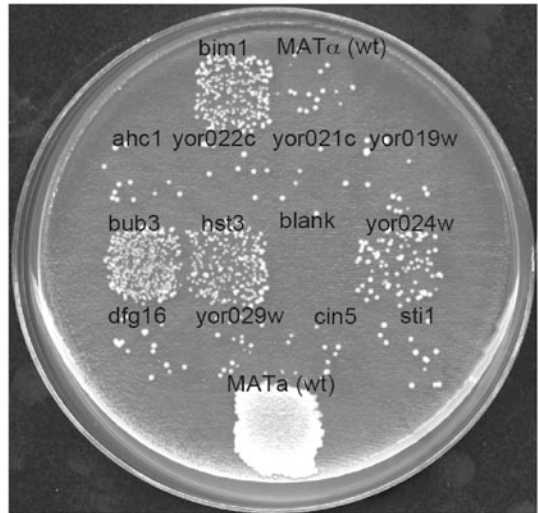
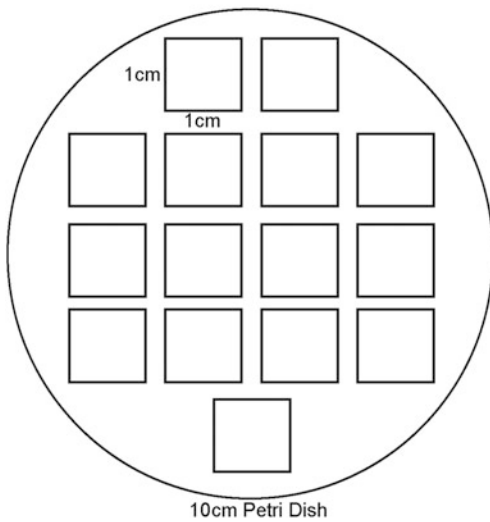
---

## 3 Methods

### 3.1 Qualitative ALF Assay

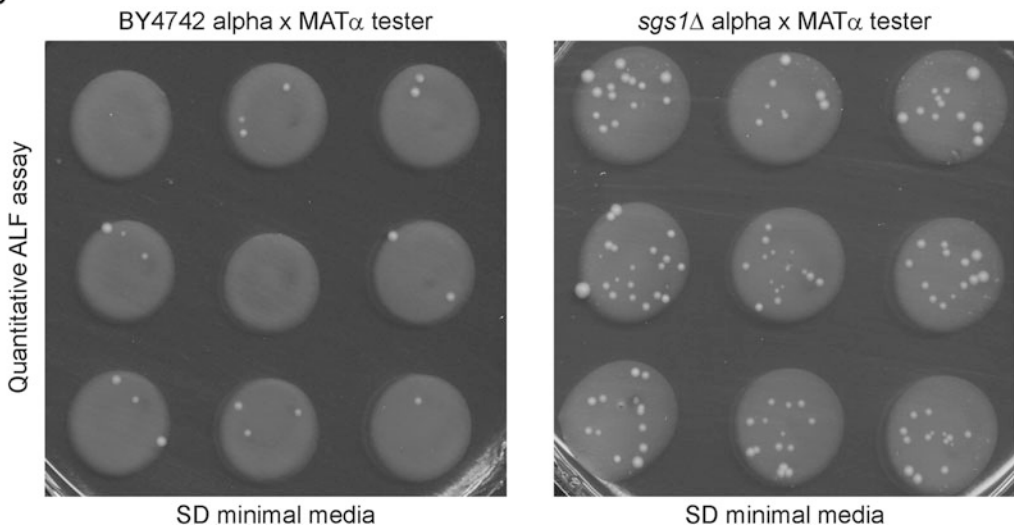
1. Day 1. Suspend  $1 \times 10^6$ – $1 \times 10^7$   $\alpha$ -tester strain cells in water, spread yeast suspension on a YPD plate, and incubate at 30 °C overnight. The plate should be fully covered with a lawn of  $\alpha$ -tester cells the following day.
2. Day 1. On a second YPD plate, using a toothpick or stick, patch *MAT $\alpha$*  haploid strains in an array of approximately 1 cm<sup>2</sup> patches. At least 15 patches can be comfortably fit onto a 10 cm petri dish (Fig. 2). The following controls can be included on each plate: wild-type *MAT $\alpha$*  (BY4742), wild-type *MAT $\alpha$*  (BY4741), and the *MAT $\alpha$  bim1 $\Delta$ ::kanMX* from the *MAT $\alpha$*  deletion collection (or another ALF phenotype positive control). Incubate at 30 °C degrees for 24 h (see Notes 3 and 4).

A



SD minimal media

B



**Fig. 2** Examples of qualitative and quantitative ALF data. (a) Sample plate setup for ALF strain patches in qualitative assay. At least 15  $1 \times 1 \text{ cm}^2$  patches can be fit on a 10 cm petri dish. At right a sample output plate showing the BY4742 and *bim1* $\Delta$  controls on top, along with various positive and negative test strains and the BY4741 control at the bottom. This panel is an excerpt reproduced from Figure S7 in ref. [7]. (b) Output SD plates from the quantitative ALF assay. A known number of wild-type BY4742 (*left*) or *MAT* $\alpha$  *sgs1* $\Delta$  (*right*) cells were mixed with an excess of *MAT* $\alpha$  tester cells and co-cultured on SD media. The *sgs1* $\Delta$  strain exhibits an increased frequency of ALF [7] which can be quantified as described in the text, by comparing to the viability counts from the associated nonselection YPD plate (not shown)

- Day 2. Mate the  $\alpha$ -tester strain with the arrayed patches by replica plating the  $\alpha$ -tester lawn, and the mutant patches, in that order on a fresh YPD plate. Incubate overnight at 30 °C (see Note 5).



4. Day 3. Replica plate the lawn of mated strains onto an SD plate. Incubate at 30 °C.
5. Day 3–5. Growth of diploid cells in colonies should appear within 48–72 h.
6. Day 5. Score the number of colonies on the SD plate for each strain. Comparing the average number of colonies in mutant patches to the number of colonies in the wild-type *MAT $\alpha$*  control patches gives a relative indication of chromosome instability.
7. Optional: retest all probable ALF phenotype hits as described above, but using 4–12 independent colonies to create patches for each mutant. Median values for colony number on the SD plates that are twofold higher than WT are usually considered ALF mutants [7, 9].

### 3.2 Quantitative ALF Assay

1. Day 1. Inoculate 9–12 independent colonies of each *MAT $\alpha$*  strain to be tested in 2–5 mL of liquid YPD broth, and incubate overnight at 30 °C (*see Note 5*).
2. Day 1. In parallel, inoculate YPD broth with 4 mL of  $\alpha$  tester cells per *MAT $\alpha$*  strain of interest, and incubate overnight at 30 °C (*see Note 6*).
3. Day 2. Separate overnight cultures by *MAT $\alpha$*  strain and follow **steps 4–11** for one strain at a time, then repeat for the rest of the mutants.
4. Day 3. Label one YPD and one SD plate with strain name.
5. Day 3. Make 1:100,000 dilutions of the overnight cultures in water (*see Note 7*).
6. Day 3. Dry YPD plate by placing a piece of ~10 cm diameter filter paper on the top of a replica plating block and pressing the surface of the plate against the filter paper. Allow water to wick into the filter paper until it is visibly wet. If necessary repeat with a new filter paper to remove more liquid (*see Note 8*).
7. Day 3. Plate 100  $\mu$ L of each independent colony dilution on the dried YPD plate to determine cell viability counts. Record plating order and orientation on the plate. *See Fig. 2* as reference for spotting position (*see Notes 9 and 10*).
8. Day 3. Dry the SD plate as described in **step 4**.
9. Day 3. Mix 100  $\mu$ L of each undiluted overnight culture with 300  $\mu$ L of overnight culture of the *MAT $\alpha$*  tester in a 1.5 mL microcentrifuge tube.
10. Day 3. Spin at 3000  $\times g$  for 5 min and resuspend each pellet in 100  $\mu$ L of H<sub>2</sub>O.
11. Day 3. Plate the mixtures from **step 9** on the dried SD plate (**step 8**) in the same order and orientation as in the YPD plate.

12. Day 3. Allow all the plates to completely absorb the liquid on the bench top, then flip and incubate at 30 °C for 24 h or 72 h for YPD and SD plates, respectively. After 2 days in culture, colonies of mated yeast should become visible on the SD plate inside the thick white circles left behind by the excess unmated cells (*see Note 11*).
13. Day 4–6. Calculate the frequency of ALF by finding the ratio between mated and plated cells on the SD plates. The number of mated cells is equal to the number of colonies grown on each spot on the SD plate. The number of cells plated on SD can be calculated from the number of viable cells in each spot on the YPD plate, multiplied by the dilution factor (*see Notes 12 and 13*).

---

## 4 Notes

1. For all solid media, autoclave at 121 °C for at least 20 min on a liquid cycle. Cool agar mixture to approximately 55–60 °C, and add other solutions, mix well before pouring 20–25 mL in each petri dish.
2. If a drug needs to be included to maintain a plasmid, ammonium sulfate should be substituted with 1 g/L monosodium glutamate.
3. Patches can grow at any desired temperature if one is looking for the effects of temperature on ALF. Standard conditions are listed for most experiments but can generally be varied.
4. It is advisable to initially fit only 15 patches on one plate to ensure even patch sizes and avoid sample cross-contamination. It can also be helpful to mark the petri dish bottom carefully with a grid and labels to ensure accurate patch size and unambiguous identification of strains to patches at the end of the experiment.
5. The density of cells in each patch should be approximately equal after the growth period. If some strains appear significantly less dense than WT, make a note of it as it will reduce the cell population transferred and thus the perceived rate of ALF colonies. In cases where many strains are sick, an extra 24 h of growth at 30 °C can be useful in allowing saturation of all patches.
6. If many *MAT $\alpha$*  mutants will be tested in one day, a larger volume of  $\alpha$  tester stain can be cultured overnight in the same flask. ~4 mL are required per strain.
7. Making at least two serial dilutions to get to a final 1:100,000 dilution of the overnight culture is recommended to achieve accurate results due to variability in pipetting small volumes.

8. The dryness of the plate will influence the spreading of the 100  $\mu$ L spots. If the plate is too dry, the liquid will not spread enough, making counting of the colonies more difficult. If the plate is not dry enough samples will run into each other or the plates will take too long to absorb in the samples. Approximately 1 g of water can be removed by the filter paper, and overdried plates should be used the same day.
9. It is recommended to fit only nine spots on the plates to avoid samples running into each other. Also, avoid making spots too close to the edges of the plate, as those areas are the hardest to dry with a replica plating tool and might be missed.
10. Using a grid template underneath the YPD and SD plates will facilitate arraying of liquid spots.
11. Some deletion mutants might take longer than 48 h to appear on the SD plate. Colony counts are made after 72 normally but can be extended to 96 h or longer for slow growers.
12. We think of ALF in terms of frequency not rate, as the ALF assay differs from mutation rate fluctuation analyses. This is because each mutation of the *MAT $\alpha$*  locus causing ALF is captured only once and fixed by mating with the mating tester. Thus, there is little possibility of jackpots. A single ALF event usually leads to a single colony and should be expressed as a frequency of marker loss.
13. ALF with other media and plasmids: The ALF assay can also be used to test the effect of plasmids or plasmid borne genes on any *MAT $\alpha$*  strain. The conditions and specific media changes will vary depending on the plasmid system. In our experience, it is important to: (a) maintain selection of the plasmid throughout patch growth, overnight culturing and mating, but it is not necessary on the final SD plate. (b) For galactose inductions the time of galactose exposure and the necessity of including other nonrepressive sugars like raffinose needs to be determined empirically.

---

## Acknowledgments

We thank Philip Hieter for helpful discussions. PCS is a Canadian Institutes of Health Research (CIHR) New Investigator and a Michael Smith Foundation for Health Research Scholar. We acknowledge operating support from CIHR (MOP-136982) and this research is funded by the Canadian Cancer Society (grant 703263).

## References

1. Herskowitz I (1988) Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiol Rev* 52(4):536–553
2. Strathern J, Hicks J, Herskowitz I (1981) Control of cell type in yeast by the mating type locus. The alpha 1-alpha 2 hypothesis. *J Mol Biol* 147(3):357–372
3. Herskowitz I (1988) The Hawthorne deletion twenty-five years later. *Genetics* 120(4):857–861
4. Liras P, McCusker J, Mascioli S, Haber JE (1978) Characterization of a mutation in yeast causing nonrandom chromosome loss during mitosis. *Genetics* 88(4 Pt 1):651–671
5. Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry R, Spencer FA (2004) S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. *Mol Biol Cell* 15(4):1724–1735. doi:[10.1091/mbc.E03-09-0637](https://doi.org/10.1091/mbc.E03-09-0637)
6. Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD (2005) Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* 120(5):587–598. doi:[10.1016/j.cell.2004.12.039](https://doi.org/10.1016/j.cell.2004.12.039)
7. Yuen KW, Warren CD, Chen O, Kwok T, Hieter P, Spencer FA (2007) Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci U S A* 104(10):3925–3930. doi:[10.1073/pnas.0610642104](https://doi.org/10.1073/pnas.0610642104)
8. Ben-Aroya S, Coombes C, Kwok T, O'Donnell KA, Boeke JD, Hieter P (2008) Toward a comprehensive temperature-sensitive mutant repository of the essential genes of *Saccharomyces cerevisiae*. *Mol Cell* 30(2):248–258. doi:[10.1016/j.molcel.2008.02.021](https://doi.org/10.1016/j.molcel.2008.02.021)
9. Stirling PC, Bloom MS, Solanki-Patil T, Smith S, Sipahimalani P, Li Z, Kofoed M, Ben-Aroya S, Myung K, Hieter P (2011) The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. *PLoS Genet* 7(4):e1002057. doi:[10.1371/journal.pgen.1002057](https://doi.org/10.1371/journal.pgen.1002057)
10. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Veronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–906

## The Chromosome Transmission Fidelity Assay for Measuring Chromosome Loss in Yeast

Supipi Duffy and Philip Hieter

### Abstract

The budding yeast *Saccharomyces cerevisiae* has served as an excellent model system for studying highly conserved biological pathways including pathways involved in genome transmission and maintenance. The Chromosome Transmission Fidelity (CTF) colony color assay was developed to assess chromosome instability (CIN) in yeast, by monitoring the loss or gain during cell division of an artificial chromosome fragment carrying a visual marker. The CTF assay monitors changes in chromosome number, allowing the detection of mutants that exhibit increased rates of chromosome nondisjunction or chromosome loss. In this article, we describe the *SUP11*-marker-based CTF assay system, and the methodologies for both qualitative analysis of mutants affecting chromosome transmission, and quantitative analysis for determining the types and rates of errors in chromosome transmission using half-sector analysis.

**Key words** Chromosome instability, Genome instability, Aneuploidy, Whole chromosomal loss, Chromosome transmission fidelity, Marker loss

---

## 1 Introduction

Chromosome replication and segregation during the mitotic cell cycle relies on the correct execution of a complex series of events, with functional determinants that act in *cis* (DNA sequence domains) and in *trans* (gene products). As these processes are essential to cell viability, functional characterization of mutants that are defective in chromosome transmission in higher eukaryotes has been challenging. In the budding yeast *Saccharomyces cerevisiae*, DNA replication and chromosome segregation occur with extremely high fidelity, as errors in chromosome transmission happen very infrequently with rates of chromosome mis-segregation, for example, occurring on the order of once per  $10^5$  cell divisions [1, 2]. However, as this rate of chromosome transmission fidelity in wild-type strains is much greater than the fidelity necessary for cell viability, it is possible to perform genetic analysis with mutants that

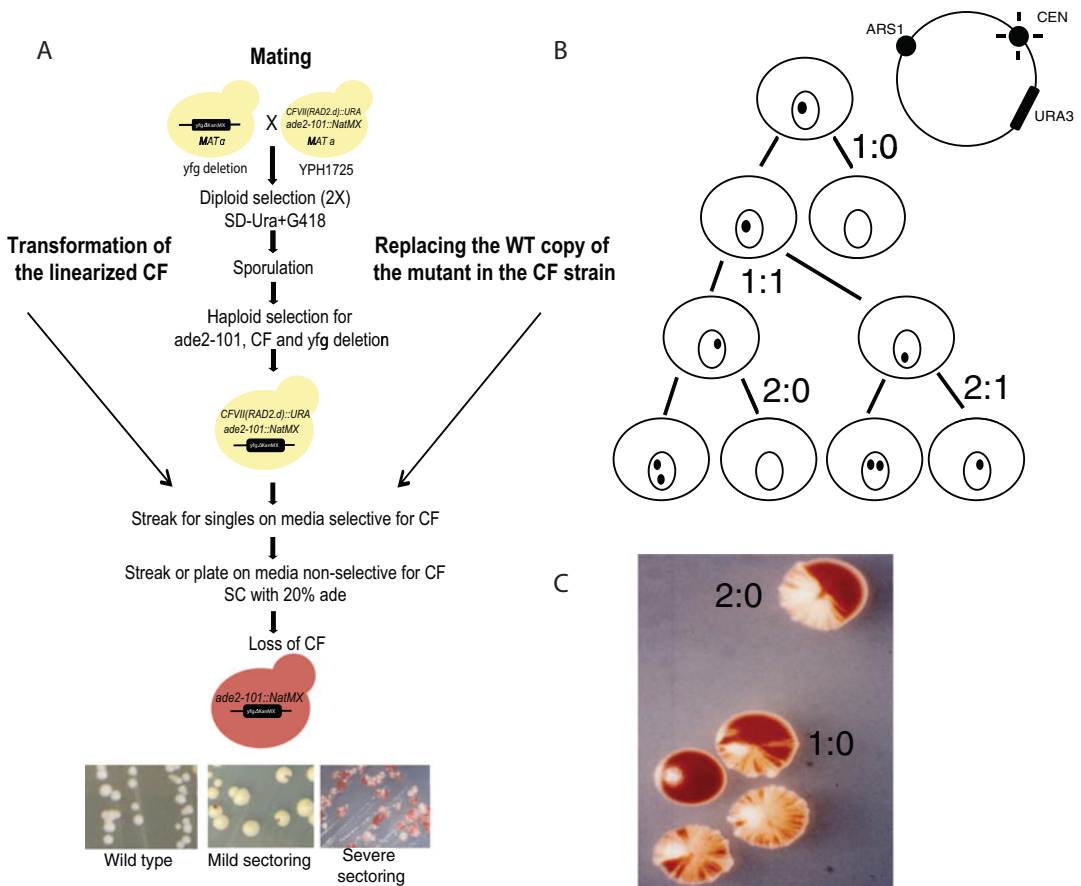
reduce chromosome transmission fidelity over a several order of magnitude range [3, 4].

Taking advantage of the findings that yeast mutants with altered chromosome transmission fidelity remain viable [5] and can tolerate aneuploidy for many individual chromosomes [6], assays were developed for monitoring changes in chromosome ploidy [7, 8]. The two visual chromosome transmission fidelity (CTF) assays, *ade3-2p* and *SUP11*, can be used to detect and analyze mutants involved in chromosome transmission. This chapter will only describe the use of the *SUP11* system for qualitative and quantitative analyses of CTF. The general principles, however, apply directly to the *ade3-2p* system.

Both CTF assays make use of mutations in the purine biosynthesis pathway [9] to facilitate the colony color assay that detects chromosome loss. Wild-type yeast cells are colorless and produce white colonies. Mutations in either the *ADE1* or the *ADE2* genes required for purine biosynthesis cause the accumulation of an intermediate that generates a red hue. The degree of redness in yeast colonies harboring these mutations depends on accumulation of biosynthetic intermediates, which leads to the formation of a red pigment [3]. In the *SUP11* system cells contain an ochre mutation in the *ADE2* gene (*ade2oc*) and are red. This mutation is suppressed by the introduction of a chromosome fragment (CF) that carries the ochre-suppressing form of a tRNA gene, *SUP11* [10–12].

The standard CF used is a 125-kilobase fragment with the short arm containing a selectable marker that can be used during initial transformation and a colony color marker (*SUP11*), which is used in the visual colony assay. The structure and genotype of the CF makes it ideal for monitoring CTF [12, 13]. As the color marker is embedded in pBR322 that bears no homology to yeast DNA on the CF short arm, it will not be lost by mitotic recombination between the CF and endogenous chromosomes [4]. The presence of the CF results in disomy for less than 1% of the yeast genome with no effects on either cell growth or endogenous chromosome fidelity; in wild-type cells, the loss rate of the CF is  $\sim 2$  in  $10^4$  cell divisions [12, 13].

In the CTF assay, a homozygous *ade2*-ochre diploid cell with zero copies of the CF produces no functional Ade2 and thus accumulates sufficient pigment to generate red colonies. A cell with only one copy of the CF will produce small amounts of Ade2, and will only suppress the ochre mutation partially, thus will generate pink colonies. A cell containing two or more copies of the CF that will be colorless will generate white colonies, because enough *ADE2* gene product will be made to prevent the accumulation of the intermediates that produce the red color. Thus, diploid cells with zero, one, and two or more copies of CF will generate red, pink, and white cell lineages respectively and the degree of red



**Fig. 1** The qualitative CTF assay and the half-sectoring assay. (a) Detailed steps involved in generating a CTF strain. For more details on generating a strain by transforming the CF or by introducing the mutant allele to the CF strain, refer to [4]. A starter strain (YPH1725) containing the *ade2-101* mutation and the *SUP11* (*CFVII* (*RAD2.d*))::*URA3*) is mated to a deletion strain of *YFG* and diploids are selected by plating onto appropriate selective media. Following sporulation, a haploid strain is generated where *yfg* $\Delta$ *KAN* is combined with the *ade2-101::NAT* mutation and the CF. Cells are then streaked onto media with the selection for the CF and after 48 h, are either streaked or plated on low adenine media plates. After growing 7–10 days at 25 °C, plates are placed at 4 °C for 2–3 days before scoring. Representative images from qualitative CTF phenotypes are shown at the bottom. (b) Segregation properties of the CF adapted from [3]. The CF includes an ARS element, a CEN element, and a selectable marker in yeast (*URA*). When the CTF strains are placed on nonselective media the pattern of sectoring in each colony reflects the inheritance of the CF. In about 99% of cell divisions, the CF replicates once and partitions equally to daughter cells (1 → 1:1 segregation). Aberrant CF transmission events (1 → 1:0, 1 → 2:0, and 1 → 2:1) are depicted here. Sample images are also included. (c) Samples from an output plate that shows *red*, *pink*, and *white* diploid colonies with 2:0 and 1:0 patterns of CF segregation

sectoring in colonies reflects the frequency of mitotic chromosome loss. In haploid cells, one copy of the CF will cause cells to appear white and a haploid cell that has lost the CF will appear red (Fig. 1a and c). The CTF assay can therefore distinguish between chromosome loss (1:0 segregation) and nondisjunction (2:0 segregation)

in diploid cells (Fig. 1b), which makes it an appropriate method for determining the types of aberrant transmissions that lead to cells with altered chromosome ploidy [7].

The qualitative CTF assay can be easily applied to large collections of loss-of-function (nonessential genes), reduction-of-function (essential genes) mutants and overexpression (essential and nonessential) mutants [14–16], whereas the half-sector analysis for determining the types and rates of chromosome aberrations is more labor intensive. It is also possible to assess CTF rates using fluctuation analysis [13]; however, this chapter will be limited to qualitative CTF and the half-sectoring assay.

To test the effects of a specific genetic perturbation, the mutation will need to be introduced into the CTF background using one of the following methods. Mutations to be tested can be introduced directly into CTF strain backgrounds (YPH1725 and YPH1726, *see* Subheading 2.3) using standard techniques. Alternatively, the CF can be generated *de novo* by transforming mutant strains of interest with the appropriate linearized plasmid [4]. Finally, strains can also be generated using mating and sporulation (Fig. 1a). Following strain generation using one of the methods above, cells are placed on low adenine containing agar media to allow accumulation of intermediates that give rise to red pigment (Fig. 1a). As wild-type cells will generate white colonies, changes in colony sectoring will reflect changes in the rates of chromosome loss and nondisjunction. The protocols described here primarily focus on genomically integrated loss-of-function or reduction-of-function alleles; however, the CTF assay can be conducted in strains with plasmids in the presence of alternate media for inducible expression.

---

## 2 Materials

### 2.1 Lab Ware

1. Disposable 10 cm plastic petri dishes.
2. Autoclaved toothpicks or sticks.
3. Sterile 1.5 mL microcentrifuge tubes.

### 2.2 Media Preparation

1. Synthetic Complete Plates with Limiting Adenine (SC+20% ade): 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, 20 g/L dextrose, 20 g/L agar (*see* Notes 1 and 2).
2. Prepare SC broth, 4% agar, and 20% dextrose in three separate autoclavable flasks/bottles as indicated above. Use at least one flask that will allow the desired final volume of media. Also place a stir bar in this flask as it will aid in the mixing process.
3. Autoclave each mixture.



4. Place the largest container on a stir plate and add the agar and the dextrose.
5. To this add 100 mL/3 L of solution containing the following supplements: adenine, 0.015%; uracil, 0.06%; L-lysine, 0.06%; L-histidine, 0.09%; L-tryptone, 0.09%; L-leucine, 0.06%.
6. Pour ~25 mL of media into 10 cm diameter plastic petri dishes after the media has cooled to a temperature of approximately 55 °C.
7. Allow agar mixture to cool and solidify before using.
8. Unused SC+1/5ade plates can be stored at 4 °C for several months.
9. Synthetic Complete agar plates (SC-URA): 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, 2 g/L—Uracil dropout mix, 20 g/L dextrose, 20 g/L agar (*see* **Notes 3** and **4**).
  - (a) Prepare SC broth, 4% agar, and 20% dextrose in three separate autoclavable flasks/bottles as indicated above. Use at least one flask that will allow the desired final volume of media. Also place a stir bar in this flask, as it will aid in the mixing process.
  - (b) Autoclave each mixture.
  - (c) Place the largest container on a stir plate and add the agar and the dextrose.
  - (d) Pour ~25 mL of media into 10 cm diameter plastic petri dishes after the media has cooled to a temperature of approximately 55 °C.
  - (e) Allow agar mixture to cool and solidify before using.
  - (f) Unused SC-URA plates can be stored at 4 °C for several months.

### 2.3 Yeast Strains

1. The CTF starter strains are YPH1725 or YPH1726, which are *MATa* and *MATalpha* strains with the *ade2-101* ochre mutation marked with NatMX and the CF marked with *URA3* [17]. The NatMX marker allows you to select for the presence of the *ade2-101* marker; therefore, the strains can be used in high-throughput screens. These strains can be used to generate final strains with genes of interest (Fig. 1a and [4]).
2. Positive control for high levels of CTF. These can be found in previously published work from the Hieter lab [12, 16–18]. The *CTF19* mutant strain used in [18] serves well as a control for the half-sectoring assay.
3. Negative control for CTF. Either YPH1725 or YPH1726 without perturbation serves as a good control.

---

## 3 Methods

### 3.1 Qualitative CTF Assay

1. Day 1. On a SC-URA plate, using a toothpick or stick, streak the CTF strains with *YFG* perturbation to achieve single colonies. Include the CTF starter strain and the *CTF19* mutant strains as controls (*see Note 5*).
2. Day 3. Pick two single colonies for each strain with a wooden stick and dilute in 1 mL of H<sub>2</sub>O. Dilute the cells 1:1000 in a second microcentrifuge tube and plate 100  $\mu$ L into SC plates with 20% of the standard adenine concentration to obtain 100–250 colonies per plate (*see Note 6*).

Alternatively, two single colonies for each strain can be streaked onto SC plates with 20% of the standard adenine concentration to achieve single colonies (*see Note 7*).
3. Day 3. Incubate plates at 25 °C for 6–7 days (*see Note 8*).
4. Day 9–10. Incubate plates at 4 °C for additional 5–7 days to enhance the development of red pigment (*see Note 9*).
5. Day 14–17. Score the number of sectored colonies on the plates. The CTF starter strain should have almost no sectored colonies, whereas the *CTF19* mutant will have many sectored colonies giving a relative indication of chromosome instability (*see Note 10*).

### 3.2 Half-Sector Assay

1. Day 1. On a SC-URA plate, using a toothpick or stick, streak out the CTF strains with *YFG* perturbation to achieve single colonies. Include the CTF starter strain and the *CTF19* mutant strains as controls (*see Note 5*).
2. Day 3. Pick 9–12 single colonies for each strain with a wooden stick and dilute in 1 mL of H<sub>2</sub>O. Dilute the cells 1:1000 in a second microcentrifuge tube and plate 100  $\mu$ L into SC plates with 20% of the standard adenine concentration to obtain 100–250 colonies per plate. Plate 100  $\mu$ L into SC plates with 20% of the standard adenine concentration to obtain 100–250 colonies per plate (*see Note 7*).
3. Day 3. Incubate plates at 25 °C for 6–7 days (*see Note 8*).
4. Day 9–10. Incubate plates at 4 °C for additional 5–7 days to enhance the development of red pigment (*see Note 9*).
5. Day 14–17. Score the number of “half-sectored” colonies on the plates to calculate the rate of chromosome loss relative to wild-type cells. Chromosome loss (1:0 segregation, Fig. 1b) will result in pink/red half-sectored colonies, whereas nondisjunction (2:0 segregation) will result in white/red half-sectored colonies (Fig. 1b). Note that colonies with greater than 1/2 red:white sectoring (e.g., 3/4 red and 1/4 white) are scored as half-sectored colonies; this rule was empirically

determined by comparisons to rates determined using fluctuation analysis.

6. The CTF rate is a ratio of “half-sectored” colonies to total CF containing colonies. Colonies that are completely red should be discounted, as these represent cells that did not have the CF when they were plated. For the *CTF19* mutant strain, the rate of chromosome loss should be ~100 times greater than wild-type and the rate of nondisjunction should be 60-fold higher than wild-type [18] (see **Notes 10** and **11**).

---

## 4 Notes

1. For all solid media (>1 L), autoclave at 121 °C for at least 30 min on a liquid cycle. Mix well after adding the solutions together.
2. Standard protocol includes 1/5 the concentration of adenine; however, it may be necessary to adjust this concentration as the red colony phenotype may vary when using different adenine stock solutions or strain backgrounds.
3. If a drug needs to be included to maintain a plasmid, ammonium sulfate should be substituted with 1 g/L monosodium glutamate.
4. To make the—Uracil dropout mix amino acids as follows. 6 g of each of serine, arginine, glycine, glutamic acid, alanine, histidine, glutamine, threonine, asparagine, phenylalanine, methionine, valine, isoleucine, proline, tryptophan, tyrosine, aspartic acid, lysine and cysteine, 12 g of leucine and 1.5 g of adenine hemisulfate.
5. It is important to maintain selection (URA) for the CF until you are ready to begin the CTF experiment, as the loss rate of the CF is higher than for native chromosomes.
6. It is assumed a single colony contains approximately  $10^6$  cells; therefore, 100  $\mu$ L from the diluted microcentrifuge tube represents approximately 100 cells. For slow growing strains it may be necessary to plate more than 100  $\mu$ L to obtain 100–250 cells. In the first pass it is advisable to plate at least two different volumes to get a final cell number of 100–250 cells/plate.
7. While streaking for singles is acceptable it is much harder to control for cell numbers using this method. It is advisable to score at least 100 single colonies during the first pass of a CTF experiment.
8. CTF starter strains can grow at any desired temperature; however, we have observed that slower growth at 25 °C enhances the accumulation of red pigment compared to growth at 30 °C.

9. This extended incubation at 4 °C is especially beneficial for detecting mild CTF phenotypes as it further enhances the red pigment development.
10. With a mis-segregation of  $10^{-4}$  approximately 30,000–100,000 single colonies will need to be counted to generate the rate of chromosome mis-segregation events for wild-type cells. Depending on the degree of the CTF phenotype it may be necessary to score at least 3000 colonies per gene tested. A rule of thumb is to plate sufficient numbers of colonies so that you score >10 independent half-sectoring colonies. Colonies with greater than 1/2 red:white sectoring (e.g., 3/4 red and 1/4 white are scored as half-sectoring) are designated as half sectoring colonies.
11. CTF with other media and plasmids: The CTF assay can also be used to test the effect of plasmids or plasmid borne genes. The conditions and specific media changes will vary depending on the plasmid system. In our experience it is important to: (a) maintain selection of the plasmid throughout growth on plates, but it is not necessary on the final SC+1/5 ade plate. (b) For galactose inductions in our experience it is better to incorporate two rounds of inductions each lasting 48 h of growth.

---

## Acknowledgments

This work was supported by grants from the Canadian Institute of Health Research (MOP 38096) and the National Institute of Health (RO1CA158162) to P.H. S.D. was supported by a Michael Smith Foundation for Health Research Trainee Award. P.H. is a Senior Fellow in the Genetic Networks program at the Canadian Institute for Advanced Research.

## References

1. Esposito MS, Maleas DT, Bjornstad KA et al (1982) Simultaneous detection of changes in chromosome number, gene conversion and intergenic recombination during mitosis of *Saccharomyces cerevisiae*: spontaneous and ultraviolet light induced events. *Curr Genet* 6 (1):5–11. doi:[10.1007/BF00397633](https://doi.org/10.1007/BF00397633)
2. Hartwell LH, Smith D (1985) Altered fidelity of mitotic chromosome transmission in cell cycle mutants of *S. cerevisiae*. *Genetics* 110 (3):381–395
3. Koshland D, Hieter P (1987) Visual assay for chromosome ploidy. *Methods Enzymol* 155:351–372
4. Shero JH, Koval M, Spencer F et al (1991) Analysis of chromosome segregation in *Saccharomyces cerevisiae*. *Methods Enzymol* 194:749–773
5. Maine GT, Sinha P, Tye BK (1984) Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics* 106(3):365–385
6. Parry EM, Cox BS (1970) The tolerance of aneuploidy in yeast. *Genet Res* 16(3):333–340
7. Hieter P, Mann C, Snyder M et al (1985) Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. *Cell* 40(2):381–392

8. Koshland D, Kent JC, Hartwell LH (1985) Genetic analysis of the mitotic transmission of minichromosomes. *Cell* 40(2):393–403
9. Roman H (1956) Studies of gene mutation in *Saccharomyces*. *Cold Spring Harb Symp Quant Biol* 21:175–185
10. Hawthorne DC, Mortimer RK (1968) Genetic mapping of nonsense suppressors in yeast. *Genetics* 60(4):735–742
11. Manney TR (1964) Action of a super-suppressor in yeast in relation to allelic mapping and complementation. *Genetics* 50:109–121
12. Spencer F, Gerring SL, Connelly C et al (1990) Mitotic chromosome transmission fidelity mutants in *Saccharomyces cerevisiae*. *Genetics* 124(2):237–249
13. Hegemann JH, Shero JH, Cottarel G et al (1988) Mutational analysis of centromere DNA from chromosome VI of *Saccharomyces cerevisiae*. *Mol Cell Biol* 8(6):2523–2535
14. Duffy S, Fam HK, Wang YK et al (2016) Over-expression screens identify conserved dosage chromosome instability genes in yeast and human cancer. *Proc Natl Acad Sci U S A* 113(36):9967–9976. doi:[10.1073/pnas.1611839113](https://doi.org/10.1073/pnas.1611839113)
15. Stirling PC, Bloom MS, Solanki-Patil T et al (2011) The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. *PLoS Genet* 7(4):e1002057. doi:[10.1371/journal.pgen.1002057](https://doi.org/10.1371/journal.pgen.1002057)
16. Yuen KW, Warren CD, Chen O et al (2007) Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci U S A* 104(10):3925–3930. doi:[10.1073/pnas.0610642104](https://doi.org/10.1073/pnas.0610642104)
17. Measday V, Baetz K, Guzzo J et al (2005) Systematic yeast synthetic lethal and synthetic dosage lethal screens identify genes required for chromosome segregation. *Proc Natl Acad Sci U S A* 102(39):13956–13961. doi:[10.1073/pnas.0503504102](https://doi.org/10.1073/pnas.0503504102)
18. Hyland KM, Kingsbury J, Koshland D et al (1999) Ctf19p: a novel kinetochore protein in *Saccharomyces cerevisiae* and a potential link between the kinetochore and mitotic spindle. *J Cell Biol* 145(1):15–28

# Chapter 3

## Measuring Mutation Rates Using the Luria-Delbrück Fluctuation Assay

Gregory I. Lang

### Abstract

The Luria-Delbrück fluctuation assay is one of the most commonly used methods for measuring the mutation rate in microorganisms. Specifically, it is used to measure the mutation rate at a particular locus or loci at which mutations give rise to a selectable phenotype. Here, I outline the essential features of performing Luria-Delbrück fluctuation assays as well as common missteps and tips for improving the accuracy of mutation rate estimates. In addition, I provide tools for analyzing data from fluctuation assays. This 96-well plate protocol has been optimized for use in yeast but should perform equally well for a range of microorganisms using standard microbiological methods.

**Key words** Mutation rate, Fluctuation test, Poisson distribution

---

## 1 Introduction

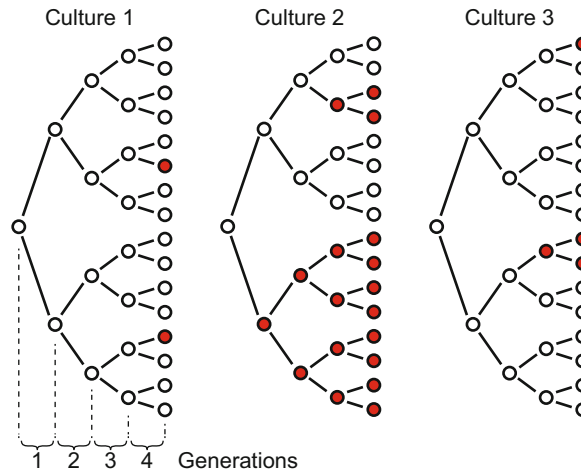
### 1.1 Principle of the Fluctuation Assay

The principle of the fluctuation assay (first introduced by Salvador Luria and Max Delbrück in their classic 1943 *Genetics* paper [1]) is simple, and if implemented properly, provides a powerful way to measure phenotypic mutation rates. The key to understanding the logic of the Luria-Delbrück experiment is to grasp the distinction between the distribution of the number of *mutation events* per culture and the distribution of the number of *mutant cells* per culture. Fig. 1 shows three hypothetical cultures that each start from a single cell and proceed through four generations of growth to produce sixteen cells. In each of these cultures exactly two mutation events occur, but the number of mutant cells differs depending upon when during the growth of the culture those mutations arose.

The distribution of the number of mutation events per culture follows the Poisson distribution and the distribution of the number

---

**Electronic Supplementary Material:** The online version of this chapter (doi: [10.1007/978-1-4939-7306-4\\_3](https://doi.org/10.1007/978-1-4939-7306-4_3)) contains supplementary material, which is available to authorized users.

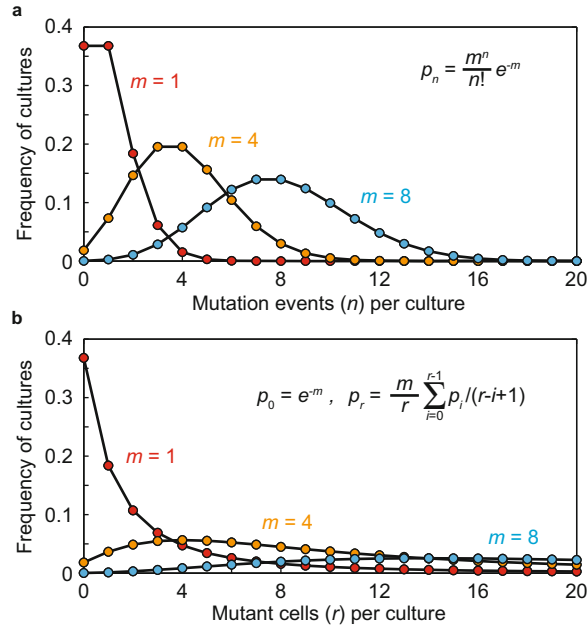


**Fig. 1** Cartoon illustrating the principle of the Luria-Delbrück fluctuation assay. During the growth of a culture the number of mutation events will follow the Poisson distribution; however, the number of mutants per culture will have a larger variance because early arising mutations produce more mutant cells. For example, each of the three cultures had two mutation events during growth; however, the number of mutant cells differs depending upon when during growth the mutations arose

of mutant cells per culture follows the Luria-Delbrück distribution (Fig. 2). Both these distributions are described by a single parameter  $m$ , the expected number of mutation events per culture. Notice that for all values of  $m$ , the zero class ( $p_0$ ) is the same for both the Poisson and the Luria-Delbrück distributions. This is because a culture will have zero mutant cells if, and only if, zero mutation events occurred. A single mutation event could produce just one mutant cell (if it arose in the last generation) but could also lead to a “jackpot” of mutant cells if it arose early in the growth of the culture. These “jackpots” are rare. Their rarity can be explained by looking retrospectively at the growth of a culture, recognizing that half of the mutation events occurred in the last generation, one quarter in the generation previous to that, etc. The “jackpot” principle was the key observation in the Luria Delbrück 1943 *Genetics* paper that proved that bacterial resistance to bacteriophage occurred through genetic mutation and not to acquired immunity following exposure to the bacteriophage at the time of plating.

## 1.2 Assumptions of the Luria-Delbrück Distribution

The formulation of the Luria-Delbrück distribution in Fig. 2 was derived by Ma, Sandri, and Sarkar [2]. In order for the number of mutant cells per culture to follow this distribution, several biological assumptions must be met. When performing fluctuation assays it is important to be aware of these assumptions because deviations could affect the accuracy of mutation rate estimates.



**Fig. 2** Distributions for the number of mutation events per culture and the number of mutant cells per culture for three values of the parameter  $m$ , the expected number of mutation events per culture. **(a)** The distribution of the number of mutation events per culture follows the Poisson distribution with an average number of mutation events per culture equal to  $m$ . **(b)** The distribution of mutant cells per culture follows the Luria-Delbrück distribution. Note that  $p_0$  is the same for both distributions; in other words, a culture will have no mutant cells if, and only if, there were no mutation events

I have listed these assumptions in the order in which they are likely to be violated in a typical fluctuation assay.

1. *Growth of each culture starts with a single cell.* This assumption will almost always be violated, but as long as the initial inoculum size is negligible relative to the final number of cells per culture, this is not a concern. This is because the number of cell divisions is equal to the number final cell number minus the inoculum size.
2. *Mutant and non-mutant cells have the same growth rate.* This will skew the distribution of the number of mutant cells that result from each mutation event. This will have less of an effect on mutation rate estimates using the  $p_0$  method than the MSS-maximum-likelihood method (*see* Subheading 4), since differential growth will not affect the zero class.
3. *No post-planting mutations.* In the case where the selection does not effectively kill or arrest non-mutant cells, additional cell divisions will produce additional mutant cells. These post-planting mutants will be Poisson-distributed.



4. *All mutants are detected.* If plating efficiency is less than 100% or if there is protein perdurance (for instance in the case of counter-selectable markers such as *CAN1* or *URA3*), the number of true mutant cells will be undercounted.
5. *No change in mutation rate during growth.* Changes in mutation rate in response to changes in growth conditions will skew the distribution of mutant cells per culture. An elevated mutation rate early would lead to more “jackpot” cultures; elevated mutation rate late in growth would increase the proportion of the lower classes of the distribution.
6. *Same final number of cells in replicate populations.* This is typically a safe assumption, though complications with the experimental setup (such as uneven evaporation across the 96-well plate) could result in this assumption being violated.
7. *No reverse mutations.* Under most circumstances, this is a safe assumption. Though it is possible, for example if a large class of suppressor mutations exists for the phenotype being assayed, that this assumption will not always hold.

### **1.3 Common Missteps in Performing Fluctuation Assays**

From my experience performing fluctuation assays and advising others on this method, I have come to recognize several common missteps that should be avoided.

1. *Not plating the whole culture.* The principle of the Luria-Delbrück fluctuation is based on accurately measuring the distribution of the numbers of mutant cells per culture. This is best determined by plating the entirety of the culture. For this reason, culture volumes should typically be kept small. If the entire volume of the culture is not plated, it is necessary to correct for this by adjusting for the dilution factor before calculating  $m$  from the data. When you have a large number of mutant cells per culture and subsampling is required, it becomes more difficult to determine how well (or how poorly) the data fit to the Luria-Delbrück distribution, which flattens out at high  $m$ .
2. *Not having a zero class.* There are two reasons why I think it is important to have a zero class. First, it allows the  $p_0$  method to be used to calculate mutation rate in addition to the MSS-maximum-likelihood (*see* Subheading 4). Second, the maximum likelihood fitting method is most sensitive to the lower end of the distribution of the data.
3. *Using too few populations.* There is no specific number of populations required, but the accuracy of the fluctuation assays (if performed properly) increases as the square root of the number of cultures [3]. Fig. 1 in [4] can be used as a guide for selecting an appropriate number. I routinely use 72 cultures to estimate  $m$ , though I have used as little as 36 and as many as 720.

---

## 2 Materials

1. 96-well plates.
2. Aluminum plate seals.
3. Whatman filter paper circles (Grade 3, Cat. No. 1003-090).
4. Replica-plating block.
5. Replica-plating velvets, sterile.
6. Liquid nonselective growth medium (*see Note 1*).
7. Agar plates containing selective growth medium (*see Note 2*).
8. Beckman Coulter particle counter (*see Note 3*).
9. Vials for Beckman Coulter particle counter.
10. ISOTON II Diluent for Beckman Coulter particle counter.
11. Sonicator.

---

## 3 Performing Fluctuation Assays

### 3.1 *Overdrying Plates*

1. Set up a replica-plating block with a sterile velvet.
2. Using sterilized forceps, place a sterile 90 mm Whatman filter paper circle on the velvet (*see Note 4*).
3. Press plate onto the replica-plating block to transfer the filter to the plate (*see Note 5*).
4. Allow the filter to sit on the plate for at least 30 min, up to overnight, to allow the filter to pull water out of the plate.
5. Repeat for each plate.

### 3.2 *Set Up 96 Parallel Cultures*

1. Grow a single overnight culture of the strain to be tested (*see Note 6*).
2. Dilute culture 1:10,000 into 40 ml of nonselective (*see Note 7*).
3. Add the appropriate volume of culture to each well of a 96-well plate (*see Note 8*).
4. Cover the plate with a plate seal to avoid evaporation.
5. Incubate the 96-well plate at 30 °C (or other appropriate temperature) until cultures reach saturation (*see Note 9*).

### 3.3 *Determining the Average Number of Cells Per Culture (N)*

1. For each 96-well plate, pool 24 of the cultures, and set the remaining 72 cultures aside to be plated on selective medium (*see Note 10*).
2. If necessary, sonicate the pooled cultures to break up cell clumps for 1–2 min in a bath sonicator or for 10 half-second pulses using a probe sonicator.
3. Make three independent dilutions of each culture into filtered ISOTON II Diluent (*see Note 11*).

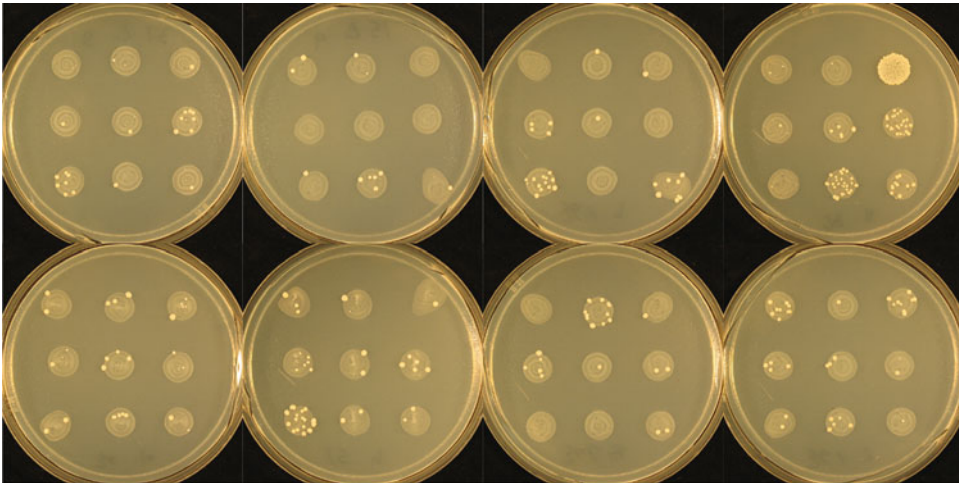
4. Count each vial three times (*see Note 12*).
5. Calculate the average number of cells per culture ( $N$ ) as the median-of-the-median cell counts, correcting for the initial dilution and culture volume (*see Note 13*).

Example data:

	Count 1	Count 2	Count 3	Median
Dilution 1	33,284	32,912	33,518	33,284
Dilution 2	33,268	33,126	32,766	33,126
Dilution 3	33,972	33,578	33,440	33,578
Median				33,284

### 3.4 Determining the Number of Mutant Cells Per Culture

1. For the remaining 72 cultures, bring the volume up to 100  $\mu$ l by adding water to each well using a multichannel pipette.
2. Plate the entire volume of each of the 72 cultures onto the overdried plates by spot-plating nine cultures per plate (*see Fig. 3 and Note 14*).
3. Allow the plates to sit at room temperature for at least a half hour or until all of the liquid is absorbed (*see Note 15*).
4. Incubate plates for 1–2 days at 30 °C until colonies are large enough to count.
5. Count the number of mutant cells per culture (*see Note 16*).



**Fig. 3** Spot plating cultures onto overdried plates. Example of a fluctuation assay on 10 $\times$  canavanine medium from ref. [7]. Seventy-two 100  $\mu$ l cultures were spot-plated onto eight canavanine plates. Colonies were counted after 2 days of growth. For larger culture volumes ( $\sim$ 200  $\mu$ l), only six cultures are spot plated onto each plate. For small volumes, cultures should be brought up to 100  $\mu$ l.

## 4 Analyzing Data from Fluctuation Assays

Many methods exist to calculate  $m$  (the expected number of mutation events per culture) from fluctuation data (reviewed in [4, 5]). In practice, only two methods should be used: the  $p_0$  method and the MSS-maximum likelihood method.

### 4.1 Calculate the Mutation Rate Using the $p_0$ Method

1. Determine the fraction of cultures with zero mutant cells. This value is  $p_0$ .
2. Calculate the expected number of mutation events per culture ( $m$ ):

$$m = -\ln(p_0) \text{ (see Note 17).}$$

3. Calculate mutation rate ( $\mu$ ):

$$\mu = \frac{m}{N}.$$

### 4.2 Calculate the Mutation Rate Using the MSS-Maximum Likelihood Method

In 1992, Ma, Sandri, and Sarkar provided a solution to the Luria-Delbrück distribution for the single parameter  $m$  [2]. This made it possible to calculate the most likely value of  $m$  based on fluctuation assay data [6]. Stewart [3] shows that this method, known as the MSS-maximum-likelihood method, is the most accurate method for determining  $m$ , and provides a formula for calculating 95% confidence intervals for estimates of  $m$  calculated using this method.

1. Calculate the expected number of mutation events per culture ( $m$ ) using the provided supplemental Matlab script “findMLm” (see Note 18) or Excel spreadsheet “fluctuationAssay\_TEMPLATE.xls” (see Note 19).
2. Calculate 95% confidence intervals on  $m$  using the following equations:

$$m_{95-} = \ln(m) - 1.96\sigma(e^{1.96\sigma})^{0.315}, \text{ and}$$

$$m_{95+} = \ln(m) + 1.96\sigma(e^{1.96\sigma})^{-0.315},$$

where  $\sigma = \frac{1.255m^{-0.315}}{\sqrt{C}}$ , and  $C$  is the number of cultures used (see Note 20).

3. Calculate mutation rate ( $\mu$ ):

$$\mu = \frac{m}{N}, \text{ where } N \text{ is the average number of cells per culture.}$$

4. Calculate 95% confidence intervals for the mutation rate:

$$\mu_{95-} = \frac{m_{95-}}{N}, \text{ and}$$

$$\mu_{95+} = \frac{m_{95+}}{N}.$$

## 5 Notes

1. This is typically a rich medium or a synthetic complete medium, for example YPD or SC for yeast.

2. The exact medium used will depend on the organism and the particular locus or loci used for the fluctuation assay. For example, in yeast, 5FOA (1 mg/ml) and canavanine (0.6 mg/ml) is used to measure mutation rates at *URA3* and *CAN1*, respectively. In the case of canavanine, it is necessary to use higher drug concentrations than is necessary for counter-selection to prevent post-plating growth, typically 60  $\mu\text{g}/\text{ml}$  [7].
3. Cells counts can also be performed using other methods such as a hemocytometer or dilution plating.
4. It is possible to sterilize the Whatman filters, but I have not found this to be necessary.
5. Press evenly so that the entire surface of the plate is in contact with the filter. The 90 mm filter circles will be slightly larger on the surface of the agar, which has diameter of  $\sim 85$  mm for standard petri dishes.
6. This can be done in a nonselective medium or in a medium that selects against the mutants. For example, when selecting for 5FOA resistance, which occurs primarily through mutations at the *URA3* locus, I perform this overnight growth in medium lacking uracil.
7. Before beginning a set of experiments it is useful to pilot the experiment to identify the optimal culture conditions. To analyze the data using the  $p_0$  method,  $\sim 20\text{--}80\%$  of the cultures should have zero mutation events, thus zero mutants. An estimate of the mutation rate will help to select a culture volume to pilot. For example, when using 5FOA resistance ( $\mu \sim 5 \times 10^{-8}$  per generation) as a selection, I typically use 200  $\mu\text{l}$  cultures with 2% glucose; when using  $\alpha\text{F}$  resistance ( $\mu \sim 5 \times 10^{-6}$  per generation) as a selection, I typically use 10  $\mu\text{l}$  cultures with only 0.2% glucose.
8. It saves time to use a multichannel pipette for this step.
9. It is not necessary to shake the plates. In fact, shaking could increase evaporation or introduce inconsistencies between wells. Side-by-side measurements of mutation rate without shaking and with shaking on a Titramax 1000 orbital shaker were indistinguishable.
10. I pool 24 wells along the diagonals, rather than two rows, to sample wells along the edge and in the center to control for possible edge effects such as uneven evaporation. This can be done without changing pipette tips. It is important that cells be resuspended well either by pipetting up and down several times or using a plate vortex such as the Titramax 1000 orbital shaker.
11. Filter ISOTON II Diluent with a 0.45  $\mu\text{m}$  filter to remove particles. Use an automatic dispenser for the solution to

maintain uniformity in volume. After dispensing the diluent, let the vials sit for a few minutes before counting to avoid counting bubbles. I aim to dilute such that I count in the range of 30,000–90,000 events (For example, I dilute a  $10^7$  culture 1:2000, 10  $\mu$ l into 20 ml).

12. Before beginning rinse and flush until the background is below 500 (ideally below 100, but I tolerate up to 1000. Sometimes this takes a while). Record the background before and after each run. It is not necessary to rinse and flush in between samples. If anything lodges in the aperture, rinse and flush, noting when this occurs. If the counts spike, stop and restart the count.
13. I do not subtract background. I use medians to ensure that these estimates are robust to non-normally distributed data.
14. If the volume is greater than 100  $\mu$ l, it may not be possible to spot-plate nine cultures per plate. For yeast fluctuation assays on 5FOA, I use 200  $\mu$ l cultures and spot-plate six cultures per plate. It is important that cells be resuspended well prior to plating. This can be accomplished by pipetting up and down several times. I have also used a Titramax 1000 orbital shaker to resuspend cells prior to plating.
15. Be careful not to disturb the plates. I usually allow an hour, but I have allowed plates to dry overnight before transferring them to the incubator.
16. It is possible to count unaided by eye; however, I prefer to use a dissection scope (10 $\times$  magnification) and a gooseneck light source. I use threshold counting. For example, for fluctuation assays on canavanine, colonies smaller than 1 mm at 10 $\times$  magnification are presumed to result from mutations that had occurred after the cells were plated and were counted separately.
17. This equation is derived from the Poisson distribution. Specifically, the probability of observing zero mutant cells (and thus, zero mutation events) is  $p_0 = e^{-m}$ .
18. The most likely value for  $m$  given fluctuation assay data can be calculated by executing the following Matlab script: **findMLm** (*data*), where *data* is a column vector containing counts of mutant cells per culture. This program requires the following Matlab scripts to run: **scoreData** and **generateLD**. **scoreData** is a Matlab script that takes as its input data from a fluctuation assay and a value for  $m$ . This program outputs the *-log* probability of observing the *data* given  $m$ . This program requires the Matlab script **generateLD**. **generateLD** takes as its input a value of  $m$  and a maximum number of mutant cells per culture (*max*) for which to calculate the probability distribution. It outputs the Luria-Delbrück distribution from 0 to *max* with

parameter  $m$ . All three scripts are provided as supplemental information and were originally published in ref. [7]. Alternatively, several web-based tools exist for analyzing fluctuation assays [8, 9].

19. Open the “fluctuationAssay\_TEMPLATE.xls” spreadsheet, provided as supplemental information. Remove the example data in blue and enter your data here. These data will be in the form of the number of cultures with a given number of mutant cells. Check that the number “# of cult.” value is the same as the number you expect. Select a value for  $m$ . The objective is to find the value for  $m$  that maximizes the value for “ $-\log P(\text{data} | m)$ ”. Using the workspace provided pick several values of  $m$ : I typically start with values of  $m$  such as 0.1, 1, 2, 3, 4. Find the minimum value. Then determine if the actual minimum is greater or less than this value. For instance if 2 is the best value of  $m$  out of 0.1, 1, 2, and 3, I next check if 1.9 or 2.1 improves the fit. If  $m$  of 2.1 is better than 2.0, I would then test all values between 2 and 3 in increments of 0.1. I repeat this process until I find the best most likely value of  $m$  given the data to three significant digits.
20. The expected error of fluctuation assays was determined by Stewart [3] and assumes that the data truly fit the Luria-Delbrück distribution. If the quality of the data is poor, calculated values for  $m$  and the 95% confidence intervals on this value will be unreliable. For this reason it is useful to compare values for the mutation rate as determined by the P0 method and the MSS-maximum likelihood method. When performing fluctuation assays one needs to be cognizant of the assumptions that underlie Luria-Delbrück distribution and how violations of these assumptions could affect estimates of mutation rates. Some efforts have been made to account for deviations such as post-plating growth [7], plating efficiency [10], and differences in the growth rate of mutant and non-mutant cells [8, 11, 12].

---

## Acknowledgments

I thank Sean Buskirk, Katie Fisher, and Dan Marad for comments on this manuscript. This work was supported by a New Investigator grant from the Charles E. Kaufman Foundation of The Pittsburgh Foundation.

## References

1. Luria S, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28:491–511
2. Ma WT, Sandri GH, Sarkar S (1992) Analysis of the Luria-Delbrück distribution using discrete convolution powers. *J App Prob* 29:255–267

3. Stewart FM (1994) Fluctuation tests: how reliable are the estimates of mutation rates? *Genetics* 137(4):1139–1146
4. Rosche WA, Foster PL (2000) Determining mutation rates in bacterial populations. *Methods* 20(1):4–17
5. Foster PL (2006) Methods for determining spontaneous mutation rates. *Methods Enzymol* 409:195–213
6. Sarkar S, Ma WT, Sandri GH (1992) On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. *Genetica* 85(2):173–179
7. Lang GI, Murray AW (2008) Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 178(1):67–82
8. Gillet-Markowska A, Louvel G, Fischer G (2015) bz-rates: a web tool to estimate mutation rates from fluctuation analysis. *G3 (Bethesda)* 5(11):2323–2327
9. Hall BM et al (2009) Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbrück fluctuation analysis. *Bioinformatics* 25(12):1564–1565
10. Stewart FM (1991) Fluctuation analysis: the effect of plating efficiency. *Genetica* 84(1):51–55
11. Stewart FM, Gordon DM, Levin BR (1990) Fluctuation analysis: the probability distribution of the number of mutants under different conditions. *Genetics* 124(1):175–185
12. Zheng Q (2005) New algorithms for Luria-Delbrück fluctuation analysis. *Math Biosci* 196(2):198–214



## Molecular Genetic Characterization of Mutagenesis Using a Highly Sensitive Single-Stranded DNA Reporter System in Budding Yeast

Kin Chan

### Abstract

Mutations are permanent alterations to the coding content of DNA. They are starting material for the Darwinian evolution of species by natural selection, which has yielded an amazing diversity of life on Earth. Mutations can also be the fundamental basis of serious human maladies, most notably cancers. In this chapter, I describe a highly sensitive reporter system for the molecular genetic analysis of mutagenesis, featuring controlled generation of long stretches of single-stranded DNA in budding yeast cells. This system is ~100- to ~1000-fold more susceptible to mutation than conventional double-stranded DNA reporters, and is well suited for generating large mutational datasets to investigate the properties of mutagens.

**Key words** Mutagenesis, Genomic instability, DNA damage, Mutation clusters, Localized hypermutation, Translesion DNA synthesis

---

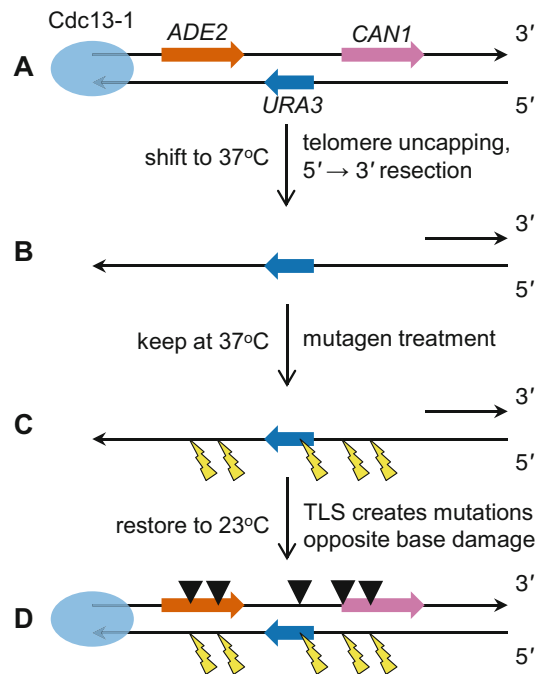
### 1 Introduction

Mutations are an important form of genetic diversity within populations of biological entities [1]. Mutations can be the end result of biochemical or physical processes that damage the nitrogenous bases in DNA. Such processes can be endogenous to cells, including spontaneous hydrolytic, oxidative, or alkylation damage [2]. Base damage can also be induced by exogenous mutagens, including chemical species [3] and high-energy radiation [4]. Model systems have been used extensively to investigate spontaneous and induced mutagenesis, yielding important insights into these processes ([5] and references therein).

Conventional systems for studying mutagenesis typically utilize a reporter gene, such that mutational inactivation or reversion of the reporter can be detected using selection media [5]. In these systems, the reporter gene can be assumed to exist in the canonical double-stranded DNA (dsDNA) state most of the time, as the

DNA should be transiently single stranded (i.e., existing as ssDNA) only during transcription, replication, or repair. dsDNA is less prone to damage than ssDNA, since the nitrogenous bases are less solvent accessible in the former state [2]. Additionally, dsDNA is less susceptible to mutation even if base damage does occur, since an undamaged strand can serve as the template for the repair of the damaged strand [6]. As a result, conventional dsDNA reporter systems can be used to elucidate the mutagenic properties of relatively strong mutagens, but can yield ambiguous results for weakly acting mutagens (*see also Discussion in [7]*).

Curiously, a number of weakly acting mutagens are also suspected carcinogens [8]. But the weak mutagenic signal induced by these agents is difficult to distinguish above the background of spontaneous mutagenesis using conventional dsDNA systems. To obtain robust data on the mutagenic properties of weakly acting agents, a much more sensitive reporter system created by the Gordenin group at the U.S. National Institute of Environmental Health Sciences has been used to good effect ([7] and *see Fig. 1*).



**Fig. 1** (a) Three reporter genes (*CAN1*, *URA3*, and *ADE2*) are embedded near the left telomere of Chromosome V in a *cdc13-1* haploid. (b) Shifting to 37 °C causes telomere uncapping, and ensuing 5' to 3' resection generates a long single-stranded overhang encompassing the reporter cassette. (c) Treatment with a mutagen creates base damage in the 3' overhang. (d) Restoration to 23 °C enables resynthesis of the resected strand. Error-prone translesion DNA synthesis (TLS) creates mutations opposite the base damage

Three reporter genes (*CANI*, *URA3*, and *ADE2*) were deleted from their native loci and reintroduced near a de novo telomere on the left arm of Chromosome V (*see* Fig. 1a). Due to the *cdc13-1* temperature sensitive mutation, the proteinaceous capping complex at the telomeres dissociates when cells are shifted to 37 °C (*see* Fig. 1b). The exposed telomere end is then enzymatically resected in the 5' to 3' direction [9]. This resection generates a long 3' overhang that encompasses the three reporter genes. The exposed ssDNA triggers the DNA damage checkpoint, arresting cells in the G<sub>2</sub> phase of the cell cycle [10]. When these cells are treated with a mutagen, base damage is induced within the 3' ssDNA overhang (*see* Fig. 1c). Upon restoration to permissive temperature (23 °C), the DNA is restored to a double-stranded state but in the process, specialized error-prone translesion DNA synthesis (TLS) polymerases create mutations opposite to the sites of base damage (*see* Fig. 1d). By plating the cells on selection media, it is possible to identify clusters of closely spaced mutations which inactivate multiple reporter genes. These mutation clusters are a form of localized hypermutation and are reminiscent of similar strand-biased clusters of mutations observed, e.g., in cancer genomes [11, 12]. The current generation reporter system is a considerable improvement over earlier, less sophisticated versions, which had been used to study clustered mutagenesis [13, 14].

The ssDNA system has been used in multiple molecular genetic studies of mutagenesis by chemical mutagens [7, 15] and cytidine deaminases [7, 16, 17]. ssDNA is some 100- to 1,000-fold more susceptible to mutation than dsDNA. As a result, the ssDNA reporter system is sufficiently sensitive to detect mutation clusters induced by weak mutagens that are essentially inert toward dsDNA [7]. Furthermore, the genetic requirements for mutagenic TLS bypass opposite to abasic sites in the ssDNA system are consistent with the consensus view in the field [16]. Since abasic sites are well-studied mutagenic lesions, this validation lends additional confidence that the ssDNA system can be used to study mutagenic TLS polymerase-mediated bypass of other kinds of base damage as well. A high frequency of selectable mutation clusters is a reliable proxy for high mutation counts elsewhere in the genome, especially within *unselected* mutation clusters near other telomeres. For example, overexpression of the human APOBEC3B cytidine deaminase generated an average of >400 unique mutations per sequenced yeast genome of ~12 megabases [17]. Also note that due to the known strandedness of the 3' ssDNA overhang, the exact identity of base substitutions can be assigned. This is an additional advantage over conventional dsDNA reporter systems, which cannot distinguish the specific substitution type that is actually induced by a mutagen (e.g., distinguishing C to A from G to T).

---

## 2 Materials

### 2.1 Day 1: Yeast Cultures

1. ySR127 ssDNA reporter yeast. This is isogenic to CG379 [18], with the following genotype: *MAT $\alpha$  his7-2 leu2-3,112 trp1-289 cdc13-1*. The reporter gene cassette (*lys2::CAN1-URA3-ADE2*) is situated ~3 kilobases from the left telomere of Chromosome V.
2. yKC023 dsDNA reporter control yeast (optional). The reporter gene cassette is in Chromosome II, 345 kilobases from the right telomere and 232 kilobases from the centromere.
3. YPDA liquid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g D-glucose, 20 g peptone, 10 g yeast extract, and deionized water to 1 L total volume. Supplement with 2 mL of 0.5% adenine sulfate solution (filter sterilized).
4. 75 mL tissue culture flasks. Microbiological culture tubes or 50 mL Falcon tubes will do as well.
5. Refrigerated incubating shaker.

### 2.2 Day 4: Mutagen Treatment

1. 10 cm petri dishes, each filled with 30–35 mL of media.
2. NaOH solution ( $\geq 5$  N).
3. Synthetic Complete solid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g D-glucose, 5 g ammonium sulfate, 1.7 g yeast nitrogen base without amino acids or ammonium sulfate, 60 mg adenine sulfate, 50 mg L-arginine HCl, 75 mg L-aspartic acid, 100 mg L-glutamic acid, 20 mg L-histidine HCl, 50 mg L-isoleucine, 100 mg L-leucine, 120 mg L-lysine HCl, 20 mg L-methionine, 50 mg L-phenylalanine, 375 mg L-serine, 100 mg L-threonine, 50 mg L-tryptophan, 50 mg L-tyrosine, 150 mg L-valine, 60 mg uracil, and deionized water to 1 L total volume. Adjust pH to 5.8 before autoclaving. Dispense into 10 cm petri dishes.
4. Canavanine, Low Adenine solid media (*see Note 1*): For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g D-glucose, 5 g ammonium sulfate, 1.7 g yeast nitrogen base without amino acids or ammonium sulfate, 20 mg adenine sulfate, 75 mg L-aspartic acid, 100 mg L-glutamic acid, 20 mg L-histidine HCl, 50 mg L-isoleucine, 100 mg L-leucine, 120 mg L-lysine HCl, 20 mg L-methionine, 50 mg L-phenylalanine, 375 mg L-serine, 100 mg L-threonine, 50 mg L-tryptophan, 50 mg L-tyrosine, 150 mg L-valine, 60 mg uracil, and deionized water to 1 L total volume. After cooling media to  $\leq 60$  °C, add 6 mL of 1% canavanine sulfate solution (filter sterilized), stir thoroughly to mix, then pour plates. Adjust pH to 5.8 before autoclaving. Dispense into 10 cm petri dishes.

5. Sterile water.
6. Cell spreader, metal or glass.
7. 95% ethanol.
8. Bunsen burner.
9. Turntable for spreading cells (optional).
10. Sterile glass beads (optional).
11. Incubator set to 23 °C.
12. Hemocytometer (optional).
13. Microscope (optional).

**2.3 Day  $\geq$  9: Colony  
Counting and  
Phenotype Verification**

1. Colony counter, automated or manual.
2. YPG solid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g glycerol, 20 g peptone, 10 g yeast extract, and deionized water to 1 L total volume. Dispense into 10 cm petri dishes.
3. Uracil Dropout solid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g D-glucose, 5 g ammonium sulfate, 1.7 g yeast nitrogen base without amino acids or ammonium sulfate, 60 mg adenine sulfate, 50 mg L-arginine HCl, 75 mg L-aspartic acid, 100 mg L-glutamic acid, 20 mg L-histidine HCl, 50 mg L-isoleucine, 100 mg L-leucine, 120 mg L-lysine HCl, 20 mg L-methionine, 50 mg L-phenylalanine, 375 mg L-serine, 100 mg L-threonine, 50 mg L-tryptophan, 50 mg L-tyrosine, 150 mg L-valine, and deionized water to 1 L total volume. Adjust pH to 5.8 before autoclaving. Dispense into 10 cm petri dishes.
4. YPDA solid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g D-glucose, 20 g peptone, 10 g yeast extract, and deionized water to 1 L total volume. Supplement with 2 mL of 0.5% adenine sulfate solution (filter sterilized). Dispense into 10 cm petri dishes.
5. Adenine Dropout solid media: For 1 L, autoclave the following mixture for 30 min at 121 °C: 20 g agar, 20 g D-glucose, 5 g ammonium sulfate, 1.7 g yeast nitrogen base without amino acids or ammonium sulfate, 50 mg L-arginine HCl, 75 mg L-aspartic acid, 100 mg L-glutamic acid, 20 mg L-histidine HCl, 50 mg L-isoleucine, 100 mg L-leucine, 120 mg L-lysine HCl, 20 mg L-methionine, 50 mg L-phenylalanine, 375 mg L-serine, 100 mg L-threonine, 50 mg L-tryptophan, 50 mg L-tyrosine, 150 mg L-valine, 60 mg uracil, and deionized water to 1 L total volume. Adjust pH to 5.8 before autoclaving. Dispense into 10 cm petri dishes.
6. Velveteen squares.
7. Replica plating block.

---

## 3 Methods

### 3.1 Day 1: Yeast Cultures

1. For each replicate culture, inoculate ySR127 yeast from a single colony into 5 mL of YPDA in each flask.
2. Shake at 23 °C and 250 revolutions per minute (RPM) for 3 days, with flasks in a vertical orientation (*see Note 2*).
3. (Optional) A control yeast strain is available (yKC023), where the triple reporter gene cassette is situated at an internal locus in Chromosome II [7]. Telomere uncapping and resection also occur at 37 °C, but the reporter DNA remains in a double-stranded state. This control can be included to assess the mutagenicity of a given agent toward dsDNA.

### 3.2 Day 4: Mutagen Exposure

1. Dilute cultures from Day 1 by tenfold using fresh YPDA (e.g., dilute 500 µL of culture with 4.5 mL of YPDA) in fresh flasks.
2. Shake at 37 °C for 3–6 h at 250 RPM (*see Note 3*).
3. Add exogenous mutagen and maintain at 37 °C, with shaking, for desired duration of treatment. Optimal duration of treatment and mutagen concentration will have to be worked out by trial and error. Set up mock treated controls in parallel (*see Note 4*).
4. Wash cells from Day 1 cultures in water three times. Determine cell titer using microscope and hemocytometer (*see Note 5*).
5. Spread appropriate dilutions from Day 1 cultures on Synthetic Complete and on Canavanine, Low Adenine plates to assess baseline survival and mutant frequency, respectively (*see Note 6*). This can be done using a cell spreader (flame sterilized with ethanol between platings) with or without a turntable, or by shaking sterile glass beads to spread cells around a plate (*see Note 7*). All platings should be done in triplicate.
6. When mutagen treatment is completed, wash treated cells and mock treated controls in water three times. Determine cell titers using microscope and hemocytometer. Spread appropriate dilutions on Synthetic Complete and on Canavanine, Low Adenine (*see Note 8*).
7. Incubate plates at 23 °C for  $\geq 5$  days (*see Note 9*).

### 3.3 Day $\geq 9$ : Colony Counting and Phenotype Verification

1. Count the number of colonies on Synthetic Complete plates. Divide this number by the number of cells plated to obtain the fraction of viable, colony forming cells within a sample. This is the plating efficiency.
2. Count the total number of colonies on Canavanine, Low Adenine plates. This is the total number of colonies with *CAN1* inactivation (*see Note 10*).

3. Count the number of red (and pink) colonies on Canavanine, Low Adenine plates. This is an estimate of the number of colonies with both *CAN1* and *ADE2* inactivated (*see Note 11*).
4. Replica the Canavanine, Low Adenine plates onto: (fresh) Canavanine, Low Adenine; YPG; and Uracil Dropout media, in that order (*see Note 12*).
5. After incubation at 23 °C, compare these replicas to their source plates. Colonies that fail to grow on YPG have lost mitochondrial respiratory function, while those that fail to grow on Uracil Dropout are inactivated for both *CAN1* and *URA3* (*see Note 13*).
6. Streak from colonies of interest (i.e., Can<sup>r</sup> Ade<sup>-</sup>, Can<sup>r</sup> Ura<sup>-</sup>, or Can<sup>r</sup> Ade<sup>-</sup> Ura<sup>-</sup>) onto YPDA and grow at 23 °C.
7. Patch a single colony from each streak onto YPDA and grow at 23 °C. ~2 cm-long patches will suffice.
8. Replica the patches onto the following media to verify phenotypes: Canavanine, Low Adenine; YPG; Adenine Dropout; and Uracil Dropout, in that order. Patches with confirmed phenotypes of interest are genetically homogeneous and suitable for downstream analyses, e.g., sequencing to identify mutations (*see Note 14*).

---

## 4 Notes

1. Be sure to omit arginine from these media to promote efficient uptake of its toxic analog, canavanine, into cells via functional Can1 protein. This is critical, as it is the basis for the selection of canavanine resistant (i.e., Can<sup>r</sup>) mutants.
2. Do not exceed 23 °C, as doing so risks sporadic telomere uncapping. Growth for 3 days effectively synchronizes the vast majority of cells in the G<sub>1</sub> phase, since the cultures are well past logarithmic growth. Growth for longer is also fine. Using cultures grown for only 2 days is possible, but at the cost of somewhat less uniform arrest at 37 °C. Growth in tubes will be slower than in flasks and cell titers likely will not be as high.
3. Vigorous shaking in flasks promotes synchronous arrest. A 6-h incubation at 37 °C yields a more uniformly arrested cell population than 3-h incubation, as the former presumably allows more cells enough time to reach the G<sub>2</sub> arresting point.
4. It is sensible to check (under the microscope) that the cultures are arrested properly prior to adding mutagen. >90% of cells should be arrested as large dumbbells, where the buds are roughly the same size as mother cells.

5. It is possible to skip cell counting and simply plate sensible dilutions of cells. This would be adequate for determining mutant frequency, but the fraction of viable cells within the sample would be unknown. This latter quantity, the plating efficiency, is important for assessing whether a treatment is cytotoxic.
6. As starting points, I recommend spreading 1,000 cells per Synthetic Complete and  $10^8$  per Canavanine, Low Adenine plate. Spreading more than  $10^8$  per selection plate is inadvisable, as pseudo-resistant papillae can grow when cells are plated at too high of a density. 3-day old starter cultures can have  $>10^9$  cells per mL. Thus if skipping cell counting, one can try spreading a  $10^6$ - and/or  $10^7$ -fold dilution on Synthetic Complete and a 10- and/or 100-fold dilution on Canavanine, Low Adenine. A culture with an unusually high *spontaneous* Can<sup>r</sup> frequency should be dropped from further consideration, as pre-existing Can<sup>r</sup> mutations in the starting culture will confound the analysis of induced mutagenesis. The spontaneous, background Can<sup>r</sup> frequency is typically  $\sim 10^{-6}$  per viable cell.
7. The glass bead method is quicker, as one can add cells to multiple plates and swirl the plates altogether as a stack. The drawback is that a non-negligible fraction of colonies would grow up near the rim of the plates and would not be included when replica plating to check for respiratory competency and inactivation of *URA3*.
8. Depending on the mutagen treatment, as few as  $10^4$  and as many as  $10^6$  cells should be spread per Canavanine, Low Adenine plate. A typical pilot experiment should try 1,000 cells per Synthetic Complete, and  $10^5$  and/or  $10^6$  cells per Canavanine, Low Adenine plate. Mock-treated controls should be tested initially with  $10^6$  cells per Canavanine, Low Adenine plate. Adjustments to platings can be made in follow-on experiments. If plating cells without counting titers, choose order of magnitude dilutions that should yield (approximately) the desired number of cells per plate as mentioned here.
9. 5 days' incubation at 23 °C is a bare minimum, but identification of red colonies can be unreliable because there may not be enough time for red pigments to build up in some mutant colonies. For the most reliable quantification of red colonies, 7 days at 23 °C followed by 7 days at 4 °C are recommended. Low temperature incubation markedly intensifies red pigmentation.
10. To estimate the number of *viable* cells plated on Canavanine, Low Adenine, multiply the total number of plated cells by the plating efficiency from the Synthetic Complete plates. To calculate mutant frequency, divide the number of Can<sup>r</sup> colonies



by the number of viable cells plated. If plating dilutions with undetermined titers, the number of viable cells plated on Canavanine, Low Adenine can still be estimated. For example, suppose 500 colonies are observed on Synthetic Complete when plated with a  $10^6$ -fold dilution. If a tenfold dilution is plated on selection media, then one would expect  $500 \times 10^5 = 5 \times 10^7$  viable cells.

11. This only gives an estimate for Can<sup>r</sup> Ade<sup>-</sup> mutation cluster frequency, as the entire reporter cassette can be lost at a frequency of  $\sim 10^{-6}$  per viable cell plated. Presumably, this is due to strand breakage within the long 3' ssDNA overhang. Cassette loss events can be recognized easily, as the red pigmentation emerges early upon incubation and is always a uniformly deep hue; these colonies tend to be small; and they are always Ura<sup>-</sup>.
12. Take care to use a light touch when replica plating, so that excessive numbers of cells are not transferred onto the replicas. This is especially important when replica plating onto Uracil Dropout media, as too heavy of an imprint will result in a high background that can make Ura<sup>-</sup> colonies hard to distinguish reliably.
13. Colonies that are respiratory deficient (i.e.,  $\rho^-$  or  $\text{rho}^-$ ) should be dropped from further consideration, as these colonies have profoundly altered metabolism. Also note that respiratory deficiency can result in reddish pigmentation, so diagnostic replicas are necessary to distinguish between Ade<sup>-</sup> and  $\rho^-$ .
14. Partial loss-of-function Ade<sup>-</sup> patches will grow slowly on Adenine Dropout. These can exhibit a range of reddish hues, so it is important to compare with YPG to verify that the redness cannot be attributed to respiratory deficiency.

---

## Acknowledgments

KC thanks Dr. Dmitry A. Gordenin for an historical perspective on the development and usage of previous generation ssDNA reporter systems. This work was supported by NSERC Discovery Grant 147465 and start-up funding from the University of Ottawa.

## References

1. Ayala FJ, Fitch WM (1997) Genetics and the origin of species: an introduction. *Proc Natl Acad Sci* 94:7691–7697
2. Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362:709–715. doi:[10.1038/362709a0](https://doi.org/10.1038/362709a0)
3. Irigaray P, Belpomme D (2010) Basic properties and molecular mechanisms of exogenous chemical carcinogens. *Carcinogenesis* 31:135–148. doi:[10.1093/carcin/bgp252](https://doi.org/10.1093/carcin/bgp252)
4. Ikehata H, Ono T (2011) The mechanisms of UV mutagenesis. *J Radiat Res* 52:115–125. doi:[10.1269/jrr.10175](https://doi.org/10.1269/jrr.10175)

5. Rogozin IB, Pavlov YI (2003) Theoretical analysis of mutation hotspots and their DNA sequence context specificity. *Mutat Res* 544:65–85. doi:[10.1016/S1383-5742\(03\)00032-2](https://doi.org/10.1016/S1383-5742(03)00032-2)
6. Boiteux S, Jinks-Robertson S (2013) DNA repair mechanisms and the bypass of DNA damage in *saccharomyces cerevisiae*. *Genetics* 193:1025. doi:[10.1534/genetics.112.145219](https://doi.org/10.1534/genetics.112.145219)
7. Chan K, Sterling JF, Roberts SA, Bhagwat AS, Resnick MA, Gordenin DA (2012) Base damage within single-strand DNA underlies in vivo hypermutability induced by a ubiquitous environmental agent. *PLoS Genet* 8:e1003149. doi:[10.1371/journal.pgen.1003149](https://doi.org/10.1371/journal.pgen.1003149)
8. International Agency for Research on Cancer (2010) Identification of research needs to resolve the carcinogenicity of high-priority IARC carcinogens
9. Booth C, Griffith E, Brady G, Lydall D (2001) Quantitative amplification of single-stranded DNA (QAOS) demonstrates that *cdc13-1* mutants generate ssDNA in a telomere to centromere direction. *Nucleic Acids Res* 29:4414–4422. doi:[10.1093/nar/29.21.4414](https://doi.org/10.1093/nar/29.21.4414)
10. Garvik B, Carson M, Hartwell L (1995) Single-stranded DNA arising at telomeres in *cdc13* mutants may constitute a specific signal for the RAD9 checkpoint. *Mol Cell Biol* 15:6128–6138. doi:[10.1128/MCB.15.11.6128](https://doi.org/10.1128/MCB.15.11.6128)
11. Roberts SA, Sterling J, Thompson C, Harris S, Mav D, Shah R, Klimczak LJ, Kryukov GV, Malc E, Mieczkowski PA, Resnick MA, Gordenin DA (2012) Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol Cell* 46:424–435. doi:[10.1016/j.molcel.2012.03.030](https://doi.org/10.1016/j.molcel.2012.03.030)
12. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jönsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerød A, Tutt A, Martens JWM, Aparicio SAJR, Borg Å, Salomon AV, Thomas G, Børresen-Dale A-L, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ, Stratton MR (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell* 149:979–993. doi:[10.1016/j.cell.2012.04.024](https://doi.org/10.1016/j.cell.2012.04.024)
13. Yang Y, Sterling J, Storici F, Resnick MA, Gordenin DA (2008) Hypermutability of damaged single-strand DNA formed at double-strand breaks and uncapped telomeres in yeast *saccharomyces cerevisiae*. *PLoS Genet* 4:e1000264. doi:[10.1371/journal.pgen.1000264](https://doi.org/10.1371/journal.pgen.1000264)
14. Burch LH, Yang Y, Sterling JF, Roberts SA, Chao FG, Xu H, Zhang L, Walsh J, Resnick MA, Mieczkowski PA, Gordenin DA (2011) Damage-induced localized hypermutability. *Cell Cycle* 10:1073–1085. doi:[10.4161/cc.10.7.15319](https://doi.org/10.4161/cc.10.7.15319)
15. Degtyareva NP, Heyburn L, Sterling J, Resnick MA, Gordenin DA, Doetsch PW (2013) Oxidative stress-induced mutagenesis in single-strand DNA occurs primarily at cytosines and is DNA polymerase zeta-dependent only for adenines and guanines. *Nucleic Acids Res* 41:8995–9005. doi:[10.1093/nar/gkt671](https://doi.org/10.1093/nar/gkt671)
16. Chan K, Resnick MA, Gordenin DA (2013) The choice of nucleotide inserted opposite abasic sites formed within chromosomal DNA reveals the polymerase activities participating in translesion DNA synthesis. *DNA Repair* 12:878–889. doi:[10.1016/j.dnarep.2013.07.008](https://doi.org/10.1016/j.dnarep.2013.07.008)
17. Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, Getz G, Gordenin DA (2015) An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat Genet* 47:1067–1072
18. Morrison A, Bell JB, Kunkel TA, Sugino A (1991) Eukaryotic DNA polymerase amino acid sequence required for 3' → 5' exonuclease activity. *Proc Natl Acad Sci* 88:9473–9477. doi:[10.1073/pnas.88.21.9473](https://doi.org/10.1073/pnas.88.21.9473)

## Analyzing Genome Rearrangements in *Saccharomyces cerevisiae*

Anjana Srivatsan, Christopher D. Putnam, and Richard D. Kolodner

### Abstract

Genome rearrangements underlie different human diseases including many cancers. Determining the rates at which genome rearrangements arise and isolating unique, independent genome rearrangements is critical to understanding the genes and pathways that prevent or promote genome rearrangements. Here, we describe quantitative *S. cerevisiae* genetic assays for measuring the rates of accumulating genome rearrangements including deletions, translocations, and broken chromosomes healed by *de novo* telomere addition that result in the deletion of two counter-selectable genes, *CANI* and *URA3*, placed in the nonessential regions of the *S. cerevisiae* genome. The assays also allow for the isolation of individual genome rearrangements for structural studies, and a method for analyzing genome rearrangements by next-generation DNA sequencing is provided.

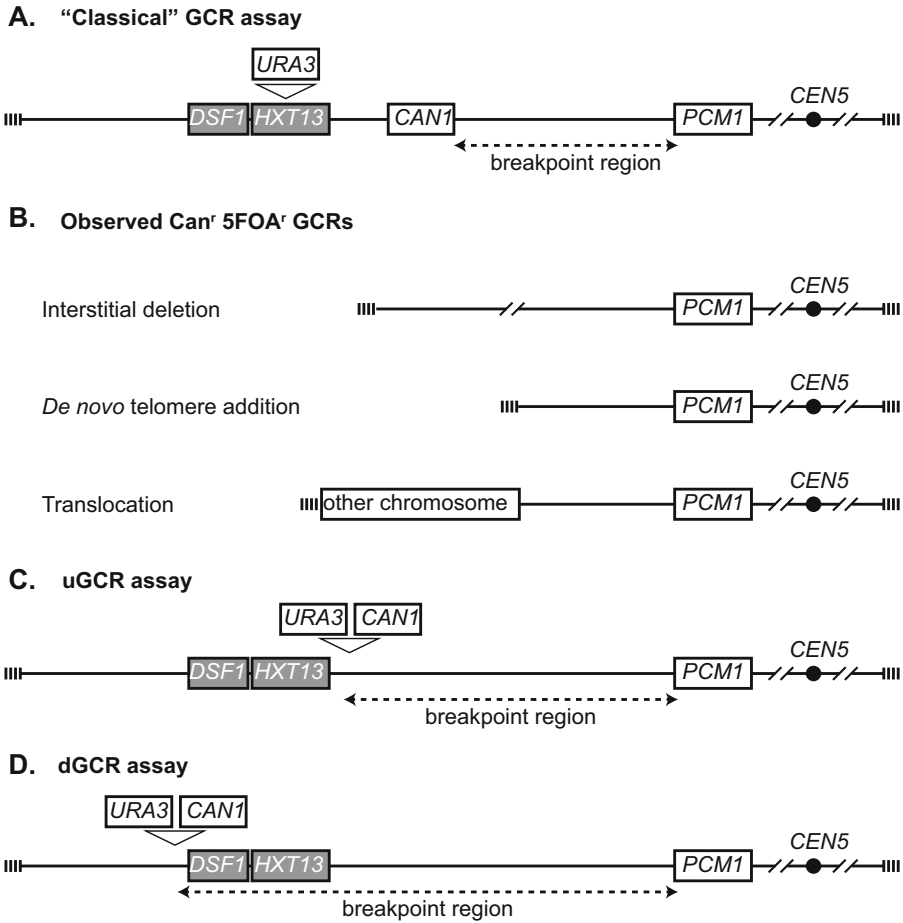
**Key words** Genome instability, Deletion, Monocentric translocation, Dicentric translocation, De novo telomere addition, Genetics, GCR rates, Whole-genome sequencing

---

### 1 Introduction

Maintaining the stability of the genome during cell division is critical for cell survival and normal cell growth. Understanding the pathways and mechanisms by which cells prevent genome rearrangements (herein called Gross Chromosomal Rearrangements or GCRs) has been greatly facilitated by the development of simple genetic methods for measuring the rate at which GCRs occur in different wild-type and mutant *S. cerevisiae* strains [1–3]. Such methods have made it possible to identify genes in which defects cause increased or decreased GCR rates, perform epistasis analysis of these mutations, and isolate independent GCRs for structural analysis (for example *see* refs. 4, 5). These types of studies have provided a growing picture of the cellular pathways that act to prevent GCRs as well as the pathways that produce GCRs.

The key observation that led to the development of a broad series of assays for measuring the rate of accumulation of GCRs was



**Fig. 1** GCR assays. (a) The “classical” GCR assay is constructed by inserting *URA3* into *HXT13*. The breakpoint region (*dashed line*) is between the most telomeric essential gene, *PCM1*, and the most centromeric marker gene, *CAN1*. (b) A variety of GCR products have been observed, including interstitial deletions, terminal deletions healed by *de novo* telomere additions, and translocations. The types of translocations are dictated by the orientation, presence, or absence of a centromere and source of the sequence joined to the broken chromosome V. (c) The unique sequence or uGCR assay contains a *CAN1 URA3* cassette centromeric to the *DSF1-HXT13* segmental duplication (*gray*). (d) The duplication-mediated or dGCR assay contains a *CAN1 URA3* cassette telomeric to the *DSF1-HXT13* segmental duplication (*gray*)

that the *CAN1*-containing portion of the left arm of chromosome V from the telomere to *PCM1* was nonessential [6]. This allowed the development of the first GCR assay (referred to as the “classical” GCR assay) in which the *URA3* gene was inserted into *HXT13* between *CAN1* and the telomere (Fig. 1; [2]). During normal mitotic growth of haploid cells containing this genetically marked chromosome, progeny arise that are resistant to both canavanine (Can) and 5-fluoroorotic acid (5-FOA). These double-drug-resistant progeny contain a rearranged chromosome V in which *URA3* and *CAN1* have been deleted by the formation of GCRs including:

(1) interstitial deletions, (2) terminal deletions of chromosome V L that are healed by the addition of a new telomere, and (3) terminal deletions of chromosome V L healed by fusion to a fragment of another chromosome, generating a translocation, which can either be stable monocentric translocations or dicentric translocations that are unstable and undergo additional rounds of rearrangement (Fig. 1) [2, 6–8]. By measuring the rate of occurrence of  $\text{Car}^r$  5- $\text{FOA}^r$  progeny, GCR rates can be determined and by determining the structure of the GCRs present in individual progeny, the rates of occurrence of specific types of GCRs can be determined.

To increase the utility of GCR assays, *CAN1* and *URA3* were inserted into a cassette. This made it possible to move the two counter-selectable genes to different chromosomal locations in haploid strains containing mutations at the *CAN1* and *URA3* chromosomal loci to evaluate the effect of chromosomal context on GCR rates. In two such assays (Fig. 1), the *CAN1 URA3* cassette was placed either centromeric (the *yel068c::CAN1/URA3* or uGCR assay) or telomeric (the *yel072w::CAN1/URA3* or dGCR assay) to the *HXT13-DSFI* segmental duplication region on chromosome V L [3]. These assays revealed strong influences of having a sequence in the breakpoint region with divergent homology with other regions of the genome on GCR rates, on the pathways that suppress GCRs and the mechanisms by which GCRs are formed [3]. In addition to these GCR assays, a number of other haploid strain-based GCR assays utilizing counter-selectable markers in nonessential regions of the genome have been described (for example *see* refs. 9, 10). In addition, GCR assays in which amplification of genetic markers is selected for in haploid strains have been described, as have a limited number of diploid strain-based GCR assays (for example *see* refs. 11–13).

GCR structures can be analyzed using a variety of methods such as PCR-based mapping and amplification of breakpoint junctions, multiplex ligation-dependent probe amplification (MLPA), pulsed-field gel electrophoresis (PFGE) combined with Southern blotting, and microarray-based comparative genome hybridization (aCGH), and usually, determining the complete structure of an individual GCR requires multiple methods and approaches [2, 3, 7, 14, 15]. A method that has been recently adapted for the analysis of GCR structures is next-generation whole-genome sequencing (NGS) (Fig. 3; [16, 17]). Analysis of GCRs using paired-end sequencing, which provides sequence from both ends of genomic DNA fragments, is particularly useful for the analysis of GCRs and should be the method used when performing NGS on GCR-containing strains. Compared to prior methods, NGS is less labor-intensive and more rapidly provides genome-wide data on copy number, point mutations, and genome rearrangements. Furthermore, the small size of the *S. cerevisiae* genome allows multiple

strains to be sequenced simultaneously (multiplexed), greatly reducing the cost of analysis.

In this article, we describe the use of GCR assays based on selection against the presence of *URA3* and *CAN1* to both measure GCR rates and isolate progeny containing unique GCRs for structural analysis. We also describe a method for preparing 600 bp multiplexed libraries from *S. cerevisiae* strains for paired-end whole-genome sequencing using an Illumina HiSeq 2500 for characterizing GCR structures (*see Note 1*). The general approach described is applicable to many other types of GCR assays.

---

## 2 Materials

### 2.1 Determining GCR Rates

1. YPD (Yeast extract Peptone Dextrose) liquid medium: For 1 l of medium, add 10 g Bacto-yeast extract and 20 g Bacto-peptone to 950 ml deionized water, and sterilize by autoclaving at 121 °C for 20 min. Cool and add 50 ml of a 40% (w/v) sterile glucose (*D*-dextrose) solution.
2. YPD plates: For 1 l of medium (40 plates), add 10 g Bacto-yeast extract, 20 g Bacto-peptone, and 21 g of Bacto-agar to 950 ml deionized water in a 2 l flask, place a magnetic stir bar in the flask and sterilize by autoclaving at 121 °C for 20 min. Cool to 60 °C and add 50 ml of a 40% (w/v) sterile glucose (*D*-dextrose) solution. Mix well by stirring and pour 25 ml per plate in 100 × 15 mm round petri dishes. Allow the plates to dry for 1–2 days at room temperature.
3. 40% (w/v) glucose solution: Dissolve 400 g glucose (*D*-dextrose) in 600 ml warm deionized water with continuous stirring, adjust the volume to 1000 ml, and sterilize by autoclaving at 121 °C for 20 min.
4. Can 5-FOA plates: Add 23–25 g Bacto-agar to 750 ml deionized water in a 2 l flask and place a magnetic stir bar in the flask. Sterilize the agar mixture by autoclaving at 121 °C for 20 min and then stir on a magnetic stir plate to allow the mixture to cool to ~65 °C (*see Note 2*). While the agar mixture cools, prepare the dropout mix as follows. Dissolve 60 mg L-canavanine sulfate (Can) and 1 g 5-fluoroorotic acid (5-FOA) in 200 ml warm deionized water. Add 2 g of dropout powder minus arginine (US Biological, D9518), 6.7 g of yeast nitrogen base without amino acids (with ammonium sulfate), and 50 ml of 40% (w/v) sterile glucose solution, and stir well to dissolve. The dropout solution can be filter sterilized, although this is generally not necessary because the combination of Can and 5-FOA is capable of killing most potential contaminants. Keep the dropout mix at 55–60 °C until the autoclaved agar mixture has cooled sufficiently. Carefully transfer the dropout mix to

the autoclaved agar mixture, avoiding the formation of air bubbles. Pour approximately 70 ml per plate in 150 × 15 mm round petri dishes or 25 ml per plate in 100 × 15 mm round petri dishes as needed. Allow the plates to dry for 1–2 days at room temperature.

5. Sterile deionized water.
6. Sterile plastic culture tubes (15 ml round-bottomed tubes, 50 ml conical tubes).
7. Sterile glass 250 ml culture flasks for larger cultures.
8. Sterile toothpicks to streak out, patch, and inoculate colonies.
9. Surgical scalpel.
10. Sterile microcentrifuge tubes—1.5 ml.
11. Sterile 4 mm glass beads.
12. Benchtop microcentrifuge.
13. Centrifuge with swinging bucket rotor.
14. Incubator set to 30 °C.
15. Shaker set to 30 °C.
16. Sterile velvets.
17. Replica-plating block.

## **2.2 For Preparing Libraries for NGS to Analyze GCR Structures**

1. YPD liquid medium.
2. YPD plates.
3. 40% (w/v) sterile glucose solution.
4. Sterile deionized water.
5. Genra Puregene Yeast/Bact. kit (Qiagen) or preferred genomic DNA extraction kit.
6. 4 mg/ml RNase A (Qiagen or equivalent).
7. Covaris microTUBEs: microTUBE-15AFA beads screw-cap vials and M220 holder XTU insert microTUBE 15 µl. For 50 µl sample volumes, use microTUBE-50AFA fiber screw-cap vials and M220 holder XTU insert microTUBE 50 µl.
8. Covaris Focused Ultrasonicator M220.
9. FlashGel DNA Cassette 1.2%, 12 + 1, single tier (Lonza).
10. FlashGel Loading Dye 5 × 1 ml, 5× concentrate (Lonza).
11. FlashGel DNA Marker 100 bp to 4 kb (Lonza).
12. FlashGel Dock System (Lonza).
13. End-It DNA end-repair kit (Epicentre Technologies).
14. Klenow DNA polymerase (3′–5′ exo<sup>-</sup>) (5 U/µl) (New England Biolabs (NEB) or equivalent).
15. dATP.

16. Quick ligation kit (NEB or equivalent).
17. Library amplification readymix (KAPA Biosystems).
18. TruSeq PCR-Free LT DNA sample prep kit set A (Illumina).
19. MinElute PCR purification kit (Qiagen or equivalent).
20. MinElute gel extraction kit (Qiagen or equivalent).
21. Certified low-range ultra agarose (Bio-Rad or equivalent).
22. 50× TAE Buffer, pH 8.0.
23. 10× Orange Loading Buffer (NEB or equivalent).
24. 100 bp DNA ladder (Bioline or similar).
25. SYBR Safe DNA gel stain (Thermo Fisher Scientific).
26. Benchtop microcentrifuge.
27. Microcentrifuge tubes—0.5 ml (clear, thin-walled; Thermo Fisher Scientific, or equivalent) and 1.5 ml.
28. Centrifuge with swinging bucket rotor.
29. Sterile toothpicks to streak out and inoculate strains.
30. Sterile culture tubes (15 ml round-bottomed tubes, 50 ml conical tubes).
31. Razor blades and forceps.
32. Nanodrop 2000 UV-Vis spectrophotometer.
33. Dark Reader transilluminator (Iso BioExpress) or other UV transilluminator.
34. Electrophoresis power supply, slab gel trays, and electrophoresis tank.
35. Thermal cycler.
36. Qubit fluorometer.
37. Qubit dsDNA HS assay kit.

---

## 3 Methods

### 3.1 Determining GCR Rates

1. Streak out the *S. cerevisiae* strains of interest containing the desired GCR assay (*see* **Note 3**) on YPD plates (or an appropriate selective medium if plasmid selection is required), and incubate the plates for 2–3 days at 30 °C to obtain well-separated single colonies. Streak out at least two independent biological isolates for each strain.
2. Using a sterile surgical scalpel (*see* **Note 4**), excise and inoculate at least seven colonies from each biological isolate into individual sterile tubes or flasks containing an appropriate volume (*see* **Note 5**) of YPD liquid medium (or appropriate selective medium, if plasmid selection is required). Ensure that the



entire colony is used for inoculation. Grow the cultures to saturation (16–18 h; *see Note 6*) in a 30 °C shaker with vigorous shaking (250–300 rpm).

3. To obtain the viable cell count, plate an appropriate dilution of each culture on YPD (or appropriate selective medium) plates as follows. For each culture, prepare five microcentrifuge tubes containing 90  $\mu$ l sterile deionized water. Transfer 10  $\mu$ l of the saturated culture into the 1st tube and vortex vigorously. Then transfer 10  $\mu$ l from the 1st tube into the 2nd tube, and so on, until the 5th tube, vortexing vigorously before each serial dilution, to obtain a tenfold dilution series. Transfer the entire 0.1 ml of the  $10^{-5}$  dilution to a YPD plate and spread well using sterile glass beads.
4. To select for cells that contain a GCR, plate an appropriate volume (*see Note 7*) of each culture on Can 5-FOA plates. Use 100  $\times$  15 mm and 150  $\times$  15 mm petri dishes to plate up to 1 ml ( $10^8$  cells) and 10 ml ( $10^9$  cells) of the culture, respectively. Use multiple plates if larger culture volumes must be plated; for instance, while plating RDKY3615 (the classical GCR assay), which requires  $\sim$ 50 ml cultures, use five 150  $\times$  15 mm Can 5-FOA plates per culture, i.e., 10 ml culture ( $10^9$  cells) per plate as follows. After plating the appropriate dilution on YPD plates (**step 3**), centrifuge the remaining culture at 3000 rpm ( $\sim$ 2000  $\times g$ ) for 10 min at room temperature and discard the supernatant. For up to 1 ml of culture, resuspend the cells in 0.1 ml of sterile deionized water and plate on a 100  $\times$  15 mm Can 5-FOA plate. For larger culture volumes, resuspend the cells in 0.2 ml of sterile deionized water per 10 ml of original culture, and plate 0.2 ml per 150  $\times$  15 mm Can 5-FOA plate.
5. Incubate the YPD and Can 5-FOA plates at 30 °C for 2–3 days and 3–5 days, respectively.
6. Count and record the number of colonies on each YPD plate (*see Note 8*) and the corresponding Can 5-FOA plates (*see Note 9*).
7. Calculate the GCR rate using the Lea-Coulson method of the median [18]. The spreadsheet illustrated in Fig. 2 provides a method to calculate the median and 95% confidence interval of the median based on the total volume of the cultures and the count of the colonies on the nonselective and selective plates. A crucial step in the calculation is the use of the number of mutants per culture,  $r$ , to estimate  $m$ , which is the most likely number of mutations per culture. The nonlinear relationship between  $m$  and  $r$  requires, in each spreadsheet, the CalcM() macro, which calculates the Lea-Coulson median estimator using the Newton-Raphson method (Fig. 2).

**A.**

	A	B	C	D	E	F	G	H
1	Culture Volume (mL)	15						
2								
3	Culture Quantitation							
4		Culture Number	Dilution Factor	Plated Volume (mL)	Number Colonies	Cells per mL		
5						=E6/D6*C6		
6		1	1.00E+05	0.1	112	1.12E+08		
7		2	1.00E+05	0.1	158	1.58E+08		
8		3	1.00E+05	0.1	184	1.84E+08	=E18/D18*C18	
9		4	1.00E+05	0.1	103	1.03E+08		
10		5	1.00E+05	0.1	98	9.80E+07		
11		6	1.00E+05	0.1	101	1.01E+08		
12		7	1.00E+05	0.1	119	1.19E+08		
13		8	1.00E+05	0.1	135	1.35E+08		
14		9	1.00E+05	0.1	137	1.37E+08	=CalcM(F18* \$B\$1)/(F6*\$B\$1)	
15	Rate Determination							
16		Culture Number	Dilution Factor	Plated Volume (mL)	Number Colonies	Mutants per mL	Rate	
17								
18		1	1	14.99	70	4.67	1.02E-08	
19		2	1	14.99	67	4.47	6.99E-09	
20		3	1	14.99	55	3.67	5.13E-09	
21		4	1	14.99	98	6.54	1.46E-08	
22		5	1	14.99	112	7.47	1.71E-08	
23		6	1	14.99	108	7.2	1.61E-08	
24		7	1	14.99	81	5.4	1.08E-08	
25		8	1	14.99	132	8.81	1.42E-08	
26		9	1	14.99	30	2	4.28E-09	
27	Summary							
28		Rate Location	Number Rates	Lower Pos 95% CI	Upper Pos 95% CI	Median Rate	95% CI Lower	95% CI Upper
29		G18:G26						
30			9	2	8	1.08E-08	5.13E-09	1.61E-08

=COUNT(INDIRECT(B30))	↑
=FLOOR((C30+1)/2-0.9789*SQRT(C30),1)	↑
=C30+1-D30	↑
=MEDIAN(INDIRECT(B30))	↑
=SMALL(INDIRECT(B30),D30)	↑
=SMALL(INDIRECT(B30),E30)	↑

**B.**

```

' Visual Basic Macro to Calculate m
' Defines the function CalcM() for use
' in spreadsheet programs
'
Function CalcM(r As Double) As Double
    Dim fm As Double, dfm As Double, m As Double
    If (r <= 0#) Then
        m = 0
    Else
        m = Exp ( Log(r) * 0.5 )
        fm = m * (1.24 + Log(m)) - r
        dfm = 2.24 + Log(m)
        While (Abs(fm) > 0.01)
            m = m - fm / dfm
            fm = m * (1.24 + Log(m)) - r
            dfm = 2.24 + Log(m)
        Wend
    EndIf
    CalcM = m
End Function

```

**Fig. 2** Spreadsheet for calculating the median GCR rate of a series of cultures and 95% confidence intervals for the median GCR rate. (a) The spreadsheet is displayed with both sample data and the underlying formulas. *Callout boxes* indicate cells that should contain formulas; the required formulas are indicated by the *callout*

### 3.2 Preparation of *S. cerevisiae* Genomic DNA Libraries for Multiplexed Paired-End NGS

1. Select strains for the analysis by NGS. Multiplexed libraries containing DNA from up to 12 strains can be sequenced on an Illumina HiSeq 2500 to obtain sufficient read depth (>20-fold) to reliably identify rearrangements, point mutations, and copy number changes. GCR-containing strains can be obtained from Can<sup>r</sup> 5FOA<sup>r</sup> colonies on plates from fluctuation tests or from patches replica plated onto Can 5-FOA medium when strains have sufficiently high GCR rates (*see Note 7*). It is critical, however, to ensure that all GCR-containing strains analyzed are independently isolated so that multiple descendants of a single GCR-containing cell are not analyzed. Hence, only one colony per culture used in fluctuation analysis or only one colony per patch should be analyzed. When sequencing genomic DNA from GCR-containing strains, it is preferable to also sequence genomic DNA from any parental strains from which the GCR-containing strains were derived so that pre-existing rearrangements and point mutations can be identified. A typical sample of multiplexed libraries from a GCR experiment will contain libraries from 1 parental strain and 11 independently isolated GCR-containing strains; however, the method can be scaled up or down in regard to the number of strains analyzed so long as sufficient coverage for each genome can be achieved.
2. Streak selected GCR-containing strains onto Can 5-FOA plates for single colonies to purify the isolate. Incubate the freshly streaked Can 5-FOA plates for 3–5 days at 30 °C. Inoculate a single colony for each GCR-containing strain into 5 ml of YPD liquid and grow in a 30 °C shaker for 12–16 h or until saturation is reached. In parallel, inoculate the parent strain(s) from which the GCR-containing isolates were derived. To generate a glycerol stock for each overnight culture of a GCR-containing strain, take 0.5 ml of the culture, mix with 0.5 ml of 40% glycerol, and freeze at –80 °C.
3. Prepare genomic DNA from the overnight cultures as follows (*see Note 10*). Harvest the cells from 1.5 ml of each culture by centrifugation at 14,000 rpm ( $\sim 21,000 \times g$ ) for 30 s to 1 min in a benchtop microcentrifuge. Carefully remove the supernatant by pipetting and discard. Suspend the cells in 1 ml of sterile

**Fig. 2** (continued) boxes. For cells F6 to F14, F18 to F26, and G18 to G26, the row specifiers in the formula should be incremented in each row (for example, F6 should have the formula = E6/D6\*C6 and F7 should have the formula = E7/D7\*C7). Most modern spreadsheets have fill functions that will automatically increment these values. **(b)** To calculate the rate in cells G18 to G26, the CalcM() Visual Basic macro must be defined for the spreadsheet. CalcM() initializes the test value for  $m$  to the square root of  $r$  and uses the iterative Newton-Raphson method to solve the transcendental equation  $r/m - \ln(m) - 1.24 = 0$ . This macro functions in both Microsoft Excel and in LibreOffice Calc

deionized water, harvest the cells again by centrifugation, and discard the supernatant. The genomic DNA purification method described here is based on the Gentra PureGene Yeast/Bact. kit with some modifications. Resuspend the cell pellet in 300  $\mu$ l Cell Suspension Solution and pipet to mix well. Add 1.5  $\mu$ l Lytic Enzyme Solution, mix well, and incubate the cells for 1 h at 37 °C. Pellet the cells by centrifugation at 14,000 rpm ( $\sim 21,000 \times g$ ) for 1 min and carefully remove the supernatant by pipetting and discard. Resuspend the cells in 300  $\mu$ l Cell Lysis solution and pipet to mix well. Add 100  $\mu$ l Protein Precipitation Solution and vortex at high speed for 20–30 s. Centrifuge the samples at 14,000 rpm ( $\sim 21,000 \times g$ ) for 3 min to pellet the proteins. Transfer the supernatant to a fresh microcentrifuge tube, add 1.5  $\mu$ l of 4 mg/ml RNase A to the supernatant, mix well and incubate at 37 °C for 1 h. Then add 300  $\mu$ l isopropanol, mix well, and centrifuge the samples at 14,000 rpm for 1 min to precipitate the DNA (the pellet may not always be visible). Discard the supernatant carefully by pouring. Wash the pellet with 300  $\mu$ l 70% ethanol, and centrifuge the samples at 14,000 rpm ( $\sim 21,000 \times g$ ) for 1 min. Remove the ethanol carefully and air-dry the pellet for 5 min. Add 50–100  $\mu$ l DNA Hydration Solution and dissolve the DNA by incubating the tubes at 65 °C for 1 h. Measure the DNA concentration using a Nanodrop spectrophotometer (*see Note 11*). If needed, the genomic DNA samples can be stored at  $-20$  °C overnight or longer.

4. Prepare sonicated DNA samples as follows. Reserve 1  $\mu$ l of each genomic DNA sample as a “pre-sonicated” control. Sonicate 1–5  $\mu$ g of the remainder of the genomic DNA to fragments with an average size of 600 bp using the Covaris M220 or preferred sonicator (*see Note 12*). To verify fragment size range, load 1  $\mu$ l of the pre-sonicated and sonicated DNA samples on a 1.2% FlashGel (use 2  $\mu$ l Flash Gel Dye +2  $\mu$ l sterile deionized water +1  $\mu$ l DNA) with 2  $\mu$ l Flash Gel Ladder, and run the gel at 200 V for 5 min. Sonication should result in a smear of fragments ranging from 200 to 800 bp. The fragment size range can be adjusted by changing the sonication time. If needed, the sonicated DNA samples can be stored at  $-20$  °C overnight or longer.
5. Prepare blunt-ended 5'-phosphorylated DNA fragments using T4 DNA polymerase and T4 polynucleotide kinase. If using the End-It DNA End-Repair kit, treat up to 5  $\mu$ g DNA in a 50  $\mu$ l reaction prepared as follows:

1–34  $\mu$ l DNA (*see Note 13*)

5  $\mu$ l 10 $\times$  End-Repair buffer

5  $\mu$ l 2.5 mM dNTP mix  
5  $\mu$ l 10 mM ATP  
1  $\mu$ l End-Repair enzyme mix  
Sterile water (for a final volume of 50  $\mu$ l)

Mix the reaction well and incubate the reactions at room temperature for 45 min. Purify the DNA using a Qiagen MinElute PCR purification kit or equivalent. If the Qiagen MinElute PCR purification kit is used, follow the manufacturer's protocol with the modification that two elutions of the columns should be performed: the first elution should use 20  $\mu$ l of elution buffer, and the second elution should use 12  $\mu$ l of elution buffer. Combine the two elutions to obtain 32  $\mu$ l of eluate. The two elution steps are required to elute most of the end-repaired DNA from the column. At this stage, all the samples should have approximately equal concentrations of DNA. If needed, the end-repaired DNA samples can be stored at  $-20^{\circ}\text{C}$  overnight or longer.

6. Adenylate the 3' ends of the DNA fragments using Klenow DNA polymerase as follows:

32  $\mu$ l end-repaired DNA  
5  $\mu$ l 10 $\times$  NEBuffer 2  
10  $\mu$ l 1 mM dATP  
3  $\mu$ l Klenow DNA polymerase ( $\text{exo}^-$ , 5 U/ $\mu$ l)

Mix well and incubate the reactions in a thermal cycler at  $37^{\circ}\text{C}$  for 30 min. Purify the DNA using a Qiagen MinElute PCR purification kit, eluting in 15  $\mu$ l of elution buffer. If needed, the A-tailed DNA samples can be stored at  $-20^{\circ}\text{C}$  overnight or longer.

7. Ligate Illumina adapters to the 5' and 3' ends of the A-tailed DNA fragments using Quick DNA ligase. Each adapter mix contains both a universal adapter and an indexed adaptor that tags the fragments so that multiplexed sequences can be assigned to specific samples after sequencing. Ensure that a different indexed adapter mix is used for each sample within one multiplexed library. Set up ligation reactions as follows:

15  $\mu$ l A-tailed DNA  
20  $\mu$ l 2 $\times$  DNA ligase buffer  
2  $\mu$ l 1:10 diluted Illumina indexed adapter mix  
3  $\mu$ l Quick DNA ligase

Mix well and incubate the reactions at  $20^{\circ}\text{C}$  for 15 min in a thermal cycler. Purify the DNA using Qiagen MinElute PCR

purification kit, eluting in 17  $\mu$ l elution buffer. If needed, the samples can be stored at  $-20^{\circ}\text{C}$  overnight or longer.

8. Purify 600–800 bp ligated samples by gel electrophoresis and gel extraction as follows. Prepare 200 ml of 2% high-resolution agarose in  $1\times$  TAE (diluted from  $50\times$  TAE in sterile deionized water), add 20  $\mu$ l Sybr Safe, and cast 2 gels (100 ml each) using 15-well combs. Load 10  $\mu$ l HyperLadder 100 bp in the first lane of the gel. Add 5  $\mu$ l Orange gel loading dye to each DNA sample from the previous step and load the first six samples on the first gel, leaving a gap of one lane after each sample to prevent cross-contamination. Load 10  $\mu$ l HyperLadder 100 bp on the final lane of the gel. Flanking the samples with markers facilitates excision of the desired fragment sizes. Similarly, load the next set of six samples and 100 bp ladder on the second gel. Run the gels at 120 V for 60 min (6–10 V/cm). View the fragments using a Dark Reader transilluminator. At this stage, it is normal to observe very weak or no DNA signal using a transilluminator. Using a fresh razor blade, excise a piece of the gel containing fragments ranging from 600 to 800 bp from each sample lane. Use a fresh blade for each sample or clean the blade thoroughly by washing in water and ethanol between samples to avoid cross contamination. Transfer the gel piece to a microcentrifuge tube using a forceps, cleaning the forceps thoroughly between samples. Purify the DNA from the excised gel pieces using the Qiagen MinElute gel extraction kit, eluting in 11.5  $\mu$ l elution buffer. The purified samples can be stored at  $-20^{\circ}\text{C}$  overnight or longer.
9. Enrich adapter-ligated DNA fragments by PCR amplification (*see Note 14*) using the following reaction mix:

11.5  $\mu$ l DNA from the previous step  
12.5  $\mu$ l  $2\times$  KAPA master mix  
1  $\mu$ l 25  $\mu$ M PCR primer mix

The sequences of the primers in the PCR primer mix are 5'-AAT-GATACGGCGACCACCGAGATCTACAC-3' and 5'-CAAG-CAGAAGACGGCATAACGAGAT-3', and are designed to hybridize to the proprietary Illumina adapter sequences.

Use the following PCR conditions:

Denaturation:  $98^{\circ}\text{C}$  for 45 s

18 cycles (*see Note 15*): denaturation ( $98^{\circ}\text{C}$  for 15 s), annealing ( $65^{\circ}\text{C}$  for 30 s), and extension ( $72^{\circ}\text{C}$  for 30 s)

Final extension:  $72^{\circ}\text{C}$  for 1 min

Hold:  $4^{\circ}\text{C}$  as long as desired

10. Purify the final libraries by running the samples on 2% agarose gels as described in Subheading 3.2, **step 8**. Excise 600–800 bp fragments as described above, and purify the fragments from the gel pieces using the Qiagen MinElute gel extraction kit, eluting the DNA in 15  $\mu$ l elution buffer. The samples can be stored at  $-20^{\circ}\text{C}$  overnight or longer.
11. Measure the DNA concentrations using an assay that is highly specific to double-stranded DNA, such as the Qubit dsDNA HS assay kit, which uses a double-stranded DNA-specific fluorescent dye. Standard spectrophotometric measurements based on the absorbance at 260 nm are often influenced by the presence of RNA and other contaminants in the sample and hence are not suitable for obtaining accurate DNA concentrations for NGS libraries. To prepare samples for the Qubit fluorometer, add 2  $\mu$ l of each sample to 198  $\mu$ l of Qubit working solution (prepared by diluting the Qubit dsDNA HS reagent 1:200 in the Qubit dsDNA HS buffer) in thin-walled, clear 0.5 ml microcentrifuge tubes. To prepare the standards, add 190  $\mu$ l of Qubit working solution to 10  $\mu$ l of each Qubit standard and mix well. Calibrate the Qubit fluorometer before measuring the sample concentrations. Calculate the molar concentration (nM) of each library based on the Qubit reading (ng/ $\mu$ l) assuming that the average fragment length is 600 bp, and hence the average molecular weight is  $600\text{ bp} \times 660\text{ g/mol/bp} = 396,000\text{ g/mol}$ . Dilute all the samples to the same concentration (5 or 10 nM but not less than 2 nM). Pool 2  $\mu$ l of each sample for multiplexing and mix well. Use 5  $\mu$ l of the pooled libraries for paired-end sequencing using the Illumina HiSeq 2500.
12. Sequencing of NGS libraries is now a regularly available service at both commercial companies and sequencing facilities at universities and research institutes. The analyses performed and details such as the file formats of the data returned are very facility-dependent. This protocol assumes that the data returned will be unmapped reads, separated by sample/index, in either a FASTQ or BAM format (which can be converted to a FASTQ format by samtools [19]).
13. Analysis of the sequencing data first requires mapping the sequencing reads to the reference sequence of the *S. cerevisiae* genome (<http://www.yeastgenome.org>) using programs such as Bowtie2 [20]. When mapping paired-end data, many read-mapping programs filter out reads that are not concordant with the reference genome, which eliminates reads that span novel junctions present in GCRs. To preserve these reads, provide the paired-end data as two sets of “unpaired” data to the mapping program (for example using the “-U” flag to Bowtie2) and add back the read identifier using programs

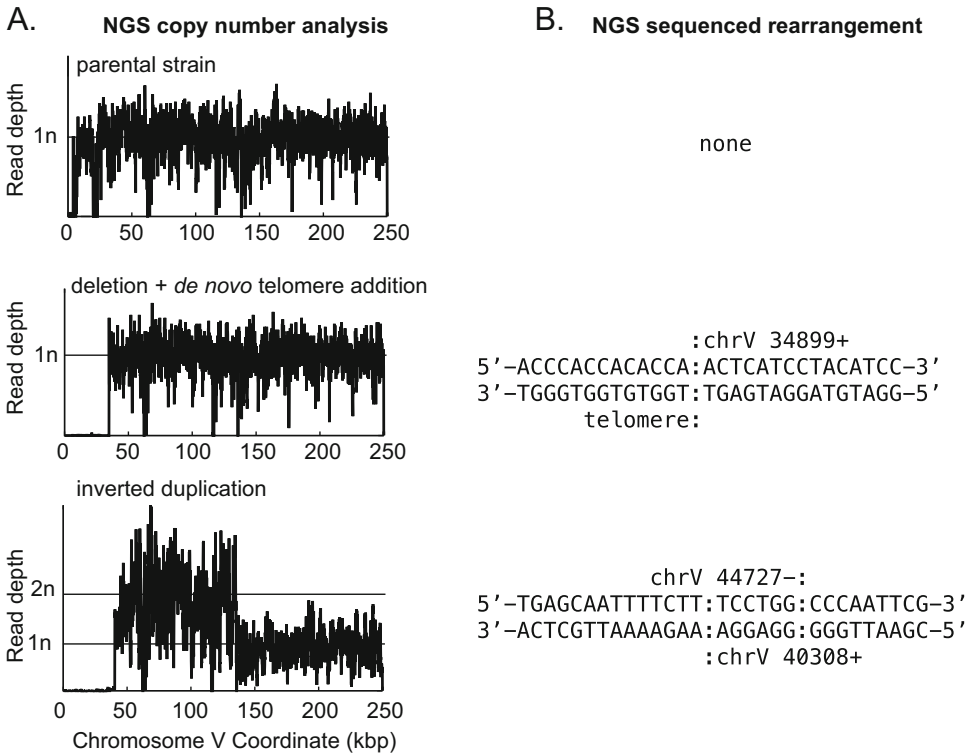
such as graft in the Pyrus suite ([16], <http://www.sourceforge.net/p/pyrus-seq>). The Pyrus suite of programs can then be used to sort the mapped reads and perform copy number analysis, sequence variant identification, and identification and sequencing of genome rearrangements by identification of anomalous paired-end alignments. Many other programs for identifying structural variants that use a variety of strategies have been developed and can also be used to perform these types of analyses [21]. The combination of copy number analysis and identification of rearrangements targeting single copy sequences in the genome can provide great insight into the structure of individual GCRs (Fig. 3), particularly if the junction sequence can be identified using programs such as comice in the Pyrus suite [16]. Many GCRs, however, are mediated by non-allelic homologous recombination between repetitive elements (such as Ty elements) that are larger than 600 bp. These rearrangements cannot be spanned by the NGS libraries described here and will not be identifiable by anomalous paired-end alignment analysis. In these cases, copy number analysis can provide hints as to the structures of these rearrangements, but proof of these structures requires the use of complementary techniques such as PCR and PFGE [7, 9].

---

## 4 Notes

1. The method for generating NGS libraries was modified from the Illumina library preparation guide (Illumina). It should be noted that there are continual improvements in the methods and reagents for NGS library generation that should be evaluated from time to time.
2. Do not allow the autoclaved agar mixture to cool too much because it will rapidly cool once the dropout mix is added and can solidify before it is poured.
3. Strains containing three widely used haploid strain-based assays include RDKY3615 (classical GCR assay; *MATa ura3-52 leu2Δ1 trp1Δ63 his3Δ200 lys2ΔBgl hom3-10 ade2Δ1 ade8 hxt13::URA3*), RDKY6677 (uGCR assay; *MATa ura3-52 leu2Δ1 trp1Δ63 his3Δ200 lys2ΔBgl hom3-10 ade2Δ1 ade8 can1::hisG yel068c::CAN1-URA3 iYEL072W::hph*) and RDKY6678 (dGCR assay; *MATa ura3-52 leu2Δ1 trp1Δ63 his3Δ200 lys2ΔBgl hom3-10 ade2Δ1 ade8 can1::hisG yel072w::CAN1-URA3 iYEL072W::hph*) [2, 3]. Mutations of interest can be introduced into these strains by methods such as PCR-based gene disruption using the following selectable markers: *HIS3*, *TRP1*, G418-resistance (e.g., *kanMX4*), and





**Fig. 3** Examples of GCR structures deciphered through analysis of NGS data. **(a)** Copy number analysis for the left arm of chromosome V is plotted for a parental strain (*top*), a de novo telomere addition GCR (*middle*), and a GCR involving an inverted duplication (*bottom*). Note that both rearranged chromosomes have deleted the region containing the *CAN1 URA3* cassette (around coordinate 25,700 in this GCR assay strain). The inverted duplication has additionally duplicated the region from around coordinate 44,700 to the Ty-related sequence *YELWdelta6* (around coordinate 138,300). **(b)** Sequence of the novel junctions involved in the observed GCR. A de novo telomere (sequence prior to the colon) healed the terminal deletion at chromosome V 34,899 for the strain displayed in the middle of the panel. For the strain displayed at the bottom of the panel, non-concordant read pairs in the NGS data revealed the presence of an inverted duplication, which was also confirmed by the increase in copy number (panel **a**) and through the sequence of the reads that are associated with the novel junction and hence did not map to the reference sequence. For this inverted duplication, there is a six-base identity (sequence between the colons) at the junction between the source and target of the inversion. The inversion was not a perfect hairpin, however, as ~4000 bp separate the two fused sequences

nourseothricin-resistance (e.g., *natNT2*). Alternatively, the classical GCR assay can be introduced into any *CAN1 ura3* strain of interest by replacing *HXT13* with *URA3* (*hxt13::URA3*) [2]. The uGCR or dGCR assay can be introduced into any *can1 ura3* strain of interest by inserting the desired GCR assay cassette (*yel068c::CAN1-URA3* for uGCR and *yel072w::CAN1-URA3* for dGCR), which can be amplified from the plasmids pRDK1379 and pRDK1378, respectively [3], at the appropriate chromosomal site. To facilitate systematic GCR studies using the BY4741 *MATa* yeast deletion

collection, *MAT*alpha query strains that contain the appropriate markers as well as the uGCR or dGCR assay have been constructed [22]: RDKY8625 (uGCR; *MAT*alpha *hom3-10 ura3Δ0 leu2Δ0 trp1Δ63 his3Δ200 lyp1::TRP1 cyh2-Q38K iYFR016::P<sub>MEAI</sub>-LEU2 can1::P<sub>LEU2</sub>-NAT yel068c::CAN1-URA3*) and RDKY7635 (dGCR; *MAT*alpha *hom3-10 ura3Δ0 leu2Δ0 trp1Δ63 his3Δ200 lyp1::TRP1 cyh2-Q38K iYFR016::P<sub>MEAI</sub>-LEU2 can1::P<sub>LEU2</sub>-NAT yel072w::CAN1-URA3*). Additional mutations can be introduced into these query strains by PCR-based gene disruption using the *HIS3* and hygromycin-resistance (e.g., *hphNTI*) markers.

4. Sterilize a scalpel/surgical blade by dipping in ethanol and flaming. Before excising colonies, ensure that the blade is sufficiently cooled by pressing into a region of agar on the edge of the plate, away from any colonies.
5. The culture volume depends on the assay used to monitor the formation of GCRs. The wild-type rates for the classical, uGCR, and dGCR assays are  $3.5 \times 10^{-10}$ ,  $2.27 \times 10^{-9}$ , and  $1.97 \times 10^{-8}$ , respectively [2, 3]. For strains with estimated GCR rates that are similar to or greater than wild-type, use 50 ml, 10 ml, and 4 ml cultures, respectively, for these three assays. For mutant strains that have estimated GCR rates that are lower than wild-type, larger culture volumes can be used if very few or no colonies are observed in patch tests (*see Note 7*) on Can 5-FOA plates. To allow sufficient aeration, the cells must be cultured in tubes or flasks with a capacity of at least four to five times the desired culture volume.
6. For strains with a severe growth defect, use longer incubation times and/or larger culture volumes.
7. The volume of culture plated depends on the estimated GCR rate, which is influenced by the sequences within the break-point region (i.e., the GCR assay used) and the mutations present in the strain of interest. Patch tests can be performed to make a preliminary assessment of the effect of any mutation of interest on the GCR rate before performing fluctuation tests. The dGCR assay is most amenable to this type of test because the rate of accumulating GCRs in the dGCR assay is sufficiently high. Patch tests can be performed as follows. The strains of interest (including a wild-type control) are streaked on YPD plates and incubated for 2–3 days at 30 °C. Using sterile toothpicks, three independent colonies of each strain are spread into individual  $\sim 1 \text{ cm}^2$  squares on YPD plates (including the wild-type strain as a control). The plates are incubated for 2–3 days at 30 °C, and the patches are then replica-plated onto GCR plates using sterile velvets. The number of papillae in each patch is counted after 3–5 days of growth at 30 °C. The wild-

type dGCR strain generally yields 1–5 papillae per patch, and different mutations can cause either increased or decreased papillation. The number of papillae can be used to make an initial estimate of the volume of saturated culture to be plated on GCR plates for fluctuation tests as follows: 1–5 papillae, 0.75–1 ml; 6–15 papillae, 0.3–0.4 ml; ~16–150 papillae, 0.1–0.25 ml; papillae that are too many or too close together to count, 0.1–0.2 ml of a 1:10 dilution of the saturated culture. If plating these volumes yields too few or too many colonies on the Can 5-FOA plates, then increase or decrease the volume plated accordingly. It is important to note that the patch test is significantly influenced by the growth rate of the strain(s) of interest, and therefore patch tests are not reliable predictors of GCR rates for strains with growth defects. Additionally, certain mutations can have assay-specific effects, for instance by increasing the GCR rate in the uGCR but not the dGCR assay (as observed when genes like *EXO1*, *RAD10*, or *RAD6* are deleted [3]). In such cases, performing patch tests using only the dGCR assay can lead to the potentially inaccurate conclusion that the mutation does not impact the GCR rate. Therefore, while patch tests are extremely useful in rapidly screening large numbers of mutations, it is crucial to perform fluctuation tests especially when the mutation of interest causes a growth defect or potentially has an assay-specific effect on the GCR rate.

8. *S. cerevisiae* strains yield on average  $10^8$  cells per ml of saturated culture in YPD; therefore, 0.1 ml of the  $10^{-5}$  dilution plated on YPD will yield approximately 100 colonies.
9. If all Can 5-FOA plates contain very few or no colonies, plate larger culture volumes.
10. The genomic DNA extraction method described here uses the Gentra Puregene Yeast/Bact. kit (Qiagen), but other genomic DNA extraction methods can be used at this step, with the modification that RNase treatment should be performed prior to DNA precipitation and resuspension.
11. If the DNA yield from 1.5 ml of the overnight culture is low, the culture volume used for the DNA isolation should be increased. For example, a 10 ml overnight culture can be split into multiple separate 1.5 ml scale DNA isolations. When multiple DNA isolations per culture are performed, the DNA can be combined in the final step by resuspending the DNA from all the tubes with the same 50  $\mu$ l of DNA Hydration Solution.
12. The Covaris microTUBE-15 and microTUBE-50 vials have a DNA input limit of 1  $\mu$ g and 5  $\mu$ g, respectively; therefore, use 1 or 2  $\mu$ g genomic DNA in a final volume of volume of 15–20

( $\pm$ ) 1  $\mu$ l and 50 ( $\pm$ ) 3  $\mu$ l, respectively, depending on the DNA concentration. Use the manufacturer settings to obtain 600 bp fragments in each case.

13. If the DNA concentration is low, use the entire 50  $\mu$ l of sonicated DNA sample and scale up the end-repair reaction to 125  $\mu$ l with 50  $\mu$ l DNA, 12.5  $\mu$ l each of the 10 $\times$  End-repair buffer, 2.5 mM dNTP mix, and 10 mM ATP, 2.5  $\mu$ l of End-It enzyme mix, and sterile water to a final volume of 125  $\mu$ l.
14. There are alternative PCR-free methods to enrich adapter-ligated fragments, as described at [www.illumina.com](http://www.illumina.com).
15. If insufficient amplification is observed, use real-time PCR to determine the optimal number of cycles.

---

## Acknowledgments

This work was supported by NIH grant GM26017 and the Ludwig Institute for Cancer Research.

## References

1. Chan JE, Kolodner RD (2011) A genetic and structural study of genome rearrangements mediated by high copy repeat Ty1 elements. *PLoS Genet* 7:e1002089
2. Chen C, Kolodner RD (1999) Gross chromosomal rearrangements in *Saccharomyces cerevisiae* replication and recombination defective mutants. *Nat Genet* 23:81–85
3. Putnam CD, Hayes TK, Kolodner RD (2009) Specific pathways prevent duplication-mediated genome rearrangements. *Nature* 460:984–989
4. Myung K, Chen C, Kolodner RD (2001) Multiple pathways cooperate in the suppression of genome instability in *Saccharomyces cerevisiae*. *Nature* 411:1073–1076
5. Pennaneach V, Kolodner RD (2004) Recombination and the Tel1 and Mec1 checkpoints differentially effect genome rearrangements driven by telomere dysfunction in yeast. *Nat Genet* 36:612–617
6. Chen C, Umezu K, Kolodner RD (1998) Chromosomal rearrangements occur in *S. cerevisiae* rfa1 mutator mutants due to mutagenic lesions processed by double-strand-break repair. *Mol Cell* 2:9–22
7. Pennaneach V, Kolodner RD (2009) Stabilization of dicentric translocations through secondary rearrangements mediated by multiple mechanisms in *S. cerevisiae*. *PLoS One* 4:e6389
8. Putnam CD, Pennaneach V, Kolodner RD (2004) Chromosome healing through terminal deletions generated by de novo telomere additions in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 101:13262–13267
9. Narayanan V, Mieczkowski PA, Kim HM, Petes TD, Lobachev KS (2006) The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. *Cell* 125:1283–1296
10. Kanellis P, Gagliardi M, Banath JP, Szilard RK, Nakada S, Galicia S et al (2007) A screen for suppressors of gross chromosomal rearrangements identifies a conserved role for PLP in preventing DNA lesions. *PLoS Genet* 3:e134
11. Koszul R, Caburet S, Dujon B, Fischer G (2004) Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J* 23:234–243
12. Umezu K, Hiraoka M, Mori M, Maki H (2002) Structural analysis of aberrant chromosomes that occur spontaneously in diploid *Saccharomyces cerevisiae*: retrotransposon Ty1 plays a crucial role in chromosomal rearrangements. *Genetics* 160:97–110
13. Zhang Y, Saini N, Sheng Z, Lobachev KS (2013) Genome-wide screen reveals replication pathway for quasi-palindrome fragility dependent on homologous recombination. *PLoS Genet* 9:e1003979

14. Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD (2005) Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* 120:587–598
15. Chan JE, Kolodner RD (2012) Rapid analysis of *Saccharomyces cerevisiae* genome rearrangements by multiplex ligation-dependent probe amplification. *PLoS Genet* 8:e1002539
16. Putnam CD, Pallis K, Hayes TK, Kolodner RD (2014) DNA repair pathway selection caused by defects in TEL1, SAE2, and de novo telomere addition generates specific chromosomal rearrangement signatures. *PLoS Genet* 10:e1004277
17. Serero A, Jubin C, Loeillet S, Legoix-Ne P, Nicolas AG (2014) Mutational landscape of yeast mutator strains. *Proc Natl Acad Sci U S A* 111:1897–1902
18. Lea DE, Coulson CA (1949) The distribution of the numbers of mutants in bacterial populations. *J Genet* 49:264–285
19. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079
20. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
21. Guan P, Sung WK (2016) Structural variation detection using next-generation sequencing data: a comparative technical review. *Methods* 102:36–49
22. Putnam CD, Srivatsan A, Nene RV, Martinez SL, Clotfelter SP, Bell SN et al (2016) A genetic network that suppresses genome rearrangements in *Saccharomyces cerevisiae* and contains defects in cancers. *Nat Commun* 7:11256

## High-Resolution Mapping of Modified DNA Nucleobases Using Excision Repair Enzymes

Monica Ransom, D. Suzi Bryan, and Jay R. Hesselberth

### Abstract

Modification of DNA nucleobases has a profound effect on genome function. We developed a method that maps the positions of the modified DNA nucleobases throughout genomic DNA. This method couples in vitro nucleobase excision with massively parallel DNA sequencing to determine the location of modified DNA nucleobases with single base precision. This protocol was used to map uracil incorporation and UV photodimers in DNA, and a modification of the protocol has been used to map sparse modification events in cells. The Excision-seq protocol is broadly applicable to a variety of base modifications for which an excision enzyme is available.

**Key words** Uracil incorporation, Base excision repair, UV photodimers, Circular ligation, High-throughput sequencing

---

### 1 Introduction

Methods for the detection of modified DNA bases include chromatographic techniques, PCR-based assays, comet assays, mass spectrometry, electrochemistry, radioactive labeling, and immunochemical methods [1]. These methods are powerful for studying single or small groups of loci but less effective for unknown targets or on a larger scale. Direct sequencing methods such as SMRT sequencing that rely on polymerase dynamics as it crosses a modification are powerful tools for detecting single types of DNA modification but do not scale well to large genomes [2].

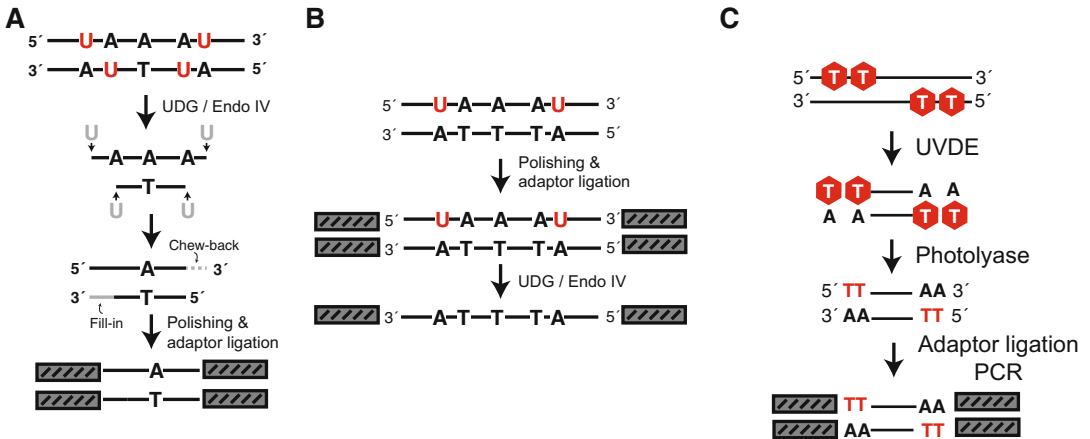
We developed a method that uses base excision repair enzymes to create nicks or gaps in DNA, which are then recovered in a sequencing library for their identification in genomic DNA. Base excision repair is a conserved repair system that recognizes and removes DNA lesions to maintain genomic stability [3]. Repair of these lesions is initiated by the recognition and cleavage of the modification by a modification-specific DNA glycosylase yielding in an apurinic site [3]. Apurinic sites are resolved to single-stranded

breaks and are repaired by additional systems involving DNA polymerase and ligase [4].

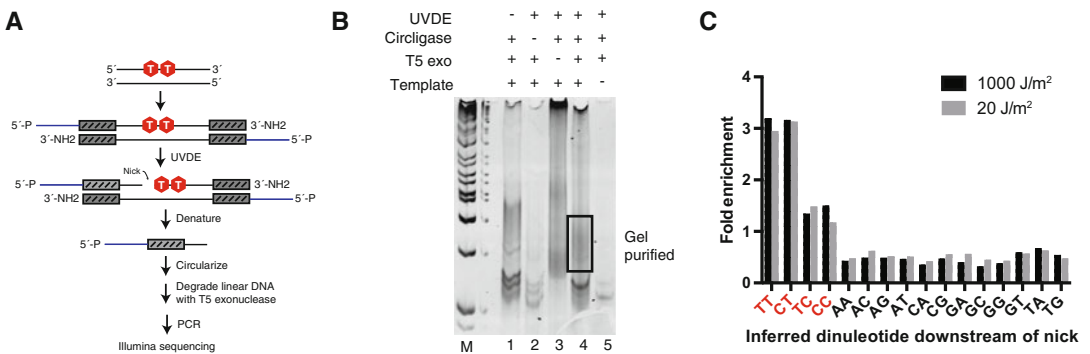
We developed a general method to map modified nucleobases in genomic DNA that couples base excision with next-generation DNA sequencing. We have successfully implemented several variations of the protocol to map different nucleobases or to produce different signals at sites of nucleobase modification<sup>1</sup>. In pre-digestion Excision-seq, samples containing modified nucleobases are first digested with a DNA glycosylase, yielding breaks at the sites of modification (Fig. 1a). Fragments released from DNA digestion are converted into Illumina libraries with standard protocols such that the 5'-end of each sequencing read corresponds to the site of DNA modification. This protocol is applicable when high levels of DNA modification (~1% or more) are expected. In addition, we developed a complementary approach termed “post-digestion” in which libraries are first created from sheared genomic DNA followed by adaptor ligation (Fig. 1b). Prior to PCR, the double-stranded DNA fragments are digested with a base excision repair enzyme, destroying fragments containing modified bases and rendering them incompetent for PCR (Fig. 2). Analysis of reads from these libraries creates a signal opposite to that of “pre-digestion” libraries. Again, this protocol is applicable to DNA with high levels of DNA modification. We used these protocols to map uracil throughout the genomes of *E. coli* and *S. cerevisiae* cells in which dUTP biosynthesis was genetically or pharmacologically perturbed [5].

To map DNA photodimers we modified the pre-digestion protocol to include a photolyase repair step following glycosylase cleavage to remove the photo dimer, which inhibits subsequent Illumina library steps. We used photolyase enzymes specific for cyclobutane pyrimidine dimers or 6-4 photodimers [6] and used this repaired DNA to generate libraries containing sites of UV damage (Fig. 1c). Again, these libraries require high modification levels because the distribution of modification determines the fragment size upon cleavage, and Illumina sequencing requires fragments <800 bp in size.

To address the limitations of the “pre-digestion” and “post-digestion” Excision-seq methods and create libraries using lower levels of genomic damage, we developed an alternative method involving DNA circularization. To generate these libraries, genomic DNA was sheared, polished, a-tailed and ligated to adapters that are appropriate for circularization. DNA was then cleaved with a DNA glycosylase generating a free 3'-OH terminus. These molecules are made single stranded and circularized with circ-ligase. Linear DNA is then degraded with T5 exonuclease and PCR is performed on intact circles (Fig. 2a). We used this protocol to generate libraries from low dose UV damaged cells (Fig. 2b). We were able to see libraries with less damage than with the basic



**Fig. 1** Excision-seq methods for mapping modified nucleobases in genomic DNA. (a) In “pre-digestion” UDG Excision-seq, DNA is sheared with UDG glycosylase and Endo IV leaving double-strand breaks. These fragments are then treated with the standard Illumina polishing and ligation reactions followed by PCR to generate Illumina libraries. Sequences derived from these libraries identify the position of base modifications. (b) In “post-digestion” UDG Excision-seq DNA is sheared mechanically and made into Illumina libraries with the standard polishing and ligation protocols. Libraries containing uracil are then cleaved with UDG and Endo IV and the remaining non-uracil containing library is PCR amplified. (c) In “pre-digestion” UV Excision-seq, UV photodimers are cleaved by the UVDE enzyme followed by the removal of the 5’ dipyrimidine base with either CPD or 6-4 photolyases. These single pyrimidine ends are then compatible with standard Illumina polishing, ligation, and PCR



**Fig. 2** Excision-seq methods for mapping individual modified nucleobases in genomic DNA. (a) In circular UV Excision-seq, DNA is mechanically sheared and ligated to modified circular adapters with standard polishing and ligation reactions. Double-stranded DNAs are then treated with UVDE to generate nicks upstream of the dipyrimidine, yielding a free 3’ OH. The DNA is denatured with heat and the strand with a free 5’ P and 3’ OH is circularized with CirLigase II (Epicentre). The remaining linear DNA is destroyed with T5 exonuclease and circle PCR generates libraries with standard Illumina sequences. (b) Representative circ-ligase preparation with controls including no UVDE (lane 1), no CirLigase II (lane 2), no T5 exonuclease (lane 3), and no template (lane 5). Doubled-stranded from lane 4 was gel purified and submitted for Illumina sequencing. (c) Analysis of low UV dosage sequencing libraries treated with 1000 J/m<sup>2</sup> (black) or 20 J/m<sup>2</sup> (gray) shows bias of dipyrimidine ends relative to genomic dinucleotide content



pre-digestion protocol but lost the specificity of the photodimer type as shown by the prevalence of pyrimidine dimers directly upstream of our sequencing libraries (Fig. 2c).

---

## 2 Materials

### 2.1 General Equipment

1. Micro centrifuge.
2. Size Select E Gel apparatus (Invitrogen).
3. E.Z.N.A. Cycle-Pure DNA kit (Omega Bio-tek).
4. E.Z.N.A. MicroElute Cycle-Pure Kit (Omega Bio-tek).
5. Qubit Fluorimeter.
6. Diagenode bioruptor (or similar DNA shearing apparatus).

### 2.2 Enzymes

1. End Repair kit(NEB).
2. Klenow Exo-.
3. Rapid T4 DNA ligase.
4. Zymolase 20T.
5. Uracil DNA glycosylase.
6. Endonuclease IV.
7. Proteinase K 20 mg/ml.
8. RNase A 10 mg/ml.
9. CircLigase (Epicenter).
10. Platinum Taq.
11. Phusion DNA polymerase (or similar high fidelity polymerase).

### 2.3 Primers

See Tables 1 and 2.

### 2.4 Cell Growth

1. LB for bacterial cultures: 10 g/l Bacto-tryptone, 5 g/l yeast extract, 10 g/l NaCl.
2. YPD for yeast cultures: Bacto peptone 20 g/l, yeast extract 10 g/l, glucose 2% (w/v).

### 2.5 Isolation of DNA and Shearing

1. Buffer Z: 1 M sorbitol, 50 mM Tris pH 8.0, 10 mM BME.
2. Lysis buffer: 1 M sorbitol, 1% SDS, 20 mM EDTA, 10 µg Proteinase K.
3. PBS-S: PBS + 1 M sorbitol.
4. Stop solution: 5% SDS, 100 mM EDTA.
5. Shear buffer: 10% glycerol, TE.
6. Spectrophotometer.
7. Stratalinker with UVC or UVB bulbs (or similar UV crosslinker).

**Table 1**  
**Oligonucleotide sequences**

JH805 BC1	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GAT C*T
JH806 BC1	/5Phos/GAT CAG ATC GGA AGA GCG GTT CAG CAG GAA TGC CGA G
JH807 BC2	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ACT G*T
JH808 BC2	/5Phos/CAG TAG ATC GGA AGA GCG GTT CAG CAG GAA TGC CGA G
JH813 BC5	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GTC A*T
JH814 BC5	/5Phos/TGA CAG ATC GGA AGA GCG GTT CAG CAG GAA TGC CGA G
JH815 BC6	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT ATC G*T
JH816 BC6	/5Phos/CGA TAG ATC GGA AGA GCG GTT CAG CAG GAA TGC CGA G
JH1139 Universal	/5Phos/GAT CGG AAG AGC ACA CGT CT
JH0804 Universal	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TC*T
JH1323 Circle	/5Phos/NNN NNN NNA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGG GAG TTC AGA CGT GTG CTC TTC CGA TCT AGC CAG CGC AGA CCG TGA GGT
JH1324 Circle	/5Phos/CCT CAC GGT CTG CGC TGG CT/3AmMO/

Adapter sequences used for this study. A “\*” denotes a phosphothioate bond. A “3AmMO” is a 3' amino modifier group from IDT used to block this end

8. Radiometer (UVX or similar).
9. 365 nM black light.
10. *E. coli* lysis buffer: 0.006% SDS, 12 µg/ml Proteinase K, 10 µg/ml RNase A.

## **2.6 Photolyase Cleavage**

1. *Vibrio* CPD buffer: 50 mM Tris-HCl pH 7.0, 50 mM NaCl, 5 mM DTT, 1 mM EDTA pH 8.0, 20% glycerol [6].
2. *Xenopus* 6-4 buffer: 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM DTT, 10% (v/v) glycerol [7].
3. *E. coli* CPD buffer: 50 mM Tris-HCl pH 7.0, 50 mM NaCl, 5 mM DTT, 1 mM EDTA pH 8.0, 20% glycerol [8].
4. *A. thaliana* 6-4 photolyase buffer: 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM DTT, 10% (v/v) glycerol [9].

**Table 2**  
**Primer sequences used for this study**

JH801 Universal	CAA GCA GAA GAC GGC ATA CGA GCT CTT CCG ATC T
JH802 Universal	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T
JH1141	CAA GCA GAA GAC GGC ATA CGA GAT CGG TAT <b>CAC</b> GGT GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT C
JH1147	CAA GCA GAA GAC GGC ATA CGA GAT CGG TCG <b>ATG</b> TGT GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT C
JH1151	CAA GCA GAA GAC GGC ATA CGA GAT CGG TGC <b>CAA</b> TGT GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT C

*Bold bases* are barcodes for demultiplexing

### 2.7 DNA Purification

1. Ampure XP beads.
2. Magnetic separator.
3. EBT: EB + 0.05% Tween.
4. DNA cycle pure kits from zymogen.
5. Invitrogen size select gel system.
6. Agilent tape station (optional) with D1000 tapes.
7. Qubit Fluorimeter and HS DNA kit.
8. Illumina HiSeq or MiSeq.

## 3 Methods

### 3.1 Uracil Pre-digestion Excision-seq for *E. coli*

1. CJ1036 *ung-1, dut-1 E. coli* were grown to log phase at 37 °C ( $\sim 2.5 \times 10^{10}$  cells) in Luria Broth (LB) and collected by centrifugation.
2. DNA was harvested following lysis in *E. coli* lysis buffer, extracted with phenol chloroform, and DNA precipitated.
3. High molecular weight DNA was digested with 5 units of UDG (1  $\mu$ l) and 10 (1  $\mu$ l) units of T4 Endonuclease IV for 2 h at 37 °C (*see Note 1*).
4. DNA cleavage was confirmed on a 1% agarose TBE gel looking for appropriate cleavage fragments from 100 to 800 bp. Uracil should make up 1–10% of the genome so the majority of DNA following cleavage should be low molecular weight.
5. Fragments were polished using T4 DNA polymerase (150 units), T4 PNK (50 units) and 1 mM dNTP for 30 min at 20 °C (*see Note 2*).

6. Fragments were cleaned up with Cycle-Pure kits and eluted in 30  $\mu$ l of elution buffer.
7. Fragments were A-tailed with Klenow Exo- fragment (15 units) and 10 mM dATP for 30 min at 37 °C.
8. Fragments were cleaned up with MicroElute Cycle-Pure kits and eluted in 10  $\mu$ l of elution buffer.
9. Fragments were ligated with pre-annealed universal adapters for 15 min at 16 °C with 3000 units of T4 Rapid DNA ligase.
10. Libraries were size selected on 2% size select E gels and fragments between 200 and 400 bp were collected in ddH<sub>2</sub>O.
11. Libraries were PCR amplified using Platinum Taq polymerase (*see Note 3*) and primers containing Illumina 6 bp barcode sequences.
12. Ampure XP beads were used to purify libraries at 1.8 $\times$  concentration and eluted in 30  $\mu$ l EBT.
13. Libraries were quantified on a qubit fluorimeter (*see Note 4*) and submitted for sequencing at 10 nM.

### **3.2 Uracil Pre-digestion Excision-seq for Yeast**

1. 50 ml of YJH516 *dut1-1 unglA* or other yeast strains ( $\sim 6 \times 10^8$  cells) were grown to mid-log phase at 30 °C in YPD and collected by centrifugation.
2. Cells were resuspended in 5 ml of Buffer Z and spheroplasted with 0.35 mg of zymolyase (20T) for 30 min at 30 °C.
3. Spheroplasts were centrifuged and resuspended in 400  $\mu$ l of lysis buffer and incubated O/N at 55 °C.
4. High molecular weight DNA was isolated by phenol/chloroform extraction followed by DNA precipitation (*see Note 5*).
5. 5–10  $\mu$ g of high molecular weight DNA was digested with 5 units of UDG (1  $\mu$ l) and 10 units of T4 Endonuclease IV (1  $\mu$ l) for 2 h at 37 °C.
6. DNA cleavage was confirmed on a 1% agarose TBE gel looking for appropriate cleavage fragments from 100 to 800 bp.
7. Fragments were polished using T4 DNA polymerase (150 units), T4 PNK (50 units), and 1 mM dNTP for 30 min at 20 °C (5  $\mu$ l of end repair kit).
8. Fragments were cleaned up with Cycle-Pure kits and eluted in 30  $\mu$ l of elution buffer.
9. Fragments were A-tailed with Klenow Exo- fragment (15 units/3  $\mu$ l) and 10 mM dATP for 30 min at 37 °C.
10. Fragments were cleaned up with MicroElute Cycle-Pure kits and eluted in 10  $\mu$ l of elution buffer.

11. Fragments were ligated with pre-annealed universal adapters for 15 min at 16 °C with 3000 units of T4 Rapid DNA ligase (5 µl).
12. Libraries were size selected on 2% size select E gels and fragments between 200 and 400 bp were collected in ddH<sub>2</sub>O (*see Note 6*).
13. Libraries were PCR amplified using Platinum Taq polymerase and primers containing Illumina 6 bp barcode sequences.
14. Ampure XP beads were used to purify libraries at 1.8× concentration and eluted in 30 µl EBT.
15. Libraries were quantified on a qubit fluorimeter and submitted for sequencing at 10 nM.

### **3.3 Uracil Post-digestion Excision-seq**

1. 50 ml of YJH516 *dut1-1 unglA* or other yeast strains ( $\sim 6 \times 10^8$  cells) were grown to mid-log phase at 30 °C in YPD and collected by centrifugation.
2. Cells were resuspended in 5 ml of Buffer Z and spheroplasted with 0.35 mg of zymolyase(20T) for 30 min at 30 °C.
3. Spheroplasts were centrifuged and resuspended in 400 µl of lysis buffer and incubated O/N at 55 °C.
4. High molecular weight DNA was isolated by phenol/chloroform extraction followed by DNA precipitation.
5. DNA was mechanically sheared in shearing buffer using a Diagenode bioruptor for 12 min (30 s on/off at maximum intensity) until DNA was mainly sheared to 200–600 bp (*see Note 7*).
6. DNA shearing was confirmed with agarose gel electrophoresis.
7. Fragments were polished using T4 DNA polymerase (150 units), T4 PNK(50 units), and 1 mM dNTP (5 µl of NEB end repair) for 30 min at 20 °C (*see Note 2*).
8. Fragments were cleaned up with Cycle-Pure kits and eluted in 30 µl of elution buffer.
9. Fragments were A-tailed with Klenow Exo- fragment (15 units/3 µl) and 10 mM dATP for 30 min at 37 °C.
10. Fragments were cleaned up with MicroElute Cycle-Pure kits and eluted in 10 µl of elution buffer.
11. Fragments were ligated with pre-annealed universal adapters for 15 min at 16 °C with 3000 units of T4 Rapid DNA ligase (5 µl).
12. Libraries were size selected on 2% size select E gels and fragments between 200 and 400 bp were collected in ddH<sub>2</sub>O.
13. Libraries were PCR amplified using Platinum Taq polymerase and primers containing Illumina 6 bp barcode sequences.

14. PCR libraries containing Uracil were sheared with 5 units of UDG (1  $\mu$ l) and 10 units of T4 Endonuclease IV (1  $\mu$ l) for 2 h at 37 °C.
15. Ampure XP beads were used to purify libraries at 1.8 $\times$  concentration and eluted in 30  $\mu$ l EBT.
16. Libraries were quantified on a qubit fluorimeter and submitted for sequencing at 10 nM.

### **3.4 UV Pre-digestion Excision-seq**

1. 50  $\mu$ l of YJH302 wt yeast were grown to an OD600 of  $\sim$ 0.4 ( $\sim$ 4  $\times$  10<sup>8</sup> cells) and arrested with alpha factor.
2. Cells were resuspended in 5 ml of Buffer Z in dark tubes (*see Note 8*).
3. Cells were irradiated in a 100 mm Petri dish until a specified UV dosage was obtained as measured by a dosimeter (*see Note 9*).
4. 0.35 mg of zymolyase (20T) was added to the cells in the dark tubes for 30 min at 30 °C.
5. Spheroplasts were centrifuged and resuspended in 400  $\mu$ l of lysis buffer and incubated O/N at 55 °C.
6. High molecular weight DNA was isolated by phenol/chloroform extraction followed by DNA precipitation.
7. UV damaged DNA was digested with 10  $\mu$ g of UVDE for 2 h at 30 °C.
8. DNA cleavage was confirmed with agarose gel electrophoresis increased UV dosage leads to increased shearing on the gel.
9. Samples were treated with either CPD or 6-4 photolyase in the appropriate buffer for 2 h under illumination of a 365 nm light ( $\sim$ 280 kJ/m<sup>2</sup>) (*see Note 10*).
10. Fragments were polished using T4 DNA polymerase (150 units), T4 PNK (50 units), and 1 mM dNTP (5  $\mu$ l of NEB end repair kit) for 30 min at 20 °C.
11. Fragments were cleaned up with Cycle-Pure kits and eluted in 30  $\mu$ l of elution buffer.
12. Fragments were A-tailed with Klenow Exo- fragment (15 units/3  $\mu$ l) and 10 mM dATP for 30 min at 37 °C.
13. Fragments were cleaned up with MicroElute Cycle-Pure kits and eluted in 10  $\mu$ l of elution buffer.
14. Fragments were ligated with pre-annealed universal adapters for 15 min at 16 °C with 3000 units of T4 Rapid DNA ligase (5  $\mu$ l).
15. Libraries were size selected on 2% size select E gels and fragments between 300 and 500 bp were collected in ddH<sub>2</sub>O (*see Note 11*).

16. Libraries were PCR amplified using Phusion DNA polymerase (*see Note 12*).
17. Ampure XP beads were used to purify libraries at 1.8× concentration and eluted in 30 µl EBT.
18. Libraries were run on D1000 lane on a TapeStation to validate library size and determine the percentage of library that was adapter dimer (*see Note 13*).
19. Libraries were quantified on a qubit fluorimeter and submitted for sequencing at 10 nM.

### **3.5 Circular Ligation of Low Dose UV Fragments**

1. Anneal circle adapters (*see Note 14*).
2. UV dose and extract DNA as in Subheading 3.4 (steps 1–6).
3. Shear for 15 min on Diagenode bioruptor (30 s on/off at max).
4. DNA cleavage was confirmed with agarose gel electrophoresis expecting a low molecular weight shear from 100 to 1000 base pairs.
5. Fragments were polished using T4 DNA polymerase (150 units), T4 PNK (50 units), and 1 mM dNTP (5 µl of NEB end repair kit) for 30 min at 20 °C.
6. Fragments were cleaned up with Cycle-Pure kits and eluted in 30 µl of elution buffer.
7. Fragments were A-tailed with Klenow Exo- fragment (15 units/3 µl) and 10 mM dATP for 30 min at 37 °C.
8. Fragments were cleaned up with MicroElute Cycle-Pure kits and eluted in 12 µl of elution buffer.
9. Fragments were ligated to pre-annealed circular adaptor 15 min at 16 °C with 3000 units of T4 Rapid DNA ligase (5 µl).
10. Ampure XP beads were used to purify libraries at 1.8× concentration and eluted in 30 µl EBT.
11. DNA was digested with 10 µg of UVDE and 10 units of EndoIV (1 µl) for 2 h at 30 °C.
12. Heat denature 95 °C for 15 min (*see Note 15*) and then keep on ice until circularization.
13. Denatured DNA was circularized with 1 mM dATP, 50 mM MnCl<sub>2</sub>, and 100 units of CircLigase (1 µl) at 60 °C for 1 h.
14. Ampure XP beads were used to purify libraries at 1.8× concentration and eluted in 30 µl EBT.
15. T5 Exonuclease was used to shear noncircular DNA with 50 units of T5 exonuclease (5 µl) in NEB buffer 4 for 2 h at 37 °C.

16. Ampure XP beads were used to purify libraries at  $1.8\times$  concentration and eluted in 30  $\mu$ l EBT.
17. Libraries were PCR amplified using Phusion polymerase (*see Note 16*).
18. Ampure XP beads were used to purify libraries at  $1.8\times$  concentration and eluted in 30  $\mu$ l EBT.
19. Libraries were quantified on a qubit fluorimeter and submitted for sequencing at 10 nM.

### 3.6 Data Analysis for Uracil Modification

1. Sequences were analyzed by alignment to a reference genome (*sacCer1*) using Bowtie 2 [10] and SAMtools [11].
2. Alignments were processed to bedGraph format using BEDTools [12], and visualized in the UCSC Genome Browser [13].
3. Coverage at each position was normalized by the number of reads aligned in the library (i.e., reads per million, RPM). Using this method, the level of coverage at a specific site or region in the genome represents the relative quantity of uracil at that position.
4. Raw and processed sequencing data (FASTQ and bedGraph formats) from **step 2** from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) [15] under accession number GSE51361.

### 3.7 Data Analysis for UV Modification

1. Sequences were aligned to the *S. cerevisiae* genome using Bowtie 2 [10] and SAMtools [11].
2. For the circular libraries the UMI tag was used to remove PCR duplicates from the library using UMIttools [14].
3. Aligned reads were separated by strand, and dinucleotide counts for the 5' ends of the reads were determined.
4. The percentages of each dinucleotide combination were normalized to the mononucleotide and dinucleotide frequencies found in *S. cerevisiae* genomic DNA to account for A:T bias in the genome.

---

## 4 Notes

1. Uracil digestion of 10  $\mu$ g of yeast DNA with 5 units of UDG and 10 unit Endo IV for 2 h at 37 °C was sufficient to yield small molecular weight DNA in strains mutant for repair enzymes as reported in [5]. Increased enzyme or time did not yield additional cleavage indicating that the majority of uracil was cleaved. For other cell types or mutant backgrounds the



cleavage conditions may need to be optimized. Small molecular weight DNA <600 bp is needed for efficient library generation.

2. The end repair module from NEB is used for polishing DNA for the most consistent results. The three enzymes can also be purchased individually.
3. Pfu DNA polymerases and their derivatives cannot be used to create Ex-seq libraries for uracil, because these polymerases are inhibited by uracil in templates.
4. Quantification of libraries for Illumina sequencing in our hands has shown the best results with the Qubit fluorimeter. We get the most accurate clustering using this method. qPCR and other methods are available for quantification and many sequencing companies offer this service.
5. Vortexing should not be done at any stage in this protocol as it shears high molecular weight DNA. All mixing should be done by hand optionally with a cut pipette tip.
6. We prefer using size select gels to purify adapter ligated libraries prior to PCR. This allows for the removal of excess adapters which will PCR downstream, as well as allowing for the isolation of a variety of fragment sizes. We use the different fragments to find the optimal library in terms of background bands and size. Also the ability to see a size selected ladder of DNA vs. a randomly sized smear can help determine real libraries from artifacts.
7. Biorupting is very cell specific with different mutants and cell types requiring different amounts of biorupting time. Prior to performing library preparation, genomic DNA as prepared above should be sheared for a variety of time points and about 10% should be analyzed via agarose gel. Smears between ~200 bp and ~1000 bp are optimal and shear time can be optimized to achieve that. Hard to shear samples can also be sheared in TE with no glycerol or in special Eppendorf tubes that improve shearing.
8. For UV experiments all the steps following UV irradiation are performed in dark tubes due to the photolyases present in wild-type cells that can repair UV photodimers upon exposure to light. Once DNA has been obtained this is no longer required.
9. UV exposure was measured with a dosimeter within the Stratalinker to confirm the exposure we saw slight variation between the Stratalinker dosage and the actual dosage given.
10. UV libraries were repaired by photolyase prior to PCR because nothing else tried allowed PCR through the adduct on the 5' end of our molecules. UVA light was used to activate these enzymes by placing the tubes within an inch of a black light.

The UVA dosage was measured by a dosimeter as  $>200$  kJ and should be sufficient for enzyme activity *in vitro*.

11. UV libraries were isolated at a higher molecular weight from size select gels due to less efficient shearing by UVDE. PCR from higher molecular weight fractions contained less dimer and more specific libraries. Also performing a laddering PCR with a series of fractions was helpful to determine real libraries from artifacts (which all PCR at the same size).
12. Phusion Polymerase was used to PCR UV libraries because both high fidelity and high activity were needed to amplify these low copy libraries.
13. Tapestation analysis helped us to confirm that these libraries contained mainly  $>400$  bp DNA fragments. Since the input is low in these libraries there is a high chance for Illumina adapter contamination to occur. We see this in our analysis so try to submit samples with low levels of adapter contamination.
14. Circular adapters have long overhanging ends and as such we perform a slower annealing process and make them relatively fresh to maintain their annealing. We have also tried adding a spacer to improve the adaptors to prevent rolling circle PCR.
15. Heat for 15 min or use 0.3 M final concentration of KOH followed by neutralization and precipitation can be used to denature the DNA.
16. To determine the yield of a circle library a variety of controls are useful. A no T5 Exonuclease step gives a strong smear following PCR as an indication of what a robust library amplification would look like. No UV and no CircLigase negative controls are useful for troubleshooting nonspecific amplification.

---

## Acknowledgments

We thank A. Sancar and D. Zhang for graciously providing purified photolyase enzymes during the development of the Excision-seq protocol. This work was supported in part by a March of Dimes Basil O'Connor Research Grant (J.H.), a Damon Runyon-Rachleff Innovation award (J.H.), a Research Scholar Grant from the American Cancer Society (RSG-13-216-01-DMC), and a Department of Defense Visionary Postdoc Award (to M.R., W81XWH-12-1-0333).

## References

1. Kumari S, Rastogi RP et al (2008) DNA damage: detection strategies. *EXCLI J* 7:44–62
2. Clark TA, Spittle KE, Turner SW et al (2011) Direct detection and sequencing of damaged DNA bases. *Genome Integr* 2:10
3. Baute J, Depicker A (2008) Base excision repair and its role in maintaining genome stability. *Crit Rev Biochem Mol Biol* 43:239–276
4. Hegde ML, Hazra TK, Mitra S (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res* 18:27–47
5. Bryan DS, Ransom M, Adane B et al (2014) High resolution mapping of modified DNA nucleobases using excision repair enzymes. *Genome Res* 24:1534–1542
6. Sancar GB, Sancar A (2006) Purification and characterization of DNA photolyases. *Methods Enzymol* 408:121–156
7. Ryoji M, Katayama H, Fusamae H et al (1996) Repair of DNA damage in a mitochondrial lysate of *Xenopus laevis* oocytes. *Nucleic Acids Res* 24:4057–4062
8. Liu Z, Tan C, Guo X et al (2011) Dynamics and mechanism of cyclobutane pyrimidine dimer repair by DNA photolyase. *Proc Natl Acad Sci U S A* 108:14831–14836
9. Li J, Liu Z, Tan C et al (2010) Dynamics and mechanism of repair of ultraviolet-induced (6-4) photoproduct by photolyase. *Nature* 466:887–890
10. Langmead B, Trapnell C, Pop M et al (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25
11. Li H, Handsaker B, Wysoker A et al (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079
12. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842
13. Karolchik D, Hinrichs AS, Kent WJ (2001) The UCSC Genome Browser. Wiley, Hoboken, NJ
14. Smith TS, Heger A, Sudbery I (2016) UMI-tools: modelling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *bioRxiv* 051755
15. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A (2012) NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research* 41 (D1): D991–D995

# Chapter 7

## Integrated Microarray-based Tools for Detection of Genomic DNA Damage and Repair Mechanisms

Patrick van Eijk, Yumin Teng, Mark R. Bennet, Katie E. Evans, James R. Powell, Richard M. Webster, and Simon H. Reed

### Abstract

The genetic information contained within the DNA molecule is highly susceptible to chemical and physical insult, caused by both endogenous and exogenous sources that can generate in the order of thousands of lesions a day in each of our cells (Lindahl, *Nature* 362(6422):709–715, 1993). DNA damages interfere with DNA metabolic processes such as transcription and replication and can be potent inhibitors of cell division and gene expression. To combat these regular threats to genome stability, a host of DNA repair mechanisms have evolved. When DNA lesions are left unrepaired due to defects in the repair pathway, mutations can arise that may alter the genetic information of the cell. DNA repair is thus fundamental to genome stability and defects in all the major repair pathways can lead to cancer predisposition. Therefore, the ability to accurately measure DNA damage at a genomic scale and determine the level, position, and rates of removal by DNA repair can contribute greatly to our understanding of how DNA repair in chromatin is organized throughout the genome. For this reason, we developed the 3D-DIP-Chip protocol described in this chapter. Conducting such measurements has potential applications in a variety of other fields, such as genotoxicity testing and cancer treatment using DNA damage inducing chemotherapy. Being able to detect and measure genomic DNA damage and repair patterns in individuals following treatment with chemotherapy could enable personalized medicine by predicting response to therapy.

**Key words** Genome stability, Nucleotide excision repair, Microarrays, 3D-DIP-Chip, DNA damage, DNA repair, Chemotherapy, Genotoxicity

---

## 1 Introduction

We have developed a method that we refer to as 3D-DIP-Chip (DNA Damage Detection by DNA Immuno-Precipitation (DIP) on microarray chips, [2]) to detect genomic DNA damage and measure lesion removal on a genome-wide scale. The method is based on the standard ChIP-chip workflow but relies on the immunoprecipitation of DNA lesions within naked DNA. We will provide details for the immunoprecipitation of three types of DNA damage (UV, cisplatin, and oxaliplatin), but the protocol can be adapted to any type of aberrant DNA base or DNA damage granted a specific

antibody against that epitope is available. Similarly, procedures for the use of chromatin and/or DNA from both budding yeast, *Saccharomyces cerevisiae*, and cultured human cell lines will be provided.

In order to analyze multiple inter- and intra-experimental datasets of this new type of repair and indeed standard ChIP-chip data, we had to develop software tools to enable this. Generally, routine ChIP-chip experiments provide genome-wide chromatin protein occupancy information through peak detection of the measured signal. This traditionally allows for binary comparison of data sets and methods are available for comparative analysis of this type of data. However, biological systems naturally contain much more information beyond that which can be described using a binary concept of presence or absence. Similarly, the nature of the abundant and heterogeneous distribution of DNA lesions such as UV-induced Cyclobutane Pyrimidine Dimers (CPDs) is not suitable for this classical approach to data analysis. This potential for more advanced analysis we have tried to fulfil by developing the Sandcastle R software package—Software for the Analysis and Normalization of Data from ChIP-chip AssayS of Two or more Linked Experiments [3]—that is an integral part of the microarray-based genomic toolset described here. This software can be used for normalization, analysis, and visual description of data in 3D-DIP-Chip and standard ChIP-chip datasets. The Sandcastle normalization procedure, described in more detail below, uniquely facilitates the comparative analysis of datasets generated under different experimental conditions, allowing the extraction of latent, biologically relevant, information.

We have used this extended workflow consisting of 3D-DIP-Chip, ChIP-chip, and Sandcastle's data analysis algorithms, to study the mechanisms and the organization of DNA repair and DNA damage-induced chromatin remodeling during Global Genome Nucleotide Excision Repair (GG-NER) in yeast [4]. This was achieved by measuring: (1) repair using 3D-DIP-Chip and UV-induced changes in chromatin occupancy of (2) DNA repair proteins, (3) chromatin remodelers, and (4) chromatin modifications [4]. The introduction of specific mutations or deletion of relevant factors can provide important mechanistic insights into the change in genome-wide chromatin occupancy, redistribution of proteins and protein complexes, and changes in relative DNA repair rates in response to DNA damage. The comparison of datasets derived from different mutant strains or data from different treated versus untreated cells can now be compared using Sandcastle to derive mechanisms of DNA repair on a genomic scale.

The ability to detect and measure DNA damage and genomic repair spectra is becoming increasingly important as recent advances in DNA sequencing technologies now permit whole genome mutation analyses [5, 6]. These sequencing projects have

revealed the presence of mutational signatures captured in individual cancer genomes derived from different types of cancer cells [7]. Efforts to determine the relationship between the distribution of genomic DNA repair rates with the mutational end-points detected in cancer genomes are emerging [8]. Progress in this field would enable the use of genomic DNA damage and repair measurements as early indicators and predictive biomarkers, including markers that predict predisposition to, and the progression of, disease, as well as the ability to determine the response of individuals to drug treatments of these diseases.

---

## 2 Materials

### 2.1 *Yeast Cell Culture and Media*

1. Glassware, relevant equipment, media, and other reagents, where necessary, were sterilized by autoclaving at 121 °C for 15 min.
2. Yeast Extract Peptone Dextrose (YPD): 2% glucose, 2% peptone; 1% yeast extract.
3. Glass beads, acid-washed (425–600 μm).

### 2.2 *Human Cell Culture and Media*

1. Dulbecco's Modified Eagle Medium (DMEM).
2. Fetal calf serum.
3. 1% L-Glutamine.
4. 1% penicillin streptomycin solution (e.g., Sigma-Aldrich).
5. Trypsin-EDTA solution (e.g., Gibco by Life Technologies, Invitrogen).

### 2.3 *DNA Purification Buffers and Solutions*

1. Sorbitol-TE: 0.9 M Sorbitol, 0.1 M Tris-HCl, pH 8.0, 0.1 M EDTA.
2. Zymolyase 20 T, 20,500 U/g, Amsbio or similar (10 mg/ml in sorbitol-TE solution).
3. β-mercaptoethanol.
4. RNase A (10 mg/ml stock).
5. 10% SDS.
6. DNA lysis buffer: 4 M Urea, 200 mM NaCl, 100 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS.
7. Pronase (20 mg/ml in TE buffer) (e.g., Roche).
8. Phenol/Chloroform 1:1 (v/v).
9. Prechilled -20 °C 100% Ethanol.
10. 100% Isopropanol.
11. TE: 10 mM Tris-HCl pH 8.0, 1 mM EDTA.

## **2.4 ChIP/IP Buffers, Solutions, and Reagents**

1. Rat anti-cisplatin modified DNA, CP9/19, antibody (ab103261, Abcam).
2. Mouse Anti-CPD antibody (CAC-NM-DND-001, COSMO Bioscience).
3. Dynabeads (Invitrogen): Sheep anti-Rat IgG, Goat anti-Mouse and ProtG dynabeads.
4. PBS-BSA: solutions of 0.1% BSA (1 mg/ml) and 1% BSA (10 mg/ml) in PBS.
5. PMSF (200 mM in Isopropanol).
6. FA/SDS: 50 mM HEPES–KOH pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton-X, 0.1% Na deoxycholate, 0.1% SDS (1 mM PMSF added fresh before use, complete protease inhibitor cocktails are optional).
7. FA/SDS 0.5 M NaCl.
8. FA/SDS 1.0 M NaCl.
9. LiCl IP buffer: 10 mM Tris–HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% NP40, 0.5% Na deoxycholate.
10. Pronase buffer: 125 mM Tris–HCl pH 7.5, 25 mM EDTA, 2.5% SDS.
11. ChIP lysis buffer: 10 mM Tris–HCl pH 8, 10 mM NaCl, 0.5% NP-40.
12. ChIP shearing buffer/SDS lysis buffer: 50 mM Tris–HCl pH 8, 10 mM EDTA pH 8, 1% SDS.
13. Protease Inhibitor cocktail (Roche, product code 11697498001).
14. ddH<sub>2</sub>O.

## **2.5 Reagents**

1. PreCR DNA repair kit (New England Biolabs).
2. T4 DNA polymerase 3 U/μl (New England Biolabs).
3. T4 DNA ligase 400 U/μl (New England Biolabs).
4. 10× NEBuffer 2 (T4 DNA Pol buffer).
5. Q5 DNA Polymerase 2 U/μl (New England Biolabs).
6. dNTPs (10 mM).
7. PureLink™ Quick PCR Purification Kit (Invitrogen).
8. iTaq™ Universal SYBR® Green Supermix (Bio-Rad).
9. Hard-Shell® 96-well semi-skirted PCR plate (Bio-Rad).
10. BioPrime Total Genome Labelling System (Invitrogen).
11. GenomePlex Complete Whole Genome Amplification kit (WGA2, Sigma-Aldrich).
12. Human Cot-1 DNA (1.0 mg/ml) (Agilent).
13. 10× Agilent Blocking Agent.
14. 2× Agilent Hybridization Buffer.

## 2.6 Equipment

1. Cell counting chamber (e.g., Neubauer by Hawksley).
2. Shaking incubator for basic microbiology cell culture (e.g., Infors HT multitron).
3. Diagenode BioRuptor (Diagenode).
4. Water bath (65 °C) and circulating cooled water bath (i.e., NESLAB Digital Plus RTE7, ThermoScientific).
5. Refrigerated mid-speed centrifuge and rotor for large volumes of 100–500 ml (e.g., Beckman-Coulter centrifuge with JA-10 and JA-20 rotor).
6. Refrigerated microcentrifuge (e.g., Beckman-Coulter Microfuge 22R or Eppendorf Microfuge 5424R).
7. Refrigerated Bench-top Eppendorf centrifuge 5810R.
8. Bench-top Eppendorf centrifuge 5804R and A-2-W-ADP rotor for 96-well plates.
9. Thermomixer (Eppendorf).
10. UV light source, VL-21G mineralight lamp (UV products, San Gabriel, CA, USA).
11. NanoDrop-1000, NanoDrop Technologies, ThermoScientific.
12. Quantitative real-time PCR equipment (e.g., CFX Connect™ Real-Time PCR detection system, Bio-Rad).
13. CFX Manager™ Software v3.1 (Bio-Rad).
14. Agilent 2200 TapeStation (or Bioanalyzer) and accompanying D1000 screentapes.

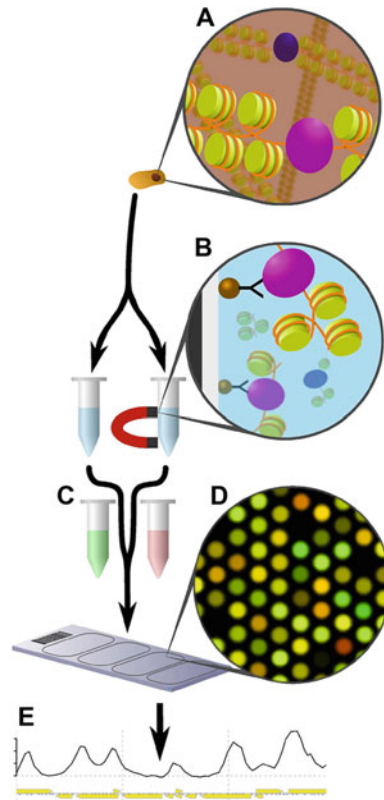
---

## 3 Methods

### 3.1 3D-DIP-Chip Procedure Overview

The 3D-DIP-Chip procedure follows the general steps of a traditional ChIP-chip workflow (Fig. 1). However, the immunoprecipitation (IP) is performed using naked DNA instead of chromatin. Hence, the initial purification of DNA (not chromatin) and shearing of the DNA are adapted for the use of naked DNA (*see* Subheading 3 for more details). In short, DNA is purified from cells treated with a DNA damaging agent and fragmented by sonication. Antibodies raised against the DNA lesion of interest are used to IP the DNA. At this stage input samples are also included for comparison, as is common practice in any ChIP-chip workflow. Next, the DNA is purified and the DNA damages are reversed so as not to inhibit amplification. Purified input and IP DNA is amplified and differentially labeled using two dyes and applied to the appropriate DNA microarray. Following the hybridization, the arrays are washed and the fluorescent signal from the dye is read out using an optical scanner. The ratio of the red and green fluorescence at each feature of the array represents the enrichment for the DNA





**Fig. 1** Representation of the ChIP-chip procedure. Proteins are crosslinked to chromatin (a) which is extracted, sonicated, and split into two samples. IP is carried out on one sample to separate out the chromatin bound to the factor of interest (b). Both the samples are purified to DNA, amplified by PCR and differentially labeled (c). They are allowed to hybridize to the microarray probes and the resulting intensity values from the scanned image (d) are converted to numerical values that can be plotted (e) and processed as required by the investigation. Figure adapted from [3]

damage detected and is converted into numerical values during a process referred to as feature extraction. The data as such are now ready for data analysis and are loaded into the R statistical programming environment for further investigation using a bespoke software package that we developed called Sandcastle [3, 9]. We describe the typical workflow for our 3D-DIP-Chip method for the detection of UV-induced CPD lesions as well as platinum adducts, resulting from cisplatin and oxaliplatin treatment of both yeast and human cells. Isolation of DNA and chromatin is described for yeast and human cells. These are followed by procedures for shearing the DNA, immunoprecipitation, and quantification by real-time PCR and microarray analysis. Finally, we present our

data analysis workflow. Our workflow is not limited to 3D-DIP-Chip but can also be applied to conventional ChIP-chip, provided that a chromatin preparation is used.

### **3.2 Yeast Cell Culture, Growth, Treatment, and Harvesting**

Work described here uses the haploid BY4742 strain from Euro-Scarf [10]. All work described in this protocol refers to this strain and mutants are derived from this background, unless stated otherwise. Yeast cell growth and handling were all performed under sterile conditions and cells were grown on Yeast Extract Peptone Dextrose (YPD) agar plates at 30 °C and stored at 4 °C for day-to-day use and short-term storage. Yeast grown on a plate was used for inoculating liquid precultures that were incubated in an Infors HT multitron standard incubator shaking at 180 rpm. Cell counting was performed on a Neubauer cell counting chamber (Hawksley) and optical density (OD<sub>600</sub>) was measured using a Jenway 6300 spectrophotometer (Bibby Scientific Ltd.) (*see Note 1* for details). Centrifugation of large liquid yeast cultures and cell suspensions was performed using a Beckman-Coulter centrifuge and JA-10 rotor at 4 °C unless stated otherwise.

1. Inoculate a 10 ml preculture of the relevant strain and incubate overnight. When the cells reach stationary phase the next day they can be stored at 4 °C.
2. A calculated amount of this preculture can now be used to inoculate a large volume culture. Account for 100 ml of culture per sample and multiply this by the total number of samples needed to establish the final volume of culture required. Follow the growth by calculating the cell density or measuring optical density (OD<sub>600</sub>) regularly, until the cells reach logarithmic-phase and a cell density of  $2 \times 10^7$  cells/ml (or OD<sub>600</sub> = 0.6–1.0).
3. Collect the cells by centrifugation at  $4000 \times g$  for 5 min at 4 °C and resuspend them in prechilled phosphate-buffered saline (PBS) at a cell density of  $2 \times 10^7$  cells/ml. Continue to **step 4** for platinum treatment or **step 5** for UV treatment.
4. Treatment of cells with platinum: Cells were treated with 2.5 mM cisplatin or oxaliplatin for 4 h and following treatment resuspended in YPD for the required repair time.
5. Treatment of cells with UV irradiation was carried out using a 254 nm UVC light (UV products, San Gabriel, CA, USA) at a dose of 100 J/m<sup>2</sup>. Set aside 100 ml of cell suspension before UV irradiation as the nonirradiated (–UV) control. Fifty milliliters of cell suspension were irradiated in a Pyrex dish ( $\Phi = 14$  cm). This process was repeated for the remainder of the cell suspension. Keep the irradiated cells in the dark in a sterile flask (to avoid photoreactivation of CPDs) and set aside a single volume of 100 ml cell suspension in PBS as the zero repair-time sample. Collect the remainder of the cells by

centrifugation ( $4000 \times g$  for 5 min) and resuspend them in fresh YPD at  $2 \times 10^7$  cells/ml and incubate them at  $30^\circ\text{C}$  for the desired repair time (e.g., 2 h).

6. For 3D-DIP-chip samples, collect the cells after the repair time by centrifugation at  $4000 \times g$  for 5 min at  $4^\circ\text{C}$  and resuspend them in ice-cold PBS and wash the cells twice. After the final wash step, the cell pellet can now be snap-frozen using liquid nitrogen and stored at  $-80^\circ\text{C}$  (*see Note 2*).
7. For standard ChIP-chip samples:
  - (a) Crosslink cells for 10 min at room temperature with continuous shaking using 3.3 ml of 37% formaldehyde for 100 ml culture. The reaction can be quenched using 5.5 ml of 2.5 M glycine for 5 min at room temperature with continued shaking. Cells are collected by centrifugation at  $4000 \times g$  for 5 min at  $4^\circ\text{C}$  and washed twice in 40 ml of ice-cold PBS.
  - (b) After the second wash step resuspend the pellet in 1 ml FA/SDS (1 mM PMSF) and transfer the suspension to a 2 ml microcentrifuge tube.
  - (c) Spin down the cells by centrifugation for 5 min at 13,500 rpm ( $13,500 \times g$ ) at  $4^\circ\text{C}$  (Beckman-Coulter Microfuge 22R). Remove the supernatant and snap-freeze the pellet using liquid nitrogen and store at  $-80^\circ\text{C}$  (*see Note 2*).

### 3.3 Human Cell Culture

Human cell culture was performed under standard sterile conditions and all equipment was autoclaved at  $121^\circ\text{C}$  for 15 min, where applicable. We have successfully applied our 3D-DIP-Chip method using immortalized human dermal fibroblasts (AG16409, Coriell Cell Repository, Camden, NJ, USA).

Fibroblasts were grown in DMEM containing 10% foetal calf serum, 1% L-Glutamine, and 1% penicillin-streptomycin (Sigma-Aldrich) as monolayers and incubated at  $37^\circ\text{C}$  in a humidified incubator under 5%  $\text{CO}_2$ . We used T75  $\text{cm}^2$  or T25  $\text{cm}^2$  cell culture flasks with 10 ml or 5 ml of growth media, respectively, to grow the cells. Fresh growth media is replaced approximately every 3 days and cells are reseeded every 4 weeks. The fibroblasts were passaged close to confluence ( $\sim 80\text{--}90\%$ ) every 7–10 days, after which they were trypsinized and  $2 \times 10^5$  to  $4 \times 10^5$  cells were reseeded.

1. Remove growth media from the flask and wash the cells with 2 ml of pre-warmed PBS.
2. Add 2 ml of  $1 \times$  trypsin-EDTA and incubate at  $37^\circ\text{C}$  for 15–30 min.
3. Resuspend the cells in the required amount of growth media for reseeded.

4. Count the cells using an improved Neubauer hemocytometer (Hawksley) using 10  $\mu$ l of cell suspension. For these cells the population doubling (PD) was  $\sim$ 0.4 PD/day and this remained constant throughout these experiments.
5. Seed  $2 \times 10^5$  to  $4 \times 10^5$  fibroblast cells (AG16409) on a tissue culture dish (100  $\times$  20 mm) and grow for 7–10 days to obtain  $\sim 2 \times 10^6$  cells. Alternatively, use a larger tissue culture dish (150  $\times$  25 mm) to yield  $4\text{--}5 \times 10^6$  cells after 10 days of incubation.
6. Grow cells to confluence before treatment.
7. Remove growth media and wash the cells twice with 2 ml of warm PBS (5 ml for larger dishes). The cells are now ready for treatment with platinum-based chemotherapeutics (*see step 8*) or UV irradiation (*see step 10*).

### **3.4 Platinum Treatment of Human Cells**

1. Treat cells with an end-concentration of 2.5 mM cisplatin and 2.5 mM oxaliplatin in serum-free media. Add the serum-free media containing the drugs and incubate the dishes at 37  $^{\circ}$ C for 2 h.
2. In addition to the treated cells, incubate the untreated cells in serum-free media for 2 h at 37  $^{\circ}$ C as well.

### **3.5 UV Treatment of Human Cells**

1. Add 5 ml PBS to the culture dish and set aside a 5 ml PBS cell suspension as untreated control.
2. Irradiate cells with 50 J/m<sup>2</sup> using a 254 nm UVC lamp (UV products, San Gabriel).

### **3.6 Human Cell Harvesting and DNA Isolation**

1. Harvest the cells after treatment using  $1 \times$  trypsin-EDTA and add PBS to the cells suspension to transfer the cells to a 15 ml Falcon tube.
2. Pellet the cells by centrifugation at 1000 rpm ( $200 \times g$ ) for 5 min at room temperature (Eppendorf centrifuge 5810R).
3. Remove the supernatant and continue immediately with DNA extraction using DNeasy Blood & Tissue Kit (Qiagen No. 69504). The DNA is now ready for sonication (*see Subheading 3.7*).

### **3.7 Human Chromatin Preparation from Nonadherent Cells**

When chromatin from human cell culture is required, the following protocol describes the cross-linking and cell lysis necessary to prepare chromatin for a human ChIP-chip workflow. We have successfully used this approach for TK6 cells and HT1080 cells. We won't go into detail on the maintenance and propagation of these cells as this is outside the scope of this method. The procedure outlined here was optimized for TK6 cells and uses four to eight T175 flasks grown to a density of  $\sim 5 \times 10^6$  cells/ml.

1. Spin down the cells at  $1000 \times g$  for 5 min at room temperature.
2. Resuspend the pellet in PBS maintaining the original cell density of  $\sim 5 \times 10^6$  cells/ml and transfer the suspension to a 2 l Erlenmeyer flask.
3. Treat the cells with 1% formaldehyde for cross-linking on an orbital shaker at room temperature for 8 min.
4. Quench the cross-linking reaction for 5 min at room temperature by adding 2.5 M glycine to an end concentration of 125 mM (*see Note 3*).
5. Cool the cells on ice and transfer the cell suspension to 50 ml tubes.
6. Centrifuge the cells at  $1000 \times g$  for 5 min at 4 °C. Keep the cells on ice during the following steps.
7. Wash the cells three times in cold PBS, pellet the cells each time by centrifugation at  $1000 \times g$  for 5 min at 4 °C.
8. Efficient shearing of TK6 cell-derived chromatin requires two lysis steps: resuspend the cells in 1 ml ( $\sim 1 \times 10^7$  cells/ml) ChIP lysis buffer, transfer the cell suspension to a 1.5 ml microcentrifuge tube, and lyse the cells by incubating them on ice for 10 min (*see Note 4*).
9. Spin down the “nuclear pellet” by centrifugation at  $1000 \times g$  for 5 min at 4 °C and remove the supernatant.
10. Resuspend pellet in 330  $\mu$ l ChIP shearing buffer. The suspension is now ready for sonication (*see Subheading 3.11*).
11. Load six tubes of chromatin for shearing and sonicate for 7 cycles of 30s on/off using “High” settings.
12. Spin down the samples at  $15,000 \times g$  for 10 min at 4 °C to remove any cell debris.
13. Chromatin preps from similar treatments can now be pooled and aliquoted for storage.
14. Snap freeze the chromatin preps using liquid nitrogen and store the samples at  $-80$  °C.

### **3.8 Human Chromatin Preparation from Adherent Cells**

1. Grow cells for chromatin extraction on 10 cm dishes and use three plates with cells grown to 80% confluence for a single large scale ChIP (yields  $\sim 200$   $\mu$ g of chromatin).
2. Cross-link the cells by formaldehyde treatment. Add 1% final concentration of FA and incubate the dishes for 3–5 min at room temperature on an orbital shaker.
3. Quench the reaction by adding 2.5 M Glycine to a final concentration of 125 mM and incubate the cell suspension for 5–10 min at room temperature.

4. Discard the media (HT1080 cells are adherent) after quenching and cross-linking and wash the cells on plate by adding 10 ml of ice-cold PBS to the corner of the dish and swirl. Avoid applying large volumes of liquid directly onto the plate to prevent the cells from peeling off.
5. Scrap the cells from the plate into 10–30 ml PBS containing a protease inhibitor cocktail.
6. Spin down the cells in tabletop centrifuge at 1200 rpm ( $290 \times g$ ) for 5 min at 4 °C.
7. Discard the supernatant and take care to remove any residual PBS because it will interfere with the lysis buffer.
8. Resuspend the pellet in SDS lysis buffer supplemented with protease inhibitor cocktail to  $1\text{--}1.5 \times 10^7$  cells/ml and aliquot in microcentrifuge tubes in volumes of 300  $\mu$ l.
9. Lyse the cells by incubating the suspension on ice for 20–30 min.
10. Shear the chromatin by sonication using a Diagenode BioRuptor. Sonicate for 5 min using ten cycles at the “high” setting cycling for 30 s on 20 s off. Place the samples on ice for a couple of minutes (*see* Subheading 3.11).
11. Spin down the sonicated chromatin prep for 10 min at 13,000 rpm ( $16,000 \times g$ ) at 4 °C.
12. Chromatin preps from similar treatments can now be pooled and aliquoted for storage.
13. Snap freeze the chromatin preps using liquid nitrogen and store the samples at  $-80$  °C.

### 3.9 Yeast DNA Preparation

1. Resuspend the untreated, platinum-treated, or UV-treated cell pellet in 5 ml of sorbitol-TE solution containing 5 mg Zymolyase and 10  $\mu$ l  $\beta$ -mercaptoethanol to the samples and mix by shaking.
2. Spheroplasts are generated by incubating the cells at 4 °C overnight or at 37 °C for 30–60 min with intermittent shaking. Spheroplast production can be monitored under a microscope (*see* Note 5).
3. Gentle centrifugation at 3000 rpm ( $800 \times g$ ) for 5 min at 4 °C will pellet the spheroplasts, which can then be washed once in sorbitol-TE and resuspended in 5 ml of DNA lysis buffer/PBS 1:1 (v/v).
4. Add 200  $\mu$ l RNase A and mix by vortexing. Incubate the samples at 37 °C for 1 h with occasional shaking.
5. Add 200  $\mu$ l pronase (20 mg/ml) and incubate the samples at 37 °C for 1 h, follow by 1 h incubation at 65 °C with occasional shaking.

6. Perform two phenol/chloroform extractions to guarantee complete deproteinization. Add 6 ml phenol/chloroform to the cell lysate (6 ml) and vortexing vigorously until homogeneous.
7. Spin down the sample at 10,000 rpm ( $12,000 \times g$ ) (Beckman-Coulter centrifuge, JA-20 rotor) for 10 min and transfer of the upper aqueous phase to a 15 ml Falcon tube.
8. Repeat **steps 7 and 8** for the second phenol/chloroform extraction.
9. Precipitate the DNA by adding two volumes (12 ml) of cold 100% ethanol ( $-20\text{ }^{\circ}\text{C}$ ) with gentle mixing and incubate overnight at  $-20\text{ }^{\circ}\text{C}$  (or 30 min at  $-80\text{ }^{\circ}\text{C}$ ).
10. Recover the DNA precipitate by centrifuging for 15 min at 4000 rpm ( $3200 \times g$ ) at  $4\text{ }^{\circ}\text{C}$  and resuspend the pellet in 1 ml TE. Allow the DNA to dissolve completely. This may take time.
11. Precipitate the DNA again using isopropanol. Add 1 ml of prechilled 100% isopropanol ( $-20\text{ }^{\circ}\text{C}$ ) and shake gently (genomic DNA is sensitive to shearing). Incubate at room temperature for 30 min with occasional, gentle mixing. Collect the DNA precipitate carefully using a pipette tip and resuspend in 1 ml TE.
12. Check the DNA for purity by agarose gel electrophoresis and quantify the DNA (i.e., using a NanoDrop UV spectrophotometer). The DNA is now ready for sonication (*see* Subheading 3.11).

### **3.10 Yeast Chromatin Preparation**

1. Thaw the cell pellet on ice and resuspended in 0.5 ml FA/SDS (1 mM PMSF) in a 2 ml microcentrifuge tube.
2. Lyse the cells by bead beating. Add 500  $\mu\text{l}$  of glass beads and vortex for 10 min at  $4\text{ }^{\circ}\text{C}$ .
3. Retrieve the cell lysate by puncturing a hole in the bottom of the 2 ml tube by using a hot needle.
4. Place the punctured tube in a 15 ml Falcon tube and centrifuge at 2000 rpm ( $300 \times g$ ) for 2 min at  $4\text{ }^{\circ}\text{C}$  (Eppendorf centrifuge 5810R).
5. Wash the glass beads with 500  $\mu\text{l}$  of FA/SDS (1 mM PMSF) and centrifuge again at 2000 rpm ( $300 \times g$ ) for 2 min at  $4\text{ }^{\circ}\text{C}$ .
6. Transfer the lysate to a fresh 2 ml tube and centrifuge at 4000 rpm ( $3200 \times g$ ) for 20 min at  $4\text{ }^{\circ}\text{C}$  (Beckman-Coulter Microfuge 22R or Eppendorf). This removes any non-crosslinked or soluble proteins.
7. Discard the supernatant and resuspend the pellet in 1 ml FA/SDS (1 mM PMSF). Split the chromatin prep over two 2 ml tubes containing 500  $\mu\text{l}$  each.

8. Sonicate the lysate using a BioRuptor sonicator (Diagenode) on “high” power settings for eight cycles of 30 s on and 30 s off at 4 °C.
9. Spin down the sonicated chromatin at 13,000 rpm ( $16,000 \times g$ ) for 10 min at 4 °C to remove cell debris.
10. Recover the supernatant in a fresh microtube and centrifuge again at max speed for 10 min at 4 °C (Beckman-Coulter Microfuge 22R). Collect the supernatant and use it to prepare an input sample (*see* Subheading 3.12, **step 12**).
11. Check the DNA for purity by agarose gel electrophoresis and quantify the DNA (i.e., using a NanoDrop UV spectrophotometer).
12. Snap freeze the chromatin prep in liquid nitrogen to prepare it for storage at –80 °C. The chromatin is now ready for immunoprecipitation, *see* Subheading 3.8.

### 3.11 Sonication of Chromatin and DNA

We use a Diagenode BioRuptor for sonication of both chromatin and naked DNA. The settings described in this protocol are standardized for four 2.0 ml microcentrifuge tubes containing 200  $\mu$ l of chromatin or DNA leaving two slots open in the sonicator's tube holder, unless stated otherwise (*see* below and individual sections for more details).

1. Human DNA is to be diluted to 100 ng/ $\mu$ l for optimal shearing, yeast DNA is sonicated using a 200 ng/ $\mu$ l concentration.
2. Samples are sonicated at “high” power settings at 4 °C using a circulating cooled water bath.
3. The duration and number of on-and-off cycles were calibrated based on the source material (DNA versus chromatin). The following numbers are a guideline (*see* **Note 6**):

Yeast DNA	20 cycles	30' on 30' off	~300–500 bp	
Yeast chromatin	8 cycles	30' on 30' off	~500 bp	
Human DNA	25 cycles	30' on 30' off	~200–300 bp	
Human chromatin	7 cycles	30' on 30' off	~500 bp	(TK6)
Human chromatin	10 cycles	30' on 20' off	~500 bp	(HT1080)

4. The size-distribution of the sonicated DNA was assessed by gel electrophoresis on a 1.2% agarose gel using the FastRuler™ low range DNA ladder (Fermentas).
5. Alternatively, size-distribution can be checked using a TapeStation (Agilent Technologies) using D1000 screentapes.



**3.12*****Immunoprecipitation***

Both proteins (ChIP-chip) and DNA lesions (3D-DIP-Chip) can now be immunoprecipitated according to the following protocol. The most important distinction is the use of naked DNA for DIP and chromatin for ChIP. For DIP, we use the anti-cisplatin modified DNA, CP9/19, antibody (ab103261, Abcam), or Mouse Anti-CPD antibody (CAC-NM-DND-001, COSMO Bioscience). Use 0.5 mg of CP9/19 antibody 50  $\mu$ l of Dynabeads for the detection of cisplatin-induced DNA damage or 1.5  $\mu$ g of antibody for the detection of oxaliplatin-induced damage (*see Note 7*).

1. Take 50  $\mu$ l Dynabeads per sample and wash three times using 500  $\mu$ l PBS-BSA (1 mg/ml).
2. Resuspend the beads in 50  $\mu$ l PBS-BSA (1 mg/ml) and add the indicated amount of antibody.
3. Incubate the antibody with the washed Dynabeads for 30 min at 30 °C in a thermomixer (Eppendorf) shaking at 1300 rpm (optional: incubate overnight at 4 °C on a rotator).
4. Wash the Dynabeads again for three rounds using 500  $\mu$ l 1 mg/ml PBS-BSA and resuspend the beads in 50  $\mu$ l PBS after the final washing step.
5. Add 30  $\mu$ l PBS-BSA (10 mg/ml), 6–10  $\mu$ g sonicated DNA *or* 3  $\mu$ g sheared chromatin and add 1 $\times$  PBS to an end-volume of 300  $\mu$ l.
6. Incubate the IP/ChIP reaction at 21 °C for 3 h shaking at 1300 rpm using a thermomixer.
7. Wash the DynaBeads. All buffers are prechilled at 4 °C and kept on ice. Resuspend the beads in 500  $\mu$ l FA/SDS, incubate on a rotator for ~1 min, precipitate the beads on a magnet.
8. Remove the supernatant and repeat this process twice using FA/SDS 0.5 M NaCl (for yeast) or FA/SDS 1.0 M NaCl (for human).
9. Next, wash the beads using LiCl solution and 1 $\times$  TE.
10. Elute the DNA from the beads by incubating them in 125  $\mu$ l pronase buffer at 65 °C shaking at 900 rpm for 30 min in a thermomixer.
11. Add 6.25  $\mu$ l pronase (20 mg/ml) to the IP samples
12. Prepare an input sample containing ~1/10 of IPed DNA by diluting DNA into 100  $\mu$ l using TE. Add 25  $\mu$ l 5 $\times$  pronase buffer and 6.25  $\mu$ l pronase to the input DNA and incubate the samples together with the IP samples, overnight at 65 °C to remove all proteins and reverse the cross-linking.
13. For platinum-treated DNA:
  - (a) Treat the DNA with NaCN to chemically remove the platinum-induced adducts. Incubate both the IN and IP

DNA samples with 0.2 M NaCN for 2 h at 65 °C. Add 5 µl RNase (10 mg/ml) for the last 45 min of the incubation.

- (b) Purify the Input and IP DNA using PureLink™ Quick PCR Purification Kit (Invitrogen) following the supplier's instructions, eluting in 50 µl elution buffer.
- (c) The DNA is now ready for qRT-PCR validation and microarray sample preparation.

14. For UV treated DNA:

- (a) Add 5 µl of RNase A (10 mg/ml) and incubate at 37 °C for 1 h.
- (b) For yeast: Purify the Input and IP DNA using PureLink™ Quick PCR Purification Kit (Invitrogen) following the supplier's instructions, eluting in 50 µl elution buffer.
- (c) For human: extract the IP and Input DNA by phenol-chloroform extraction.
- (d) Treat 40 µl sample (IN and IP) with the PreCR DNA repair kit (New England Biolabs) according to the supplier's instructions.
- (e) Purify the Input and IP DNA using PureLink™ Quick PCR Purification Kit (Invitrogen) following the supplier's instructions, eluting in 50 µl elution buffer.
- (f) The DNA is now ready for qRT-PCR validation and microarray sample preparation.

### 3.13 Quantitative Real-time PCR

To assess the successful enrichment of the IP compared to the input DNA before continuing with the microarray, perform qRT-PCR using your favorite locus and RT-PCR system and reagents. Here, we provide details for the use of the iTaq™ Universal SYBR Green Supermix and CFX Connect™ Real-Time PCR detection system (Bio-Rad). We used the *MFA2* locus for yeast 3D-DIP-Chip experiment for the detection of DNA damages. Work on human cells made use of the 28S ribosomal RNA gene to verify successful DIP. The volumes of samples and dilutions listed here are enough for enrichment detection at 3 loci, with samples run in triplicate.

1. Dilute the input and IP samples 10× by adding 5 µl sample to 45 µl ddH<sub>2</sub>O.
2. Create an internal dilution series of input DNA ranging from 1/10 to 1/100,000 (5 µl in 45 µl ddH<sub>2</sub>O).
3. Create a master mix by adding forward and reverse primers to the SYBR Green Supermix to an end concentration of 1 µM diluting from 100 µM stock.
4. Aliquot 5 µl of the SYBR Green Supermix and primer master mix in a 96-well PCR plate (Bio-Rad). Account for samples

(input and IP) to be run in triplicate and include a set of three non-template controls (NTC). Add 5  $\mu\text{l}$  of sample to each well.

5. Vortex the 96-well PCR plate and spin down for 2 min at 300 rpm at room temperature to collect the sample at the bottom of the wells.
6. Load the 96-well PCR plate into the CFX Connect™ Real-Time PCR detection system (Bio-Rad).
7. Generate melt-curves to check for nonspecific amplicons.
8. Analyze the data using the CFX Manager™ Software v3.1 (Bio-Rad). Eliminate data points that deviated more than 0.5 units of the threshold (Ct).

### **3.14 Yeast DNA Labeling and Microarray Hybridization**

Yeast DIP and ChIP and accompanying input DNA samples can now be processed per the Agilent Technologies Yeast ChIP on chip protocol (Agilent Technologies Yeast ChIP-on-chip Analysis Protocol, version 9.2). A short summary of the steps is described below. The method for each of these steps has been described in detail previously [11] and is described extensively in the Agilent Yeast ChIP on chip protocol.

1. Blunt-end the DNA using T4 DNA polymerase.
2. Perform a phenol/chloroform extraction and precipitate the DNA using ice-cold 100% ethanol.
3. Ligate the linker hybrid to the input and IP DNA using T4 DNA ligase.
4. Amplify the DNA using two rounds of ligation-mediated PCR (LM-PCR).
5. Measure the DNA concentration in a NanoDrop spectrophotometer (ThermoScientific) and normalize the concentration of all samples to 150 ng/ $\mu\text{l}$  using ddH<sub>2</sub>O.
6. Differentially label 10.5  $\mu\text{l}$  the IP and IN DNA with the Cy5 and Cy3 fluorophores, respectively, using the BioPrime Total Genome Labelling System (Invitrogen) for 2 h at 37 °C (*see Note 8*).
7. Purify the DNA using the PureLink PCR Purification Kit (Invitrogen) and elute in 55  $\mu\text{l}$  elution buffer E1.
8. Apply 5  $\mu\text{l}$  of labeled sample on a NanoDrop spectrophotometer (ThermoScientific) using the MicroArray Measurement Software Module to establish the labeling efficiency.
9. Combine the labeled IP and IN samples (100  $\mu\text{l}$ ) and purify the DNA using ethanol precipitation (*see Agilent Yeast ChIP on chip protocol for details*).
10. Resuspend the DNA in 39  $\mu\text{l}$  ddH<sub>2</sub>O and add the hybridization mixture (end volume is 110  $\mu\text{l}$ ).

11. Apply 100  $\mu\text{l}$  of the hybridization mixture to each Agilent yeast  $4 \times 44$  k microarray (see Agilent Yeast ChIP on chip protocol for details).
12. Hybridize for 24 h at 65 °C at 20 rpm. For larger format arrays, e.g., 4x44K the hybridization time is 40 h.

### **3.15 Human DNA Labeling and Microarray Hybridization**

Successful amplification of human samples was performed using the GenomePlex Complete Whole Genome Amplification kit (WGA2, Sigma-Aldrich). This method was introduced for these samples because the LM-PCR-based method relies on two rounds of amplification, which is a potential source of amplification bias particularly when working with larger genomes [12]. This step of the 3D-DIP-Chip workflow is therefore slightly different for yeast and human-derived DNA samples and details are described below. The manufacturer's instructions were followed for amplification of human DNA using the WGA2 kit, with the following alterations:

1. No fragmentation was necessary as our DNA is fragmented by sonication already (*see* Subheading 3.11).
2. Use the entire DIP sample and add 2  $\mu\text{l}$  of  $1 \times$  library preparation buffer and 1  $\mu\text{l}$  of library stabilization buffer.
3. Incubate the mixture at 95 °C for 2 min and directly after, cool the sample on ice.
4. Add 1  $\mu\text{l}$  of library preparation enzyme and incubate in a thermal cycler as follows; 20 min at 16 °C, 20 min at 24 °C, 20 min at 37 °C, and 5 min at 75 °C.
5. Next, add  $10 \times$  amplification master mix, 47.5  $\mu\text{l}$  nuclease-free  $\text{H}_2\text{O}$ , and 5  $\mu\text{l}$  WGA DNA polymerase to each sample. Mix and incubate in a thermal cycler for amplification using: 3 min at 95 °C, 14–16 cycles of 15 s at 95 °C, and 5 min at 65 °C.
6. Purify the amplified DNA using the PureLink PCR Purification Kit (Invitrogen) and elute in 15  $\mu\text{l}$  nuclease-free  $\text{H}_2\text{O}$ .
7. Quantify the DNA using a NanoDrop spectrophotometer (ThermoScientific). A yield between 1 and 5  $\mu\text{g}$  is to be expected and is sufficient for microarray analysis.
8. Label the library DNA as described previously using the Bio-Prime Total Genomic Labelling System (Invitrogen, Subheading 3.14, steps 6–9) but now incubate the reaction for 3 h at 37 °C.
9. Combine the IP and IN samples and apply 100  $\mu\text{l}$  of the hybridization mixture to a custom-designed Agilent Technologies human  $1 \times 244$  k microarray.
10. Hybridize for 40 h at 65 °C at 20 rpm.

### 3.16 *Microarray Washing, Scanning, and Feature Extraction*

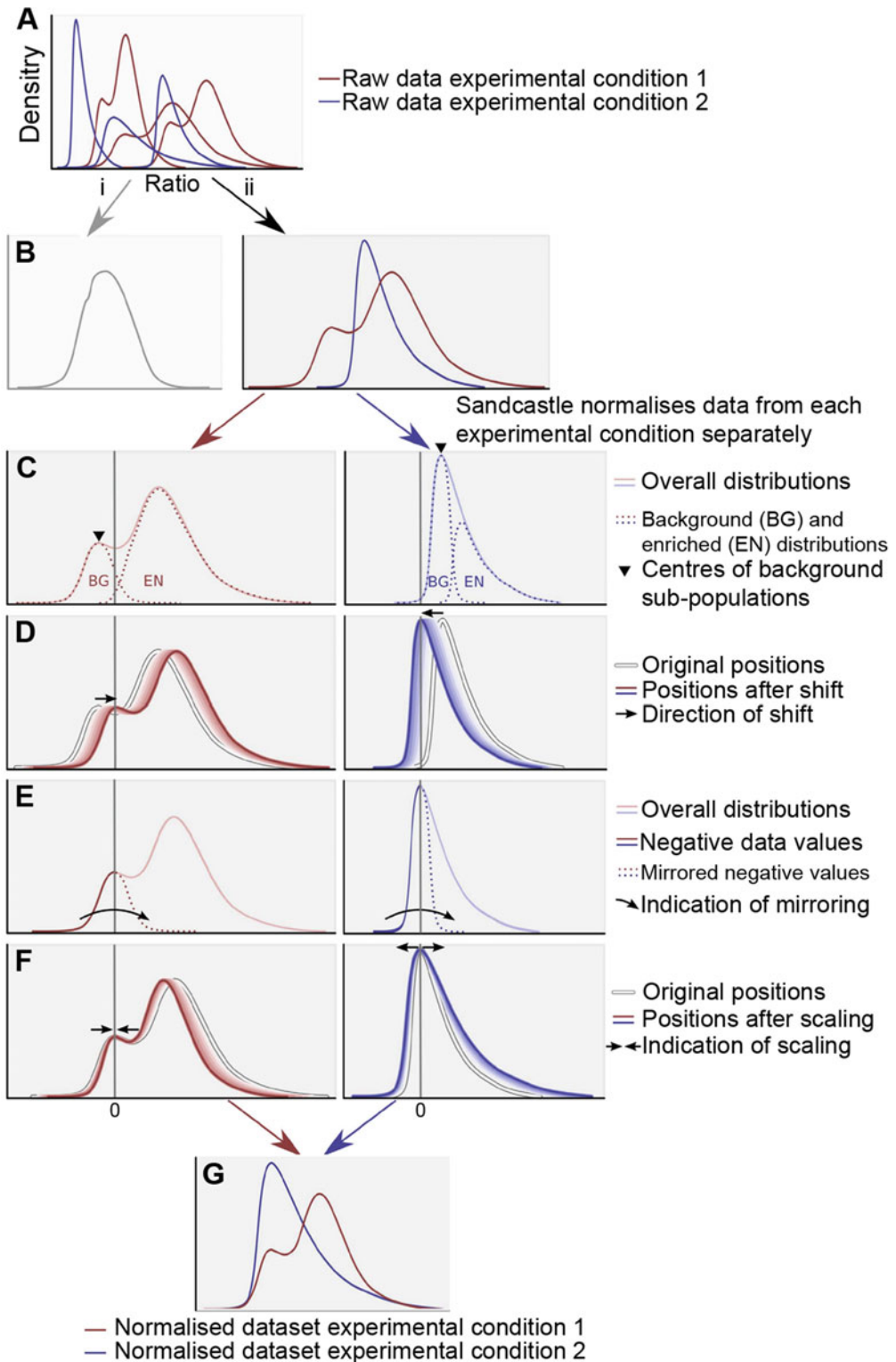
1. Both human and yeast arrays are washed twice using SSPE washing buffers as described in detail in the Agilent Technologies Yeast ChIP on chip protocol (Agilent Technologies Yeast ChIP-on-chip Analysis Protocol, version 9.2).
2. A washed and dried microarray slide is now ready for scanning.
3. The scanned image of the array records the fluorescence of both the Cy3 and Cy5 dyes in a TIFF image, the Agilent Feature Extraction software retrieves the raw data for analysis.
4. The raw data files can now be loaded into the R statistical programming language environment [9] by using the built-in functions of Sandcastle developed in our lab [3].

Details of data loading, normalization, visualization, and peak detection are discussed in the next section.

### 3.17 *Overview of Data Analysis Using Sandcastle*

The following section describes a typical data analysis workflow using Sandcastle. The Sandcastle software package is written in the statistical programming language of R [9] and provides a novel normalization procedure, enrichment detection, and graphical tools for plotting data. For full details on how to use each command, we refer to the Sandcastle instruction document (<http://reedlab.cardiff.ac.uk/sandcastle/sandcastleinstructions.pdf>) and the Sandcastle documentation and vignettes that can be accessed from within R.

In short, the normalization procedure performed by Sandcastle consists of four stages (Fig. 2) and works only when the background and enrichment signal can be clearly differentiated (*see* Fig. 2). The first step of preprocessing (1) removes unused or unusable data points, ensuring that all datasets contain the same amount of useable data points before moving forward. It should be noted that this step only removes a small set of the data points. (2) During the second stage of the normalization process, *within*-condition quantile normalization is applied to reduce variation among replicate datasets derived from the same experimental condition. The third and fourth stages of normalization introduce the unique *between*-condition normalization by (3) shifting and (4) scaling the data and can only be applied to data of which the background population can be identified. Shifting the datasets is achieved by centering the background subpopulation to the  $y$ -axis ( $x = 0$ ) for each individual dataset. This generates the same mean of zero for each of the background subpopulation of data points for all datasets. Next, to be able to scale the data, the left-hand, negative portion of the background data is mirrored along the  $y$ -axis to create a symmetrical dataset that acts as a simulated background distribution (Fig. 2). The background population can now be scaled to a standard normal distribution with a mean of zero and



**Fig. 2** Representation of the normalization procedure. (a) Raw density profiles of datasets from two experimental conditions (*red and blue*), each with three replicates. Differences in the shapes of the profiles indicate

SD of 1. This transforms each dataset to have the same theoretically consistent background subpopulation against which it is now possible to assess the enriched, biological data in relative terms.

### 3.18 Loading Data and Compatible Data Formats

Sandcastle can use raw ChIP-chip data derived from Agilent's Feature Extraction software or tab-delimited files from other sources. When the data are loaded by Sandcastle the `arrayData` object that is created contains the  $\log_2$  IP/IN ratios for each probe or feature of the microarray, including information on the genomic coordinates and probe annotation.

1. The `findArrayFiles` function will identify files suitable for loading into R in your R working directory, giving each file a numbered identifier that can be used when loading the data.
2. First, a quality control can be performed on the raw data. Using `checkData`, files found by the `findArrayFiles` function will be put through QC analysis. A "check.pdf" file will be generated in your working directory that can be printed out for logging the experiment. This file provides fields to write down the experimental conditions, as well as pseudo images of the red and green channel derived from the scanned array for visual inspection. The distribution of signal in the dataset is presented as a box-plot, density profile and scatter plot.
3. Using the `loadArrayFiles` function, data will be loaded into an `arrayData` object. Making use of the numbers assigned to each

**Fig. 2** (continued) experimentally induced biologically relevant changes, but these cannot be compared in their raw state. **(b)** Quantile normalizing all datasets together (1) removes much of the experimentally induced, biologically relevant differences between them. This is not desirable, as these differences cannot then be investigated. Sandcastle quantile normalizes the datasets from each experimental condition separately, to maintain these biological differences. Quantile normalization makes each of the datasets follow the same distribution, meaning all density profiles from each experimental condition overlap each other (2). This reduces intra-condition—but not inter-condition—technical variations. **(c)** Each dataset consists of two overlapping subpopulations (*dashed lines*), background (BG), and enriched (EN). These cannot be fully discerned in the data and only the overall population (*solid lines*) is known. Sandcastle performs inter-condition normalization based on estimated background subpopulations. This requires the central (modal) point of the background subpopulations to be identifiable (*marked with triangles*). If this central point cannot be discerned (for example, if the background subpopulation is too small), then the Sandcastle normalization cannot be applied. **(d)** Data are first shifted to center the modal point of the estimated background subpopulation on zero (indicated by *arrows*). **(e)** To estimate the properties of the whole background subpopulation all negative values (the *left-hand side* of the estimated background subpopulation following the shift step) are mirrored into the positive (indicated by *arrow*; *dashed lines* show mirrored data). This allows the standard deviation of the estimated background subpopulation to be calculated. **(f)** Data are scaled to make the calculated standard deviation of the estimated background subpopulation 1 (indicated by *arrows*). **(g)** The resulting fully normalized datasets have estimated background subpopulations with the same mean (0) and standard deviation (1). Comparisons of data between conditions can now be made relative to this common background. For clarity axis labels are only shown in **(a)**—all other *x*- and *y*-axes are ratio and density values respectively. *Vertical gray lines* indicate 0, which are only labeled in **(f)**

dataset by *findArrayFiles* the order and grouping of the datasets to be loaded can be specified here.

4. Next, the user needs to load the annotation data from the yeast ensemble database to an annotation data object using *loadAnnotation*. The annotation data contains information about all annotated genes used for genomic analysis and graphical representation of genomic data. The default data loaded is that of yeast but human annotation data can also be loaded (*see* the Sandcastle documentation for more details).

### 3.19 Normalization

1. Data loaded as an *arrayData* object can now be normalized by calling the *normalize* function.
2. The “batches” argument can be used with the normalization function to combine the replicate datasets that need to be quantile normalized to eliminate the “within” condition variation. After this step the group(s) of quantile normalized data will be subjected to the scaling and shifting algorithms to allow the “between” condition comparison.
3. A graphical representation of the individual steps of the normalization procedure can be generated using the “plot” argument, allowing the user to follow exactly each step of the normalization process.
4. Knowledge of the dataset and nature of the signal detected in the 3D-DIPChip or ChIP-chip experiment are important in determining whether the dataset is amenable to Sandcastle’s full normalization procedure (Fig. 2). By default, all the four steps of the normalization procedure, described earlier (Sub-heading 3.17), are performed. However, for 3D-DIP-Chip data that does not contain a discernible background signal, the scaling and shifting steps must be omitted by setting the “shift” and “scale” arguments to FALSE.
5. After applying the appropriate normalization to the data, the *normalisationAssumptions* function provides a means of validating whether two datasets that were normalized in different batches (e.g., treated and untreated) are suited to the normalization procedure (Fig. 2).

### 3.20 Enrichment Detection

After following the previous steps enrichment detection and peak detection can be performed by calling the *enrichmentDetection* function. Enrichment detection describes the probe enrichment, whereas peak detection finds local maxima and generates a *sandcastlePeaks* list object, identifying potential binding regions of the protein of interest in the case of ChIP-chip data analysis (*see* Note 9). 3D-DIP-Chip data describing UV-induced CPDs and their removal is not amenable



of peak detection since CPD distribution is random and dispersed and is not expected to generate defined local peaks at significant resolution.

### 3.21 Plotting

Sandcastle provides means to plot data in a linear, genomic fashion as well as aggregating data into composite plots around genomic features such as ORFs, TSSs, TESs but also peaks or other genomic coordinates supplied by the user.

1. *genomePlot* plots the data for the first chromosome, but can also be instructed to display user-defined sections of a chromosome or export a whole-genome plot to a PDF file. This function uses the annotation data loaded through *loadAnnotation* (see Subheading 3.18, step 4) to plot the location and direction of genes. Using standard R-commands, data can be group together and averaged to plot results in different context.
2. *profilePlot* allows users to generate so-called *composite plots* that aggregate genomic data from a user-defined selection of genomic positions or structures and compiles them into a single plot. These genomic features can, for instance, be protein binding sites, peaks detected using Sandcastle's *enrichmentDetection* (Subheading 3.20) or ORF-features extracted from the ensemble annotation data.

---

## 4 Notes

1. Following yeast cell growth using OD<sub>600</sub> can vary based on the yeast strain grown and spectrophotometer used. Cell-counting is more reliable but also more time consuming. Yeast cell doubling time can vary from 1.4 to close to 2.0 h depending on the genetic background or mutant used.
2. Snap-freezing and storing yeast cell pellets at  $-80\text{ }^{\circ}\text{C}$  after treatment improves the efficiency of cell lysis.
3. When performing cross-linking in medium (i.e., for the HT1080 cells), quenching the formaldehyde cross-linking reaction using glycine changes the color of the medium due to the pH of the glycine stock ( $\sim\text{pH } 2.5$ ) if a pH indicator is present.
4. In yeast only use PMSF during lysis because of high protease activity in yeast extracts. Use of protease inhibitor cocktails is possible but not required. Use of protease inhibitor cocktails (i.e., from Roche) is recommended for human cell lysis and chromatin preparation.

5. Monitoring spheroplast production:
  - (a) Add 100  $\mu$ l 10% SDS to 100  $\mu$ l cell suspension.
  - (b) Deposit 5–10  $\mu$ l of this mixture together with 5–10  $\mu$ l of cell suspension on an object slide.
  - (c) The absence of the cell wall should be clearly discernible comparing the cell suspension with or without SDS. In the presence of SDS the cells are very faintly visible compared to the untreated suspension.
  - (d) Spheroplast production is complete if most cells are poorly visible in the presence of SDS.
6. Processing sonication human DNA samples for microarray analysis revealed that a shorter fragment length was required to improve microarray resolution. For yeast studies, an average fragment length of 300–500 bp was used, while for human studies a shorter fragment length of 200–300 bp was required.
7. To increase the overall DNA yield in the human DIP samples, DIP samples can be processed in parallel and two samples pooled together prior to the qRT-PCR stage. Multiple labeled IP DNA samples can also be combined prior to microarray hybridization as a means of boosting the microarray signal.
8. You can store the IP and IN DNA after labeling for any length of time prior to purification if 5  $\mu$ l of 0.5 M EDTA is added to the reaction mixture to quench the reaction.
9. Enrichment detection power improves by providing replicate datasets, but detection of enrichment or peaks in individual datasets can be performed as well.

## References

1. Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362(6422):709–715
2. Powell JR et al (2015) 3D-DIP-Chip: a microarray-based method to measure genomic DNA damage. *Sci Rep* 5:7975
3. Bennett M et al (2015) Sandcastle: software for revealing latent information in multiple experimental ChIP-chip datasets via a novel normalisation procedure. *Sci Rep* 5:13395
4. Yu S et al (2016) Global genome nucleotide excision repair is organized into domains that promote efficient DNA repair in chromatin. *Genome Res* 26(10):1376–1387
5. Nik-Zainal S et al (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534(7605):47–54
6. Morganella S et al (2016) The topography of mutational processes in breast cancer genomes. *Nat Commun* 7:11383
7. Alexandrov LB et al (2013) Signatures of mutational processes in human cancer. *Nature* 500(7463):415–421
8. Perera D et al (2016) Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature* 532(7598):259–263
9. R Core Team R (2015) A language and environment for statistical computing
10. Brachmann CB et al (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14(2):115–132
11. Teng Y et al (2011) A novel method for the genome-wide high resolution analysis of DNA damage. *Nucleic Acids Res* 39(2):e10
12. Powell JR (2014) Thesis, Cardiff University ORCA portal. <http://orca.cf.ac.uk/id/eprint/65976>

## Study of UV-induced DNA Repair Factor Recruitment: Kinetics and Dynamics

Sarah Sertic, Stefania Roma, Paolo Plevani, Federico Lazzaro, and Marco Muzi-Falconi

### Abstract

The local UV irradiation technique enables detection, kinetic measurements of recruitment, and quantification of DNA Damage Response (DDR) proteins at the site of UV-induced DNA damage.

Using Isopore filters with high density pores of a broad range of sizes, it is possible to UV irradiate and damage only a very small portion of the nucleus of a cell by letting UV light pass only through the pores. Immunofluorescent analyses of modified DNA nucleotides, proteins, or fluorescently tagged versions of target factors can be used as markers to label and study UV-induced lesions and their repair.

**Key words** UV-induced DNA damage, Immunofluorescence, Protein recruitment, Sequential assembly, NER

---

### 1 Introduction

Nucleotide Excision Repair (NER) is involved in the removal of bulky lesions on DNA, and in human cells is the only process that repairs UV-induced DNA injuries. Several genetic and biochemical studies during the last few decades have defined details of paramount importance for the mechanistic characterization of NER [1]. It is a multistep process involving many enzymes starting from recognition of DNA damage, unwinding of double helix, endonucleolytic cleavage 5' and 3' to the lesion, gap-refilling and ligation.

However, characterization of the sequential assembly, dynamic and kinetics of complexes involved in the different steps of DNA lesion removal and repair has been possible thanks to the (r)evolution of sophisticated fluorescence microscopy techniques. In particular, different laboratories [2–4] developed a very powerful technique, based on the induction of Local UV Damage (LUD). The added value of this technique resides in its simplicity.

UV-light produces bulky lesions on DNA, which are very dangerous if not properly repaired as they can be led to double-strand breaks and block both DNA replication and transcription. Mutations in human genes coding for NER factors are responsible for the onset of severe pathologies, among which are xeroderma pigmentosum (XP), Cockayne syndrome (CS), and Trichothiodystrophy (TTD). The principal and common symptom to all these conditions is extreme sensitivity to UV irradiation, while other manifestations can vary. XP mutations have been grouped into eight complementation groups (from XP-A to XP-G and XP-V). The LUD technique has been very important to define who is first recruited to the UV lesion and the sequential assembly of all the factors [4]. For example, it has been used to demonstrate that the incisions at the 5' side of the lesions by XPF and 3' side of the lesion by XPG are sequential and they occur in this order [5]. Using XP cells from patients, recruitment at LUDs has been used to identify the genetic dependencies for the recruitment at UV lesions of different new factors, such as XRCC1-Lig3, MutS $\beta$ , MDC1-53BP1, Pol  $\kappa$ , and EXO1 [6–10].

Here, we describe the technical procedure to generate LUDs that can be used to define details for the recruitment and involvement of proteins in UV-induced DNA damage processing and repair.

---

## 2 Materials

### 2.1 Cell Cultures

1. Cell dishes, multiwell plates.
2. PBS sterile for cell culture.
3. Specific cell culture medium complemented with Pen/Strep, Glutamine, Fetal Bovine Serum (at different concentrations).
4. Glass coverslips 20 × 20 mm.

### 2.2 Local UV Irradiation

1. Thin tweezers.
2. Isopore filters 5  $\mu$ m (Merck-Millipore TMTP04700).
3. UV dosimeter.
4. Timer.

### 2.3 Immunostaining

1. PBS.
2. 4% Paraformaldehyde (2 $\times$ ).
3. Formaldehyde 37% (10 $\times$ ).
4. Permeabilization Buffer: 0.5% Triton-X100 in PBS (4 °C).
5. Blocking Buffer: 10% BSA in PBS (4 °C).
6. Antibodies incubation buffer: 0.1% Tween-20 in PBS.
7. ProlongGold with DAPI.

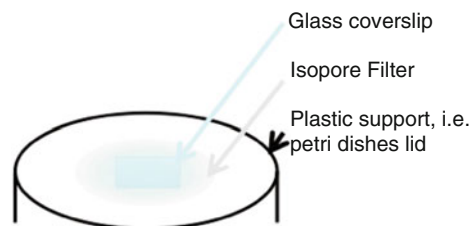
## 2.4 Detection and Analysis

1. Widefield Fluorescence Microscope.
2. High-resolution Spectral Confocal Microscope.

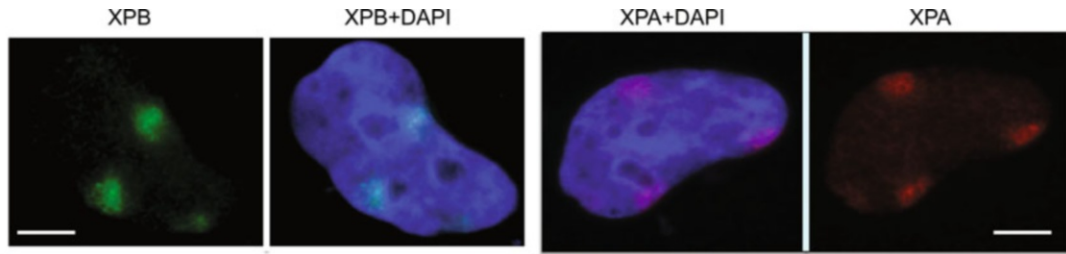
## 3 Methods

### 3.1 Local UV Irradiation

1. Day 1: cell seeding (*see Note 1*) on coverslip, one for each treatment, UV dose or/and different kinetics.
  2. Day 2: transfect cells or treat them with any chemical that needs 24 h incubation.
  3. Day 3: Local UV irradiation and immunofluorescence.
  4. Day 4: Fluorescence microscopy analyses.
1. Place UV-Lamp 254 nm on a proper support to ensure appropriate distance between lamp and sample, so that cells receive a UV dose of 0.5–1 J/m<sup>2</sup>/s (*see Note 2*). Switch on lamp.
  2. Control UV dose with a dosimeter after a lamp warm-up time of at least 15–20 min.
  3. Prepare petri dish lid covered with parafilm.
  4. Wash cells with PBS.
  5. Lift the coverslip with tweezers and position it on the parafilm covered lid, with cells on top.
  6. Wash the Isopore filter with PBS in a petri dish.
  7. Place the filter on the top of the coverslip (Fig. 1).
  8. Set an alarm for the seconds of irradiation needed to reach the desired UV dose (e.g., 40 J/m<sup>2</sup>, corresponding approximately to a total irradiation of 2 J/m<sup>2</sup>, when using 5 μm pores).
  9. Place the petri dish lid holding the coverslip and filter under the warmed UV lamp.
  10. After the appropriate time, when desired UV dose is reached, remove it and lift Isopore filter with tweezers with one hand, while adding PBS with a hand pipette using the other hand (*see Note 3*).
  11. Lift coverslip and put it back in a multiwell plate.



**Fig. 1** Local UV damage casting setup



**Fig. 2** Example of two different NER factors at LUDs by immunofluorescence. MRC5VI cells were seeded and irradiated with  $40 \text{ J/m}^2$  through Isopore filter of  $5 \mu\text{m}$  dimension and incubated for 30 min. Scale bar =  $5 \mu\text{m}$

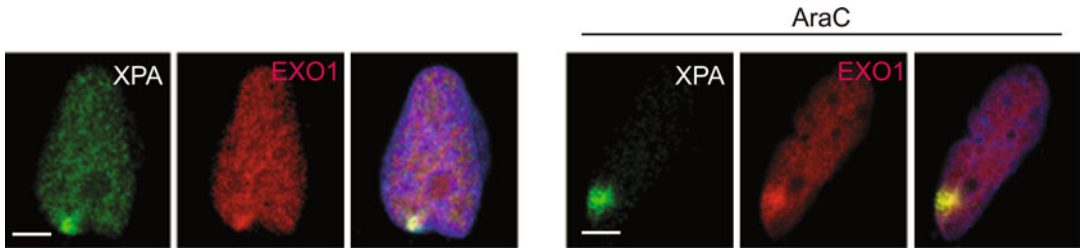
12. Incubate for the desired period of time (e.g., 1 h) in a  $37^\circ\text{C}$  humidified incubator.
13. Wash cells twice with PBS.
14. Fix cells with 2% paraformaldehyde for 20 min at room temperature or 3.7% formaldehyde for 15 min.
15. Wash cells twice with PBS (*see Note 4*).
16. Add ice-cold permeabilization buffer for 5 min.
17. Add blocking buffer for 30–60 min at room temperature.
18. Incubate with primary antibody (diluted in PBST) at the appropriate dilution for 1.5–2 h at room temperature in a humidified chamber.
19. Wash three times with PBST.
20. Incubate with fluorescent secondary antibody (AlexaFluor<sup>®</sup>, Molecular Probes) for 1 h at room temperature (diluted in PBST) protected from light.
21. Wash three times with PBST.
22. Mount coverslip with ProLong Gold mounting solution with DAPI.
23. Let coverslip dry overnight at room temperature (*see Note 5*).

### 3.2 Microscopic Analysis

Samples can be analyzed under a widefield microscope or confocal microscope depending on the need. A  $63\times$  or  $100\times$  immersion oil objective can be used for the analysis. Example images are presented in Figs. 2 and 3.

## 4 Notes

1. Be careful of seeding cells at a maximum confluence of 70% the day of Local UV damage protocol. Higher confluence would give rise to a bad LUD accumulation and immunofluorescence, while lower confluence would favor cell loss when removing Isopore filters, leaving few cells for the analyses.



**Fig. 3** Application of LUD technique. MRC5VI cells were seeded and transfected with EXO1-mCherry tagged protein. The following day, cells were Local UV irradiated at  $100 \text{ J/m}^2$  in the presence or absence of AraC. Representative images, acquired under a confocal microscope (Leica SPE) for each condition, are shown. Scale bar =  $5 \mu\text{m}$

2. An appropriate UV dose ( $\text{J/m}^2/\text{s}$ ) is suggested. If the dose is too high reproducibility can be affected, while too low doses will need long incubation times of living cells under the lamp in an almost dry condition. We generally use a 6 W 254 nm lamp (VL-6.C) placed 35 cm away from the cells.
3. Be careful when lifting the filter, pipetting PBS simultaneously is essential to avoid scraping cells off the coverslip.
4. If needed, fixed cells can be stored at  $4^\circ\text{C}$  for up to 1 week.
5. Do not place coverslip under an immersion oil objective before drying is completed, since wrong movement of the sample might result in scratching of cells and morphological changes.

## References

1. Gillet LCJ, Schärer OD (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* 106:253–276
2. Katsumi S, Kobayashi N, Imoto K et al (2001) In situ visualization of ultraviolet-light-induced DNA damage repair in locally irradiated human fibroblasts. *J Invest Dermatol* 117:1156–1161
3. Moné MJ, Volker M, Nikaido O et al (2001) Local UV-induced DNA damage in cell nuclei results in local transcription inhibition. *EMBO Rep* 2:1013–1017
4. Volker M, Moné MJ, Karmakar P et al (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* 8:213–224
5. Staresinic L, Fagbemi AF, Enzlin JH et al (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J* 28:1111–1120
6. Moser J, Kool H, Giakzidis I et al (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell* 27:311–323
7. Hong Z, Jiang J, Hashiguchi K et al (2008) Recruitment of mismatch repair proteins to the site of DNA damage in human cells. *J Cell Sci* 121:3146–3154
8. Ogi T, Limsirichaikul S, Overmeer RM et al (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol Cell* 37:714–727
9. Sertic S, Pizzi S, Cloney R et al (2011) Human exonuclease 1 connects nucleotide excision repair (NER) processing with checkpoint activation in response to UV irradiation. *Proc Natl Acad Sci U S A* 108:13647–13652
10. Marteiijn JA, Bekker-Jensen S, Mailand N et al (2009) Nucleotide excision repair-induced H2A ubiquitination is dependent on MDC1 and RNF8 and reveals a universal DNA damage response. *J Cell Biol* 186:835–847

## Inserting Site-Specific DNA Lesions into Whole Genomes

Vincent Pagès and Robert P. Fuchs

### Abstract

Here, we describe a methodology that allows the insertion of site-specific DNA lesions into genomes in living cells. The technique involves the integration of a plasmid containing a site-specific lesion engineered *in vitro* into a precise location in the genome via the site-specific recombination reaction from phage lambda. The notion of DNA lesion is not restricted to chemically modified nucleotides but also refers to unusual DNA structures. This method will be instrumental to study qualitatively and quantitatively the genetic consequences of site-specific lesions *in vivo*; moreover, it does also allow analyzing the molecular structure of stalled replication forks at well-defined locations.

**Key words** DNA adducts, Unusual DNA structures, DNA damage tolerance, Site-specific recombination, Translesion synthesis, Damage avoidance, Homology-directed recombination

---

### 1 Introduction

In cells, replication-blocking lesions are tolerated via two major pathways, mutagenic Translesion synthesis (TLS) and error-free Damage Avoidance (DA). Currently, most molecular and genetic studies of TLS and mutagenesis are achieved by using plasmids probes carrying site-specific single adducts. However, it became clear that plasmids are inappropriate tools for the study of DA mechanisms [1, 2]. Indeed, replication fork uncoupling [3, 4] that occurs locally at the site of a transiently blocked replication fork leads to complete unwinding of the parental strands in plasmids, thus preventing DA pathways to be functional. To overcome this difficulty we developed a method to insert a construct containing a single lesion into a whole genome by harnessing the site-specific recombination machinery of phage lambda. The methodology, described in the present chapter for the insertion of a single lesion into the *E. coli* genome, can be transposed to yeast or mammalian cells. The methodology can be applied to the study of chemical lesions (UV induced, lesions, adducts formed by chemical carcinogens) or to sequences potentially forming unusual DNA



structures (G4 quadruplex, H-DNA, cruciforms, etc.) or to short ribonucleotides stretches embedded in DNA. This approach not only allows the genetic consequences induced by a site-specific lesion in a genome to be analyzed, it also allows the structure and the dynamics of a single well-defined stalled replication fork to be dissected at the molecular level. The genetic and molecular consequences as a function of the site of stalling (close to replication origins, to replication termini, etc.) can be studied.

---

## 2 Material

1. *E. coli* strain EC100D pir-116 (from Epicentre Biotechnologies—cat# EC6P0950H).
2. LB liquid media.
3. LB agar plates.
4. SOC (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose).
5. IPTG (Isopropyl β-D-1-thiogalactopyranoside). Stock at 200 mM in H<sub>2</sub>O.
6. X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). Stock at 80 mg/ml in N,N-Diméthylformamide.
7. Ampicillin. Stock at 100 mg/ml in H<sub>2</sub>O.
8. Kanamycin. Stock at 50 mg/ml in H<sub>2</sub>O.
9. Chloramphenicol. Stock at 30 mg/ml in ethanol.
10. EtOH/Na-acetate solution: dissolve 21.8 g of Sodium Acetate (MW = 136) into 20 ml of water. Complete to 1 l with 100% Ethanol.
11. 5× GM buffer:
12. 0,5 ml 1 M TrisHCl pH 7 (50 mM final).
13. 0,1 ml 5 M NaCl (50 mM final).
14. 0,1 ml 0,5 M EDTA (5 mM final).
15. H<sup>2</sup>O complete to 10 ml.
16. Filter sterilize and store at 4 °C.
17. 10× Gap Duplex Ligation Buffer: 500 μl 1 M Tris–HCl pH 7,5 (100 mM final), 500 μl 1 M MgCl<sub>2</sub> (100 mM final), 500 μl 1 M DTT (100 mM final), 3500 μl H<sub>2</sub>O.
18. Isopropyl-alcohol/TE/CsCl saturated: Saturate TE solution with CsCl. Then add Isopropyl-alcohol.
19. MaxiPrep kit (Qiagen).
20. 30 ml glass tubes, type Corex 30.
21. Ultracentrifugation tube 8 × 35 mm (Beckman # 361082).

22. Electroporation apparatus: BioRad Gene Pulser.
23. Ultracentrifuge type Beckman Optima™ L-90K and rotor NVT 65.2.
24. Ultracentrifuge type Beckman Optima™ Max-XP with rotor TLN120.
25. Automatic serial diluter and plater EasySpiral Dilute (Interscience).
26. Scan 1200 automatic colony counter (Interscience).

---

## 3 Methods

### 3.1 Construction of the Recipient Cell

The recipient cell must contain a unique attR integration site in fusion with the 3' end of *lacZ* gene. We use MG1655 strain in which the original  $\lambda$  attB site was replaced by an artificial promoterless operon carrying attR fused to the 3' end of *lacZ* upstream of the *aadA* gene (conferring Spectinomycin resistance) and between *ybhC* and *ybhB* [5]. This locates the integration site around minute 17 of the chromosome (Fig. 1a). Two parental strains were obtained, FBG151 where the integration of the lesion occurs on the lagging strand, and FBG152 where *lacZ* is inverted to locate the integrated lesion on the leading strand (Fig. 1c).

pVP135 plasmid expresses the *xis/int* operon from phage lambda under the control of a *trc* promoter that has been weakened by the mutation in the -35 and the -10 regions [6] to obtain optimal expression level of integrase. This plasmid is transformed into this recipient strain and maintained by growing the cells with 50  $\mu\text{g}/\text{ml}$  of kanamycin.

### 3.2 Control Plasmid to Measure Transformation Efficiency

To normalize integration efficiency between the damaged vector and the non-damaged control, we use an internal standard (pVP146) in the electroporation experiment. pVP146 is derived from pACYC184 plasmid where the chloramphenicol resistance gene has been deleted by *Bsa*AI digestion and re-ligation. This vector, which carries only the tetracycline resistance gene, serves as an internal control for transformation efficiency.

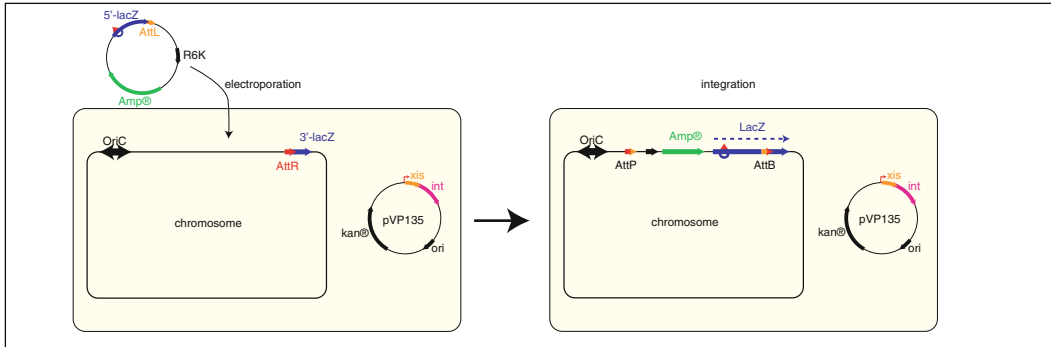
### 3.3 Construction of the Damaged Integrating Vector by the "Gapped Duplex" Method [7]

Two plasmids are required for the construction of the damaged integrating vector by the gapped duplex method. One plasmid named "parental" plasmid contains the sequence where the lesion-containing oligonucleotide will hybridize. The other plasmid named "helper" is lacking the sequence of the oligonucleotide allowing generating a precise gap into which the oligonucleotide will be incorporated.

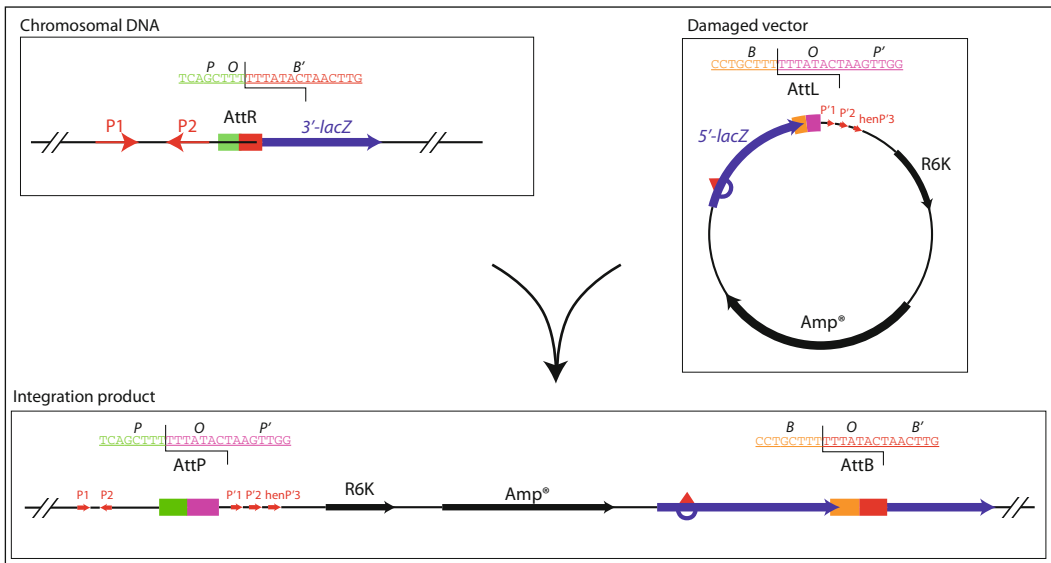
#### 3.3.1 Description of the Parental Plasmids

In the following example, we will use the combination of plasmid pVP141 (helper) and pVP142 (parental) that allows us to

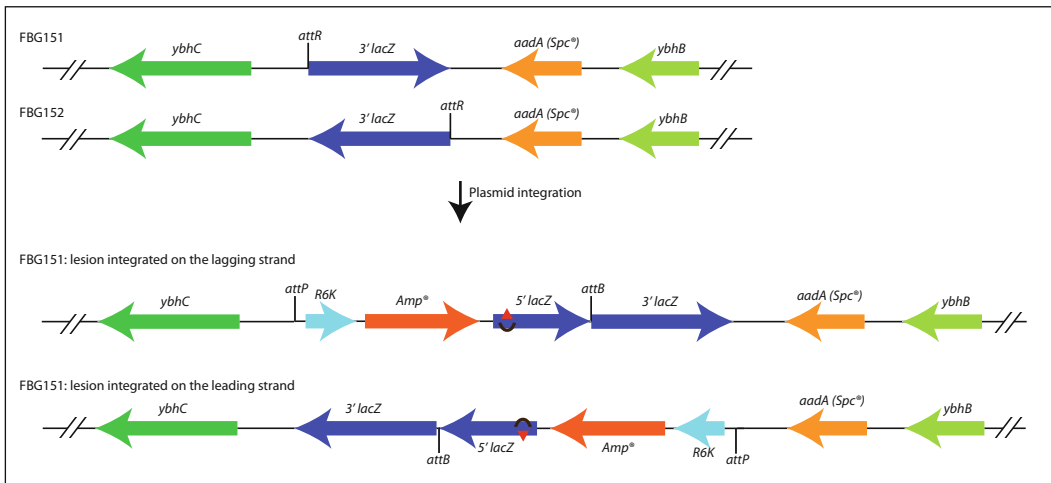
**A. Outline of the integration system**



**B. Detail of the phage lambda integration system**



**C. Map of the strains**



**Fig. 1 (a)** Outline of the integration system: The recipient strain contains a single *attR* integration site in fusion with the 3' end of *lacZ* gene at min 17 in the *E. coli* chromosome. Following ectopic expression of phage

probe non-frameshift TLS events at the *NarI* locus modified by the G-AAF lesion [1, 2].

Plasmids pVP141 and pVP142 plasmids contain the following characteristics:

- the *bla* gene that confers resistance to ampicillin thus allowing for the selection of the integration events.
- the R6K replication origin that prevents autonomous replication of the plasmid in the recipient strain. This plasmid can be propagated in a strain that carries the *pir* gene [8].
- the 5' end of *lacZ* gene in fusion with the attL site of phage lambda. The P3' site of attL has been mutated (AATGATTAT to AATTATTAT) to avoid excision of the plasmid once integrated [9].

### 3.3.2 Gap Duplex Assembly of the Damaged Vector

Production of the vectors:

Plasmids are produced in strain EC100D *pir*-116 (*see Note 1*).

Plasmids are purified using the MaxiPrep kit.

To obtain plasmid of higher purity, we further purify the DNA by cesium chloride/Ethidium bromide equilibrium centrifugation on Beckman Beckman Optima™ L-90K ultracentrifuge with rotor NVT 65.2 [10].

Digestion of the vectors (Fig. 2a):

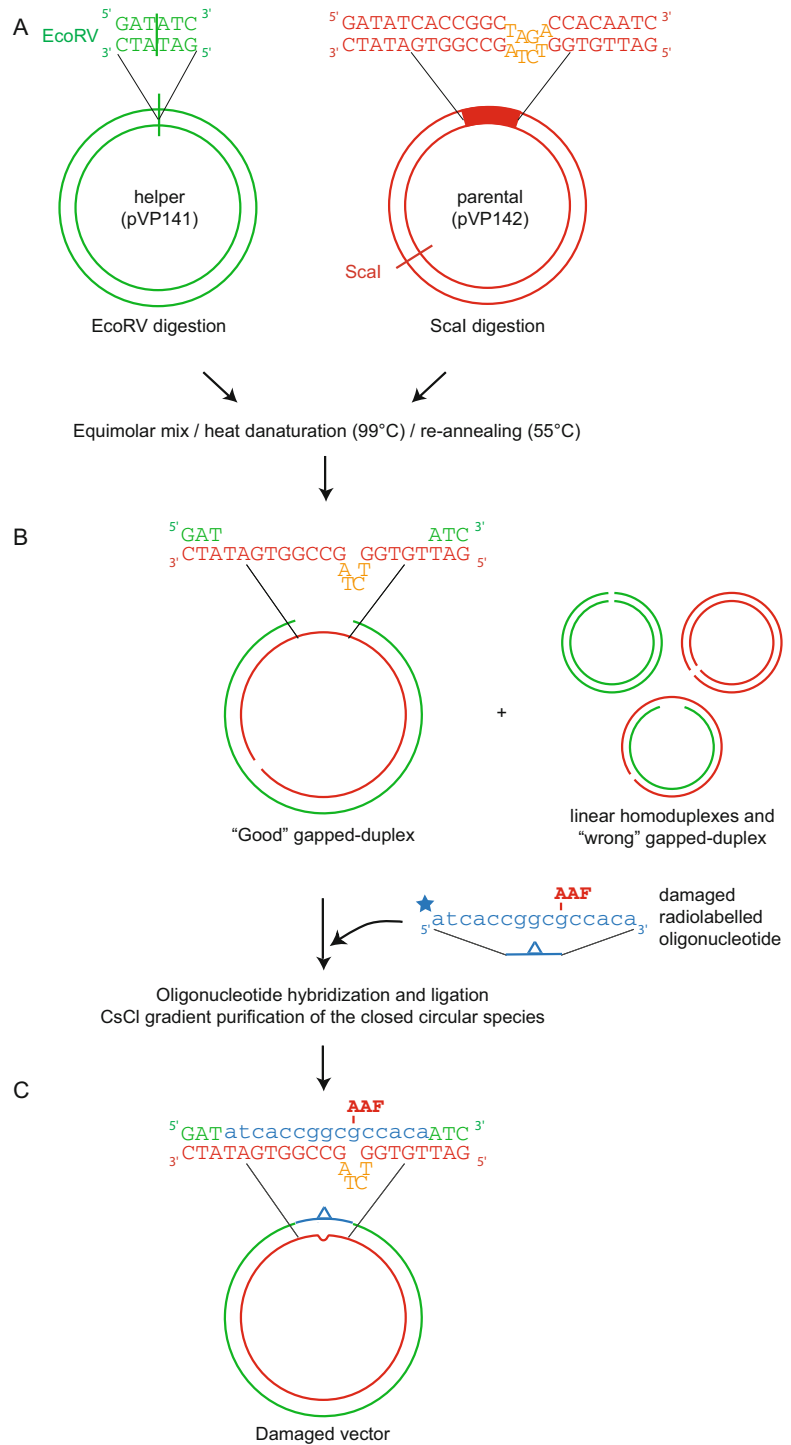
Helper plasmid (pVP141) is digested with EcoRV, whereas parental plasmid (pVP142) is digested with ScaI.

pVP141	pVP142
50 µl pVP141 at 1 µg/µl (=50 µg)	50 µl pVP142 at 1 µg/µl (50 µg)
40 µl NEB buffer cutsmart	40 µl NEB buffer cutsmart
4 µl EcoRV-HF (20 U/µl)	4 µl ScaI-HF (20 U/µl)
306 µl H <sub>2</sub> O (final 400 µl)	306 µl H <sub>2</sub> O (final 400 µl)

1. Incubate for 3 h at 37 °C.
2. Take a sample to check that the digestion is complete on agarose gel electrophoresis.

←

**Fig. 1** (continued) lambda int–xis, the lesion-carrying construct is introduced by electroporation. Its attL site will recombine with the chromosomal attR, leading to the integration of the entire lesion-containing construct. Integration events are selected on the basis of their resistance to ampicillin. Integration at nucleotide level resolution restores a functional *lacZ* gene allowing these events to be monitored on X-gal indicator plates. **(b)** Detail of the phage lambda integration system: The figure shows the detail of the core sequences of the site-specific recombination sites of phage lambda: attR, attL, attB, and attP. **(c)** Map of the strains before and after integration. In strain FBG151, the lesion is integrated in the lagging strand. In strain FBG151, attR-*lacZ* has been inverted to integrate the lesion in the leading strand



**Fig. 2** Gapped duplex method to produce vectors with a single lesion. **(a)** Helper and parental plasmid are linearized by a restriction endonuclease generating blunt ends. **(b)** Equal amounts of the two plasmids are mixed, heat-denatured and allowed to re-anneal to form to gapped-duplex molecules. Two molecular

3. Ethanol precipitate by adding 1.2 ml of EtOH/Na-acetate solution.
4. Spin and wash the pellet with 1 ml 70% Ethanol.
5. Spin and dry the pellet under vacuum.
6. Resuspend in 50  $\mu$ l of water.

Gapped-duplex formation (Fig. 2b):

All the following steps are executed in duplicate, one being used for the non-damaged control, the other for the lesion-containing plasmid.

1. Mix in a 30 ml glass tube (*see Note 2*):

pVP141 digested EcoRV (1 $\mu$ g/ $\mu$ l)	25 $\mu$ l
pVP142 digested ScaI (1 $\mu$ g/ $\mu$ l)	25 $\mu$ l
GM buffer 5 $\times$	1 ml
H <sub>2</sub> O	3.9 ml

→ freeze a 20  $\mu$ l aliquot for electrophoresis checking (STEP 0).

2. Incubate for 3.5 min at 99 °C with constant and vigorous shaking.
3. Immediately after, cool off in a mix of water/crushed ice while shaking for 2 min.
4. Leave three additional minutes without shaking.

→ freeze a 20  $\mu$ l aliquot for electrophoresis checking (STEP 1).

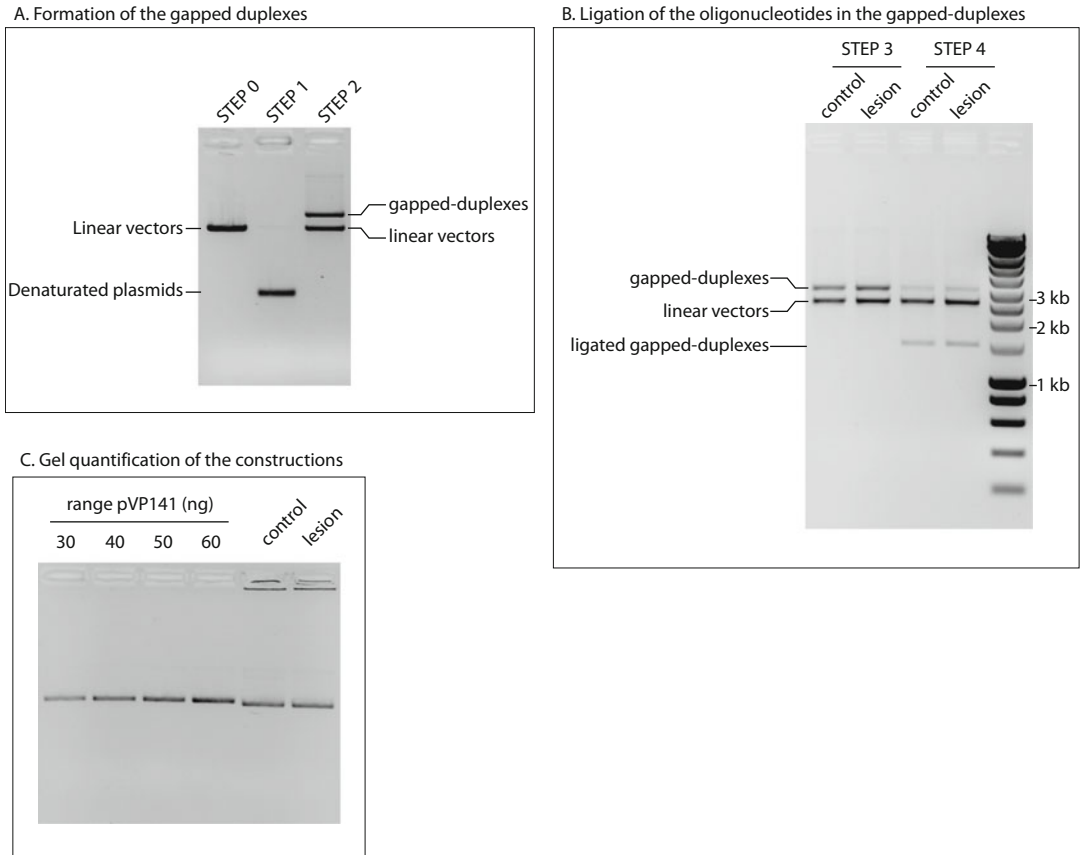
5. Add 40  $\mu$ l of 5 M NaCl, vortex.
6. Incubate at 55 °C in a water bath.
7. After 2 h incubation, shut off the water bath and allow it to slowly cool down to room temperature overnight.

→ freeze a 20  $\mu$ l aliquot for electrophoresis checking (STEP 2).

8. Control the efficiency of denaturation and gapped-duplex formation by migration of the aliquots of Steps 0, 1, and 2 on a 1.4% agarose gel electrophoresis without ethidium bromide (to better resolve the gapped duplex from the linear forms). After migration, stain the gel in TAE buffer, 0.5  $\mu$ g/ml ethidium bromide (Fig. 3a).



**Fig. 2** (continued) forms of gapped-duplexes will be generated, one with a gap that is complementary to the damaged oligonucleotide (*denoted "good" gapped duplex*) and one that has the same sequence as the oligonucleotides (*denoted "wrong" gapped duplex*). (c) Ligation of the control or damaged oligonucleotide in the gapped-duplex and isolation of the closed circular species by CsCl gradient purification allows recovering high purity damaged vector



**Fig. 3** Controlling the steps of vector production on agarose gel electrophoresis. **(a)** Visualization of the three steps leading to the formation of the gapped duplexes. *Step 0*: linear vectors after digestion. *Step 1*: denatured plasmids. *Step 2*: after re-annealing, formation of ~50% of linear homoduplexes and ~50% of gapped heteroduplexes. **(b)** Control of the ligation of the oligonucleotides (control and damaged) into the gapped-duplex molecule. Successful ligation events generate circular, covalently closed constructs that migrate faster during agarose gel electrophoresis in the presence of ethidium bromide. **(c)** Quality control and quantification of the construction: a range of the helper plasmid (pVP141) allows estimating the concentration of the construction after CsCl gradient purification. (Note that the electrophoretic migration of the circularized duplex is slightly faster than the pVP141 plasmid. Indeed, pVP141 was produced *in vivo* and therefore contains negative supercoiling, whereas the constructed duplex was ligated *in vitro* and contains thus no supercoils)

Ligation of the oligonucleotides (Fig. 2c).

The 15mer oligonucleotide containing the single lesion (G-AAF in our example) and the non-damaged control oligonucleotide are phosphorylated and purified on a polyacrylamide gel electrophoresis.

1. To the 5 ml gapped-duplex reaction (still in the glass tube), add a 1.5 to 3-fold excess of oligonucleotide (*see Note 3*).
2. Add 0.5 ml of 10× Gap Duplex ligation buffer.

3. Incubate for 2 min at 65 °C.
4. Incubate for 15 min at 16 °C.
5. Add 50 µl of 0.1 M ATP and 25 µl of 10 mg/ml BSA.  
→ take a 20 µl aliquot for electrophoresis checking (STEP 3).
6. Add 2 µl of T4 DNA ligase (400 U/µl).
7. Vortex briefly and incubate for 3 min at 16 °C (*see Note 4*).
8. Stop the reaction on ice.  
→ take a 20 µl aliquot for electrophoresis checking (STEP 4).
9. Control the ligation on a 0.8% agarose gel electrophoresis with ethidium bromide (Fig. 3b).
10. Inactivate the ligase by incubating for 10 min at 65 °C.
11. Transfer the reaction to a 50 ml falcon tube.
12. Precipitate the DNA by adding 15 ml of Ethanol/NaAcetate solution and incubate overnight at -20 °C.

Purification of the ligated gapped-duplex on CsCl gradient:

1. Recover the precipitated DNA by centrifugation for 30 min at 12,000 × *g*.
2. Resuspend in 1.2 ml TE/CsCL/Ethidium Bromide mix.
3. Transfer to an ultracentrifugation tube 8 × 35 mm and seal the tube.
4. Centrifuge for 5 h at 15,000 × *g* rpm at 20 °C on Beckman Optima™ Max-XP ultracentrifuge in rotor TLN120.
5. After centrifugation, recover the lower band under UV light (312 nm) using a syringe and a needle. The recovered volume is usually around 100 µl.

Removal of the ethidium bromide:

1. Add one volume of isopropyl-alcohol/TE/CsCl saturated.
2. Vortex.
3. Eliminate the upper phase where the ethidium bromide is dissolved.
4. Repeat two additional times.
5. Add two volumes of water to the aqueous phase.
6. Precipitate by adding two volumes (of the total volume) of 100% EtOH and incubate for 1 h at -20 °C.
7. Recover the precipitated DNA by centrifugation.
8. Wash the pellet with 150 µl of 100% EtOH.
9. Dry and resuspend the pellet in TE.



- Quantify and adjust the concentration of the damaged and non-damaged plasmid by quantification on agarose gel electrophoresis (Fig. 3c—Note 5).

### 3.4 Integration Protocol

#### 3.4.1 Competent Cells

- Inoculate 3 ml of LB, 50 µg/ml kanamycin with the recipient strain and grow O/N at 37 °C with shaking.
- Inoculate 100 ml of LB, 50 µg/ml kanamycin, 200 µM IPTG with 1 ml of the O/N starter culture.
- Grow at 37 °C under shaking until  $OD^{600} = 0.4$ .
- Spin and wash the cells with 100 ml of ice-cold water.
- Spin and wash the cells with 50 ml of ice-cold water.
- Spin and wash the cells with 2 ml of ice-cold 10% glycerol.
- Spin and resuspend the cells in 200 µl of ice-cold 10% glycerol.
- Divide the competent cells in 40 µl aliquot and freeze at -80 °C.

#### 3.4.2 Electroporation

- Add 1 µl of plasmid mix (gap-duplex + transformation control plasmid pVP146) at 1 ng/µl each to the 40 µl of thawed competent cell.
- Transfer to an electroporation cuvette (2 mm gap) and Pulse 2.5 kV, 25 µF, 200 Ω (time constant should be ~5 ms).
- Immediately add 1 ml of SOC, 200 µM IPTG (pre-warmed at 37 °C). Mix by pipetting and transfer 500 µl to 2 ml of LB, 200 µM IPTG (pre-warmed at 37 °C) (*see Note 6*).
- Incubate at 37 °C for 40 min with shaking.
- Dilute and plate the cells.

Integrations events: on LB-agar, 50 µg/ml Ampicillin, 80 µg/ml X-gal.

Transformation control: on LB-agar, 30 µg/ml Chloramphenicol.

- Incubate O/N at 37 °C.

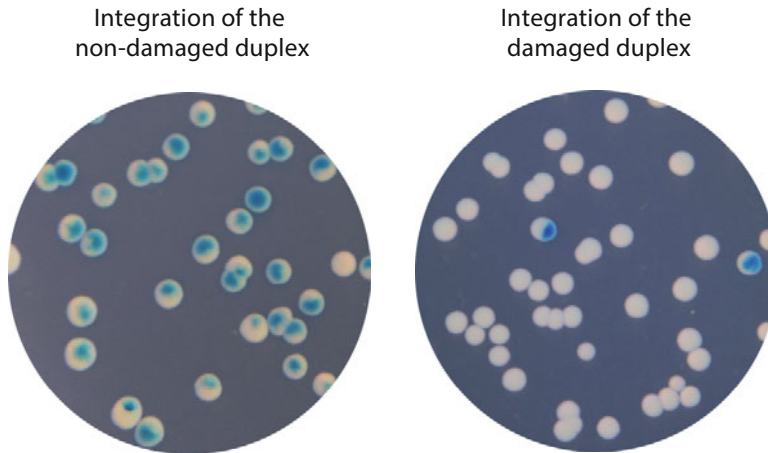
#### 3.4.3 Analysis of the Plates (Fig. 4)

Blue and white colonies are counted.

Blue colonies represent TLS events. The relative integration efficiencies of damaged vectors are compared to their non-damaged homologs, and normalized by the transformation efficiency of pVP146 plasmid in the same electroporation experiment, allow the overall rate of tolerance of the lesion to be measured.

### 3.5 Modification of the System

As discussed above, different DNA lesions can be inserted into the genome of living cells by using the gapped duplex methodology. In addition to the lesion of choice, genetic markers (i.e., mismatches) can be introduced upstream and/or downstream from the lesion to



**Fig. 4** Colonies following integration. After the integration of the non-damaged construct, all colonies exhibit *blue/white* sectors; the *white* and *blue* sectors represent the bacterial progeny that stems from the replication of the *lacZ*- and *lacZ*+ strands of the initial construct, respectively. After the integration of the single lesion-containing vector, only cells in which a TLS event occurred give rise to a sectored *blue/white* colony. Phenotypic monitoring of DA events requires a specific construct [11]

study strand-switching/recombination reactions. Care should be taken to identify proper mismatches that will be resistant to mismatch repair correction *in vivo*. For that purpose mismatch repair defective strains should be tested. We recently implemented a modified version of the above-described system that allows us to monitor homologous recombination events at the lesion site. This was achieved by the introduction of an additional mismatch upstream of the lesion site, allowing us to distinguish the damaged from the non-damaged strand, and to evidence any exchange of genetic information between the two strands [11].

---

## 4 Notes

1. Vectors containing the R6K origin can only be replicated in a strain that expresses the *pir* gene. To increase the yield of our plasmid prep, we use the strain EC100D *pir*-116, in which the *pir*-116 allele supports a higher copy number of R6K origin plasmids.
2. It is important to perform all the following steps in glass tubes, as plastic containers do not allow sufficient thermal exchange to insure proper plasmid denaturation efficiency.
3. Expected yield: the heat denaturation/renaturation cycle will produce both the two reconstituted linear plasmid forms (50%) and gapped-duplexes (~50%). Two molecular forms of gapped-duplexes will be generated with gaps of complementary sequences (denoted as “good” and “wrong” in Fig. 2b).

Thus, the maximal yield of gapped duplex carrying the proper gap will be  $\approx 25\%$ . From an initial amount of 50  $\mu\text{g}$  of parental plasmids (25  $\mu\text{g}$  each), the expected yield of proper gapped-duplex will correspond to  $\sim 7$  pmol of DNA (parental plasmid size 2.5 kb). Therefore, 10–21 pmol of oligonucleotide is added.

4. Since only three nicks have to be sealed, a 3 min ligation at 16 °C is enough. Further incubation will lead to re-circularization of the linear forms that we want to avoid.
5. There is a slight difference in migration between the construct and the control plasmid (Fig. 3c). This difference is due to differences in the supercoiling status, the covalently closed circular (ccc) construct having essentially no supercoils.
6. Recovery of the cells after electroporation in SOC only or LB only greatly reduces the integration efficiency. For this reason, cells are resuspended in SOC after electroporation and diluted in LB for the incubation.

## References

1. Pagès V, Mazon G, Naiman K, Philippin G, Fuchs RP (2012) Monitoring bypass of single replication-blocking lesions by damage avoidance in the *Escherichia coli* chromosome. *Nucleic Acids Res* 40:9036–9043
2. Naiman K, Philippin G, Fuchs RP, Pagès V (2014) Chronology in lesion tolerance gives priority to genetic variability. *Proc Natl Acad Sci U S A* 111:5526–5531
3. Pagès V, Fuchs RP (2003) Uncoupling of leading- and lagging-strand DNA replication during lesion bypass in vivo. *Science* 300:1300–1303
4. Higuchi K, Katayama T, Iwai S, Hidaka M, Horiuchi T, Maki H (2003) Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication in vitro. *Genes Cells* 8:437–449
5. Esnault E, Valens M, Espéli O, Boccard F (2007) Chromosome structuring limits genome plasticity in *Escherichia coli*. *PLoS Genet* 3:e226
6. Weiss DS, Chen JC, Ghigo JM, Boyd D, Beckwith J (1999) Localization of FtsI (PBP3) to the septal ring requires its membrane anchor, the Z ring, FtsA, FtsQ, and FtsL. *J Bacteriol* 181:508–520
7. Koehl P, Burnouf D, Fuchs RP (1989) Construction of plasmids containing a unique acetylaminofluorene adduct located within a mutation hot spot: a new probe for frameshift mutagenesis. *J Mol Biol* 207:355–364
8. Inuzuka M (1985) Plasmid-encoded initiation protein is required for activity at all three origins of plasmid R6K DNA replication in vitro. *FEBS Lett* 181:236–240
9. Bauer CE, Hesse SD, Gumpert RI, Gardner JF (1986) Mutational analysis of integrase arm-type binding sites of bacteriophage lambda. Integration and excision involve distinct interactions of integrase with arm-type sites. *J Mol Biol* 192:513–527
10. Heilig, J.S., Elbing, K.L. and Brent, R. (2001) Large-scale preparation of plasmid DNA. *Curr Protoc Mol Biol*, Chapter 1, Unit 1.7
11. Laureti L, Demol J, Fuchs RP, Pagès V (2015) Bacterial proliferation: keep dividing and don't mind the gap. *PLoS Genet* 11:e1005757

# Chapter 10

## A qPCR-Based Protocol to Quantify DSB Resection

Matteo Ferrari, Shyam Twayana, Federica Marini, and Achille Pelliccioli

### Abstract

The nucleolytic degradation of the 5'-ending strand of a Double-Strand DNA break (DSB) is necessary to initiate homologous recombination to correctly repair the break. This process is called DNA end resection and it is finely regulated to prevent genome rearrangements. Here, we describe a protocol to quantify DSB resection rate by qPCR, which could be applied to every organisms whenever the break site and its flanking region sequences are known.

**Key words** Double-strand breaks, DNA end resection, Single-strand DNA, qPCR

---

### 1 Introduction

DSBs can be repaired thorough the alternative processes of ends joining (NHEJ, non homologous ends joining) or homologous-directed recombination (HR). NHEJ consists of a ligation of two DNA ends without needing any homology, while a homologous sequence is used to repair the break through HR. The coordinated actions of multiple nucleases and associated factors mediate the nucleolytic degradation of the 5'-ending strands of a DSB (also called DSB resection). This process generates the 3'-ended single-strand DNA (ssDNA) filament, which is the first intermediate of all the HR-mediated DSB repair mechanisms. When a DSB is resected, NHEJ-mediated repair is no longer possible. Therefore, all those mechanisms that regulate DSB resection are fundamental to promote HR and to determinate the choice between NHEJ and HR repair events. Indeed, CDK1-mediated phosphorylation of critical factors restricts efficient DSB resection and HR in S and G2/M cell cycle phases [1].

In *S. cerevisiae* it is possible to induce a DSB through the over-expression of the HO endonuclease, which normally cuts a specific sequence at the *MAT* locus. However, in the last years, a number of strains have been generated in which the HO-cut site has been moved to different genomic loci, thus allowing testing different

aspects of DSB processing and repair [2]. Therefore, *S. cerevisiae* has been considered an ideal model organism to study DSB response and different methods have been described to test DSB resection in this organism. Early methods were based on Southern blotting approaches to visualize either the formation of the 3' ssDNA filament or, alternatively, the disappearance of dsDNA fragments in the flanking region of a DSB, using specific probes [3]. These methods are very useful to visualize DSB resection, but are not always easy and reproducible to measure and quantify resection rate. Therefore, a qPCR-based method has been described, which can give an accurate quantification of DSB resection speed [4]. A comparison between this qPCR-based method and a Southern blotting approach to study the processing of the HO-cut at the *MAT* locus can be seen in our recent works [5, 6].

Here, we illustrate the protocol to apply the qPCR-based technique to measure DSB resection in yeast; however, thanks to its power and versatility this method has been recently applied to other organisms. Remarkably, the same protocol can be applied, with minor changes, to *Azi*SI-induced DSBs in human genome [7, 8].

---

## 2 Materials

### 2.1 Yeast Cells Growing and In Vivo DSB Induction

1. *S. cerevisiae* JKM139 derivative strains (J. Haber, Waltham University, USA).
2. Yeast liquid growth media (YP).
3. 30% Raffinose stock in water, filter sterilized.
4. 30% Galactose stock in water, filter sterilized.
5. 10% Sodium Azide solution in water, filter sterilized.
6. Flasks.
7. Incubator.

### 2.2 Genomic DNA Extraction, Digestion, and Purification

1. Reagents to prepare genomic DNA.
2. *Sty*I, *Rsa*I restriction enzymes and appropriate buffers.
3. 2-Propanol.
4. 70% EtOH solution in water.
5. TE 1×, filter sterilized (10 mM Tris-HCl pH 7.5, 1 mM EDTA-NaOH pH 8).

### 2.3 qPCR Reagents

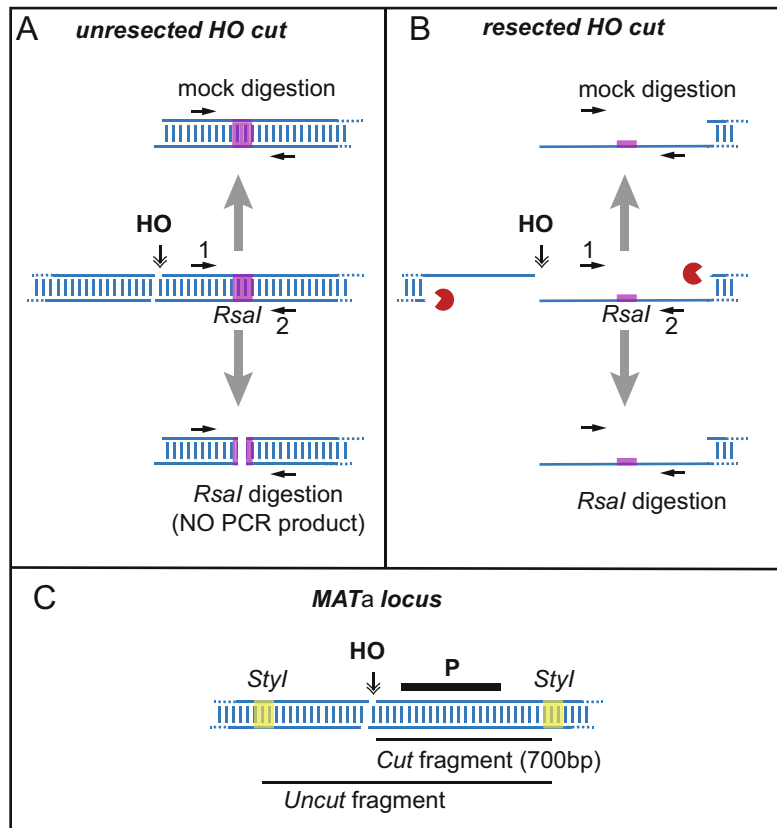
1. Real Time PCR machine.
2. 96-well plates with appropriate adhesive sealing films.
3. 2× Master Mix with SYBR Green.
4. Ultrapure water (MilliQ).
5. Customized DNA primers (*see* qPCR reaction step).

## 2.4 Southern Blot Reagents

Although we do not describe the Southern Blotting procedure, materials to perform this step of the protocol are needed.

## 3 Methods

Following the HO break induction, genomic DNA is purified and prepared for the analysis of end resection. The technique we describe here relies on the ability of restriction enzymes to cut selectively dsDNA but not ssDNA; therefore, the ssDNA generated by DSB resection is resistant to endonucleolytic cleavage and can be used as a template for the amplification of short DNA fragments either close or far from the break. The rationale of the procedure is shown in Fig. 1. The genomic DNA is either cut or mock cut with appropriate restriction enzyme (in our example in *S. cerevisiae* we



**Fig. 1** Schematic representation of the qPCR method used to monitor DSB resection of one HO cut induced at *MATa* locus in *S. cerevisiae*. Expected outcomes if the DSB ends are not resected (a) or resected (b). The relative positions of the *StyI* and HO sites at the *MATa* locus are shown in a scheme (c). “P” is the probe used to visualize the Uncut and Cut fragments by Southern blotting

used *RsaI*, an enzyme that digests the yeast genome frequently, therefore is a useful tool to study resection at many locations). Another sample of the genomic DNA is cut with a different enzyme (in our example we used *StyI*, which cuts opportunely at *MATa* locus), to measure the HO cut efficiency through a Southern Blotting procedure. Indeed, HO cut efficiency can vary among different strains and experiments; therefore, it is necessary to test the HO cut efficiency to normalize the results.

### 3.1 Cell Sample Collection

1. Inoculate JKMI39 derivative strains in YP + 3% Raffinose, grow O/N at 28 °C until cells reach the exponential phase.
2. The day after, synchronize the cell culture in G2/M with Nocodazole (*see Note 1*).
3. Add 2% galactose to induce HO endonuclease expression (in the indicated strain HO cuts only in the *MATa* locus, *see Fig. 1*).
4. Collect 50 ml of cell culture sample for each time point in Falcon tubes. Immediately, add Sodium Azide to the sample (0.1% final) and keep the sample on ice till the end of the collection of all the samples to block whole cell metabolisms and freeze all the DSB processing enzymes.
5. After the collection of all the samples, centrifuge the tubes at around  $3200 \times g$  for 5 min and eliminate the supernatant. Keep the pellets on ice at 4 °C until you are ready to process samples (*see Note 2*).

### 3.2 Extraction and Digestion of the Genomic DNA

1. Prepare genomic DNA from each sample by a standard procedure (*see Note 3*).
2. Divide each genomic DNA sample (50–100 ng/ $\mu$ l) into 2 aliquots of 25  $\mu$ l each in 1.5 ml tube.
3. Prepare the following mix:

Mix 1	Mix 2 (Mock digestion)
25 $\mu$ l genomic DNA	25 $\mu$ l genomic DNA
5 $\mu$ l 10 $\times$ Buffer	5 $\mu$ l 10 $\times$ Buffer
0.5 $\mu$ l <i>RsaI</i> enzyme	0 $\mu$ l <i>RsaI</i> enzyme
19.5 $\mu$ l MilliQ H <sub>2</sub> O	20 $\mu$ l MilliQ H <sub>2</sub> O

Incubate at 37 °C for 6 h.

4. Precipitate DNA with an equal volume (50  $\mu$ l) of 2-Propanol and centrifuge at 16,000 RCF for 30 min at RT.
5. Eliminate the supernatant and wash the pellet with 1 ml of cold 70% EtOH.

6. Centrifuge for 5 min at 16,000 RCF, eliminate the supernatant, and dry the pellet in a vacuum pump centrifuge.
7. Resuspend the pellets in 100  $\mu\text{l}$  TE 1 $\times$ .

### 3.3 qPCR Reaction

1. Dilute the genomic DNA in MilliQ water at a working concentration of  $\sim 0.1$  ng/ $\mu\text{l}$ . Normally, we have a genomic DNA concentration of 50–100 ng/ $\mu\text{l}$ ; therefore, we dilute it 500–1000 $\times$ . We also prepare three different dilutions of the mock-digested time 0 sample to perform the standard curve (*see* **Note 4**).
2. We suggest using primers with a similar melting temperature and an amplicon length of 150–250 base pairs. Design primers at the desired distance from the DSB (primers 1 and 2; *see* Fig. 1). All the amplicons should have a *RsaI* site. Moreover, an amplicon on a different chromosome in which neither *RsaI* nor HO are cutting is essential to normalize all the PCR values. For this analysis you should design specific primers (3 and 4, at the *PRE1* gene locus on chromosome V in our case). *See* **Note 5** for the sequence of all the primers used in our practical examples.
3. Prepare the following 2 PCR reaction mixes:

CTRL MIX (as a negative control)	DSB MIX
12.5 $\mu\text{l}$ Master mix with SYBR green (2 $\times$ )	12.5 $\mu\text{l}$ Master mix with SYBR green (2 $\times$ )
0.5 $\mu\text{l}$ Primer 3 (10 $\mu\text{M}$ )	0.5 $\mu\text{l}$ Primer 1 (10 $\mu\text{M}$ )
0.5 $\mu\text{l}$ Primer 4 (10 $\mu\text{M}$ )	0.5 $\mu\text{l}$ Primer 2 (10 $\mu\text{M}$ )
6.5 $\mu\text{l}$ MilliQ water	6.5 $\mu\text{l}$ MilliQ water

4. Add 5  $\mu\text{l}$  of each diluted sample in the 96-well PCR plate (we suggest analyzing all the samples in triplicate).
5. Add 20  $\mu\text{l}$  of the reaction mix in each well and seal the plate (for Master Mix details, *see* **Note 6**).
6. Perform the real-time PCR with the following program (considering the annealing temperature of our primers we use a two-step program):
  - (a) 95  $^{\circ}\text{C}$  for 1' 30"
  - (b) 95  $^{\circ}\text{C}$  for 15"
  - (c) 66  $^{\circ}\text{C}$  for 40"
  - (d) GO TO 2 for 49 times
  - (e) Melt curve 60–95  $^{\circ}\text{C}$ , increment 0.5  $^{\circ}\text{C}$



- Analyze the PCR results with the specific software of your qPCR machine. Check the linearity of the PCR reaction performing a standard curve analysis (efficiency of the reaction between 90 and 105%) and the specificity of the DNA amplification by melting curve analysis.

---

## 4 Data Analysis

### 4.1 Example A

In this section, we illustrate a practical example of how to calculate the ssDNA accumulation at 0.15 Kb from one HO cut at the *MATa* locus in wild-type yeast cells. For each PCR reaction you will obtain a *Ct* value, which is the cycle in which the fluorescence of the sample became higher than the background signal and then detectable by the machine (threshold cycle). Consequently, the higher is the amount of DNA in the sample the lower is the *Ct* value.

To calculate the amount of ssDNA at a defined time and position, four *Ct* values are required: two obtained using primers 1 and 2, at the DSB locus, and two using primers 3 and 4, at the control locus using mock and *RsaI* digested DNA in the PCR reactions (*see* Fig. 1).

Use the following formula:

$$\Delta Ct_{\text{digested}} = Ct_{\text{digested DSB}} - Ct_{\text{digested CTRL}}$$

$$\Delta Ct_{\text{mock}} = Ct_{\text{mock DSB}} - Ct_{\text{mock CTRL}}$$

Calculate the  $\Delta\Delta Ct$  of each time point and distance from the break as:

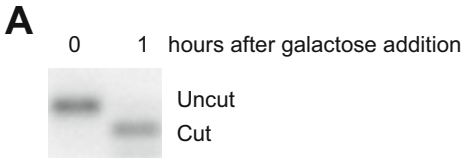
$$\Delta\Delta Ct = \Delta Ct_{\text{digested}} - \Delta Ct_{\text{mock}}$$

$$\%ssDNA = \left\{ 100 / \left[ (1 + 2^{\Delta\Delta Ct}) / 2 \right] / f \right\} = \%DSB \text{ resected}$$

$f$  = HO cut efficiency.

There are different procedures to measure the  $f$  value (HO cut efficiency). Here, we describe a Southern blotting-based method. We recommend collecting two cell samples (one at time 0 and one within 1 h after the HO cut induction) to get a reliable  $f$  value by this procedure.

- Prepare genomic DNA from each sample by a standard procedure (*see* Note 3).
- Digest 3–5  $\mu\text{g}$  of genomic DNA with *StyI* enzyme.
- Load an aliquot of the *StyI* digested DNA and run it at 50 V for 5 h.
- Transfer the DNA on a nylon membrane using a standard Southern blotting protocol.
- After hybridization with the desired probe, that has to be inside the HO locus, quantify the signal for normalization (*see* Fig. 1c for a scheme of the *MATa* locus and probe).

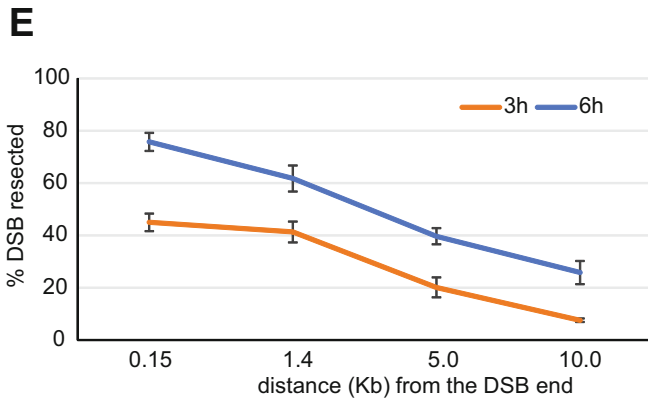
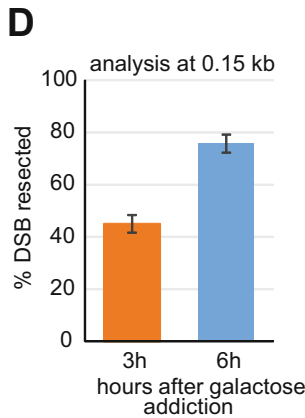


**B** Example: cut efficiency ( $f$ ) calculation

	densitometric value	$f$
Uncut 1h	157701.71	
Cut 1h	1893839.58	0.92313013

**C** Example: % DSB resected calculation

	$Ct$ Digested	$\Delta Ct$ Digested	$Ct$ Mock	$\Delta Ct$ Mock	$\Delta\Delta Ct$	Ratio	$f$	% of DSBs resected
DSB 0h	26.89	4.68	23.94	-1.4	6.1	67.6	0	0
DSB 3h	21.52	1.13	20.45	-0.91	2	4.11	0.92	42.38
DSB 6h	20.51	-0.18	20.94	-1.25	1.1	2.1	0.92	69.90
Ctrl 0h	22.21		25.34					
Ctrl 3h	20.39		21.36					
Ctrl 6h	20.69		22.19					



**Fig. 2** Data analysis: Example A. The Southern blotting procedure to measure the HO cut efficiency is shown in (a), together with the calculation of the  $f$  value in (b). The densitometry of the bands is calculated using the ImageJ software. (c) Data obtained from qPCR reactions and calculation of the % DSB resected. “DSB” indicates PCR reactions close to the HO cut site (primers 1 and 2); “CTRL” indicates PCR reactions obtained using primers 3 and 4.  $\Delta Ct$  values and the % DSB resected are calculated using the formula shown in the text. (d) The plot shows the mean values of three different experiments in which DSB resection has been measured at 0.15 kb from the HO cut, 3 and 6 h after galactose addition. Error bars show the Standard Error of the Mean (SEM). (e) Measurement of the % DSB resected at different distances from DSB end, 3 and 6 h after galactose addition. The mean of three different experiments is plotted. Error bars indicate the SEM

6. Perform a densitometric analysis and calculate the  $f$  value as follows:

$$f = \text{Cut value} / (\text{Cut value} + \text{Uncut value}) = \text{HO Cut efficiency}$$

In our example (see Fig. 2), we show the data and the graph of the % DSB resected at 0.15 kb from the HO cut. Using the same procedure, we also calculate resection rate at different positions far from the break (see Fig. 2e).

## 4.2 Example B

This qPCR-based protocol has been originally described to analyze a DSB in *S. cerevisiae*. However, it can be applied to every organism whenever the break site and its flanking region sequences are known. To give an example, in this section, we apply the protocol to one DSB induced in chromosome I in U2OS human cell line, stably expressing the *AsiSI* enzyme fused to the oestrogen receptor (*AsiSI-ER*). Upon tamoxifen (4OHT) treatment, *AsiSI-ER* enters into the nucleus and cuts DNA at defined positions [7]. After treating or mock treating cells with 4OHT, genomic DNA is extracted with a standard commercial kit. Normally, for each time point we use 90% confluent cells grown in a well of a 6-well plate and obtain a final concentration of DNA around 100 ng/ $\mu\text{l}$ . Then, 15  $\mu\text{l}$  of genomic DNA is digested or not with *BsrGI* enzyme, as follows:

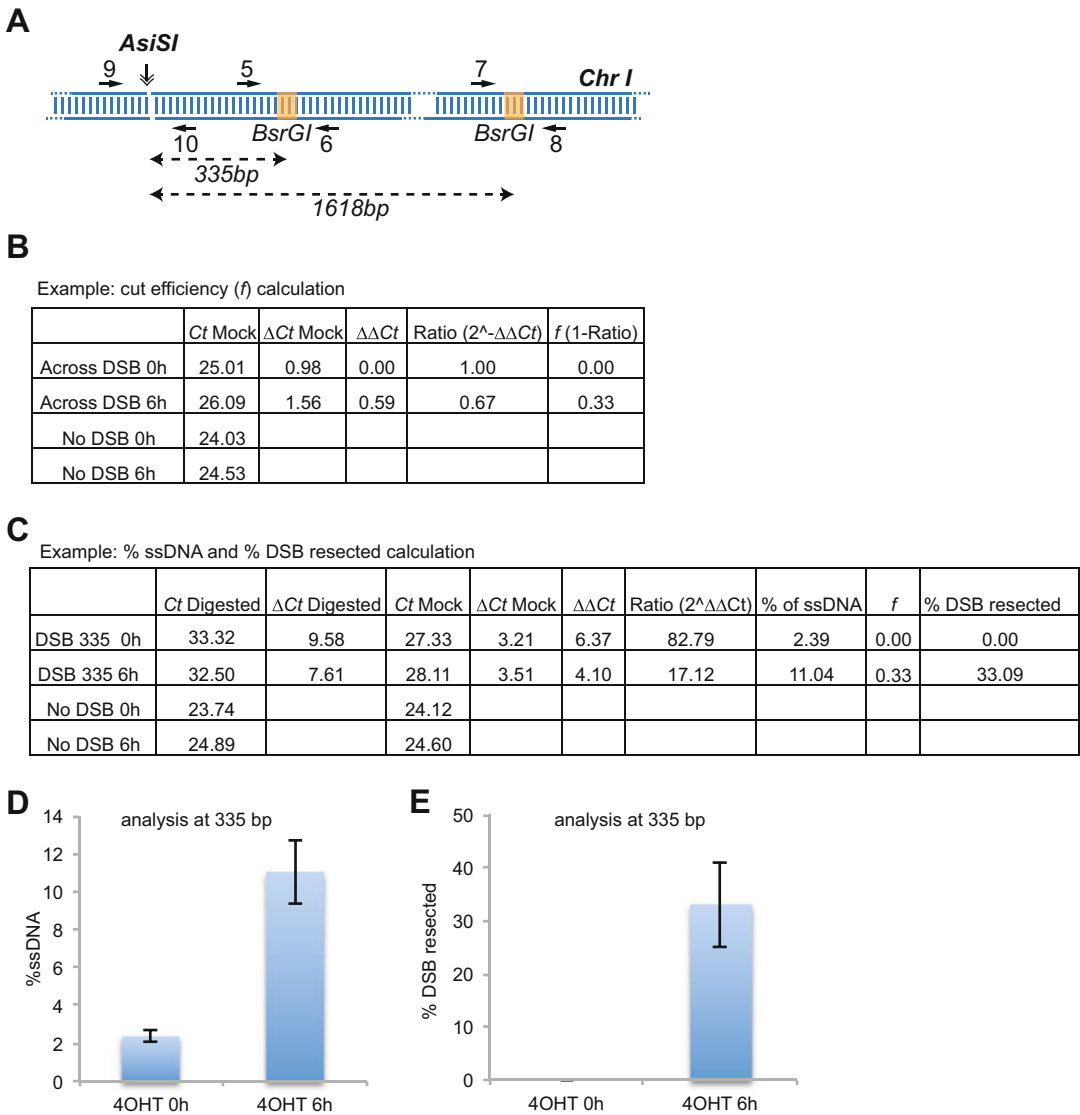
Mix 1	Mix 2 (Mock digestion)
15 $\mu\text{l}$ genomic DNA	15 $\mu\text{l}$ genomic DNA
9 $\mu\text{l}$ 10 $\times$ Buffer	9 $\mu\text{l}$ 10 $\times$ Buffer
2 $\mu\text{l}$ <i>BsrGI</i> enzyme	0 $\mu\text{l}$ <i>BsrGI</i> enzyme
64 $\mu\text{l}$ MilliQ H <sub>2</sub> O	66 $\mu\text{l}$ MilliQ H <sub>2</sub> O

For the qPCR reaction 3  $\mu\text{l}$  (approximately 50 ng of genomic DNA) of each mix is used. Master Mix, DSB-Mix, and CTRL-Mix are prepared as described above. See **Note 5** for the sequences of the primers used in U2OS cells.

The *AsiSI-ER* cut efficiency ( $f$ ) is determined by PCR, using specific primers (9 and 10) that anneal at 3' and 5' of the cut site (Fig. 3a).

The % DSB resected is calculated taking into consideration the  $f$  value, with the same formula describe above in example A.

It is important to underline that in the literature, people who use the *AsiSI-ER-U2OS* system often show the % ssDNA obtained with the same procedure without normalization with the  $f$  value in their graphs. However, contrary to the HO system in yeast, the cut efficiency of *AsiSI-ER* changes significantly among the experiments, at different times after the enzyme induction and at different genomic sites. Therefore, the percentage of ssDNA could significantly differ from the percentage of DSB resected value and the normalization with the  $f$  value is determinant for the reliability of the results.



**Fig. 3** Data analysis: Example B. (a) The relative positions of *AsiSI* and *BsrGI* sites analyzed by qPCR on Chromosome I in U2OS cells are shown in a scheme. Arrows indicate the primers used for the qPCR. Primers 9 and 10 were used to determine the cut efficiency ( $f$ ). (b) Data obtained from qPCR reactions to calculate the  $f$  value. “Across DSB” indicates  $Ct$  values obtained using primers 9 and 10, while “No DSB” indicates  $Ct$  values achieved using primers 11 and 12 (control genomic site where *AsiSI* does not cut) at different time points as indicated. The template DNA was not digested with *BsrGI* (Mock).  $\Delta Ct$  is obtained by subtracting  $Ct$  of “No DSB” from  $Ct$  of “Across DSB”. (c) Data obtained from qPCR reactions and calculation of % ssDNA and % DSB resected. “DSB 335” indicates the  $Ct$  values obtained using primers 5 and 6, “No DSB” indicates  $Ct$  values as in (b), with template DNA taken from either Mix1 (*BsrGI* digested) or Mix2 (Mock digested) at different time points as indicated.  $\Delta Ct$  values, % of ssDNA, and % of DSB resected are calculated using the formula shown in the text. The data calculated in (c) at 335 bp from the *AsiSI* cut, 0 h and 6 h after of 4OHT addition were plotted as % ssDNA (d) and % DSB resected (e). Error bars indicate Standard Error of Mean (SEM)

## 5 Notes

1. As described in the introduction, resection is cell cycle regulated; therefore, it is important to analyze it in cells blocked in a specific cell cycle phase.
2. Here one can eventually keep cells on ice overnight.
3. There are many different protocols to extract genomic DNA from yeast cells. Some of them are hand-made protocol, while others are based on commercial kits. Based on our experience, the protocol used to extract genomic DNA is not critical for the DSB resection analysis; therefore, you can perform this step according to your lab indication. We normally use an adapted Teeny prep protocol derived from the original protocol by Dr. J. Boeke (NYU). We obtain DNA at a final concentration of 50–100 ng/ $\mu$ l.
4. Analysis of a standard curve is required to verify that the concentration of each sample is within the linear range of the PCR reaction. We use three different concentrations of the mock digested sample to test the linearity of the PCR reaction. The results of the PCR can be considered suitable for the resection analysis only if the efficiency of the reaction is between 90 and 105% and the  $C_t$  value of all the samples analyzed is encompassed in the standard curve range.
5. List of primers in yeast:

1\_ CCTGGTTTTGGTTTTGTAGAGTGG;  
2\_ GAGCAAGACGATGGGGAGTTTC;  
3\_ CCCACAAGTCCTCTGATTTACATTCG;  
4\_ ATTTCGATTGACAGGTGCTCCCTTTTC

List of primers in U2OS human cells.

5\_ AATCGGATGTATGCGACTGA;  
6\_ AAAGTTATTCCAACCCGATCC;  
7\_ TGAGGAGGTGACATTAGAACTCAGA;  
8\_ AGGACTCACTTACACGGCCTTT;  
9\_ GATGTGGCCAGGGATTGG;  
10\_ CACTCAAGCCCAACCCGT;  
11\_ ATTGGGTATCTGCGTCTAGTGAGG;  
12\_ GACTCAATTACATCCCTGCAGCT

6. Many companies sell Master Mix for qPCR protocols. We use a Master Mix that contains all the reagents for qPCR, including the SYBR green. Other qPCR protocols require TaqMan enzyme and specific primers.

---

## Acknowledgments

This work was supported by grants from Associazione Italiana per la Ricerca sul Cancro [AIRC\_IG Grant n.15488 to A.P.; CARIPLO [2013-0790 to A.P.]; a fellowship from Fondazione “Gabriella Dolfin Voyasidis”-Accademia Nazionale dei Lincei to M.F.

## References

1. Symington LS (2016) Mechanism and regulation of DNA end resection in eukaryotes. *Crit Rev Biochem Mol Biol* 51:195–212
2. Sugawara N, Haber JE (2012) Monitoring DNA recombination initiated by HO endonuclease. *Methods Mol Biol* 920:349–370
3. White CI, Haber JE (1990) Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *EMBO J* 9:663–673
4. Zierhut C, Diffley JF (2008) Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. *EMBO J* 27:1875–1885
5. Ferrari M, Dibitto D, De Gregorio G, Eapen VV, Rawal CC, Lazzaro F, Tsabar M, Marini F, Haber JE, Pellicoli A (2015) Functional interplay between the 53BP1-ortholog Rad9 and the Mre11 complex regulates resection, end-tethering and repair of a double-strand break. *PLoS Genet* 11:e1004928
6. Dibitto D, Ferrari M, Rawal CC, Balint A, Kim T, Zhang Z, Smolka MB, Brown GW, Marini F, Pellicoli A (2015) Slx4 and Rtt107 control checkpoint signalling and DNA resection at double-strand breaks. *Nucleic Acids Res* 44 (2):669–682
7. Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube G (2010) High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J* 29:1446–1457
8. Zhou Y, Caron P, Legube G, Paull TT (2014) Quantitation of DNA double-strand break resection intermediates in human cells. *Nucleic Acids Res* 42:e19

# Chapter 11

## Alkaline Denaturing Southern Blot Analysis to Monitor Double-Strand Break Processing

Chiara Vittoria Colombo, Luca Menin, and Michela Clerici

### Abstract

Generation of 3' single-stranded DNA (ssDNA) tails at the ends of a double-strand break (DSB) is essential to repair the break through accurate homology-mediated repair pathways. Several methods have been developed to measure ssDNA accumulation at a DSB in the budding yeast *Saccharomyces cerevisiae*. Here, we describe one of these assays, which is based on the inability of restriction enzymes to cleave ssDNA. Digestion of genomic DNA prepared at different time points after DSB generation leads to the formation of ssDNA fragments whose length increases as the 5' strand degradation proceeds beyond restriction sites. After the separation by electrophoresis on alkaline denaturing agarose gel, these ssDNA fragments can be visualized by hybridization with an RNA probe that anneals with the 3'-undegraded DSB strand. This assay allows a direct and comprehensive visualization of DSB end processing.

**Key words** DNA double-strand breaks, Resection, Single-stranded DNA, HO endonuclease, *MAT* locus, Southern blot, Electrophoresis, Alkaline denaturing conditions, RNA probe

---

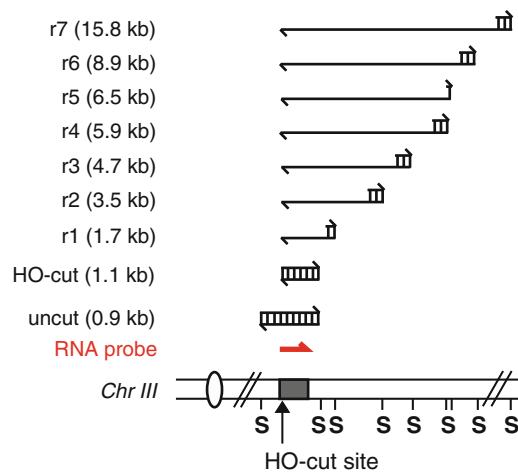
### 1 Introduction

DNA double-strand breaks (DSBs) can be repaired either by non-homologous end-joining (NHEJ), which directly rejoins the two broken ends together, or by homologous recombination (HR) that uses intact homologous duplex DNA sequences as a template for accurate repair. HR initiates with the nucleolytic degradation of the 5' DNA strand on both the DSB sides (a process referred to as resection) to yield 3' single-stranded DNA (ssDNA) tails. One of these tails invades the homologous duplex and primes reparative DNA synthesis [1, 2]. The initiation of resection is thought to channel DSB repair to HR and is tightly regulated [3, 4].

Much of our knowledge of the DNA end resection mechanism and regulation comes from genetic and biochemical studies in *Saccharomyces cerevisiae*, where DNA end processing of a DSB located at a specific genomic locus can be followed by both PCR-based and electrophoretic methods. Among these methods, a

Southern blot analysis of genomic DNA that has been run on an alkaline denaturing agarose gel [5] allows a direct and comprehensive visualization of resection. When genomic DNA from a time-course experiment is used, the kinetics of both resection initiation and elongation can be determined in a single gel blot. This assay is based on the inability of many restriction enzymes to cleave ssDNA. As resection proceeds beyond restriction sites, additional, slower migrating bands are revealed by an RNA or ssDNA probe that anneals to the 3'-unresected strand on one side of the break (Fig. 1).

Here, we describe the protocol to monitor DNA end processing in haploid yeast cells where a DSB is generated at the *MAT* locus by the homothallic switching endonuclease HO, whose expression is regulated by a galactose inducible promoter [6]. The same method can be used to study ssDNA generation at DSBs located at different genomic loci and/or generated by different nucleases. This method was successfully employed to monitor end resection of both an HO-induced DSB at *LEU2* locus [7] and a Spo11-induced DSB at the *YCR048W* meiotic recombination hot-spot [8]. Furthermore, a modified version of this assay that used denaturing polyacrilamide gels allowed us to measure ssDNA accumulation at telomeres [9]. Resection analysis at DSBs located in different genomic loci requires a detailed analysis of each locus to



**Fig. 1** System to detect DSB end resection at the *MAT* locus. Schematic representation of the region immediately centromere-distal to the *MAT* HO site (*bottom*), and of the DSB and 5'-to-3' resection products (*top*) detectable with the indicated RNA probe after alkaline gel electrophoresis of *SspI* (S)-digested DNA. The probe reveals a 1.1 kb fragment representing the uncut *MAT* locus. When HO cuts this locus, a smaller HO-cut fragment is produced. 5'-to-3' nucleolytic degradation progressively eliminates *SspI* sites, generating larger ssDNA *SspI* fragments (r1–r7) detected by the probe



determine both the most appropriate restriction enzyme and the probe sequence for Southern blot.

Extensive resection of a DSB induced at *MAT* locus can be monitored in JKM139 derivative strains [6]. In these strains the *GAL-HO* fusion is stably integrated in the genome and allows the expression of the HO endonuclease when galactose is added to the media. As transcription from the *GAL* promoter is repressed by glucose, cells are grown in raffinose-containing media before galactose addition. The HO endonuclease recognizes a 24-bp cleavage sequence, which is unique in the yeast genome and is located adjacent to the Y region at the *MAT* locus. HO cleavage yields 4-bp, 3'-overhanging ends [10]. JKM139 derivative strains also carry the deletion of both the *HML* and *HMR* cassettes, which are homologous to the *MAT* locus and are normally used to repair the HO-induced DSB by a gene conversion event that triggers mating type switching [10]. Therefore, HO expression in these cells generates an irreparable DSB, whose 5' end degradation proceeds for several kilobases.

With the aim of highlighting some critical steps, here we divide the protocol into three main parts: (1) DSB induction in yeast cells; (2) genomic DNA extraction; (3) alkaline denaturing electrophoresis and Southern blot analysis. Indeed, we experienced that, beside the alkaline Southern blot, the efficient and simultaneous DSB induction in the yeast population, as well as the preparation of a high quality genomic DNA, is crucial to obtain a clear result with good resection kinetics.

---

## 2 Materials

Prepare all the solutions using filtered and deionized ultrapure water (ddH<sub>2</sub>O; resistivity 18.2 MΩ·cm at 25 °C) and analytical grade reagents. Prepare and store all solutions at room temperature (unless indicated otherwise).

### 2.1 Double-Strand Break Induction in Yeast Cells

1. YEPD: 2% Bacto peptone, 1% yeast extract, 2% D-(+)-glucose monohydrate, 0.005% adenine hemisulfate salt. Dissolve in ddH<sub>2</sub>O. Autoclave.
2. YEP + raffinose: dissolve 2% Bacto peptone, 1% yeast extract, 0.005% adenine hemisulfate salt in ddH<sub>2</sub>O. Autoclave. Add sterilized raffinose from 30% solution to 2%.
3. 30% raffinose: dissolve D-(+)-raffinose pentahydrate in ddH<sub>2</sub>O. Autoclave or sterilize by filtration.
4. 30% galactose: dissolve D-(+)-galactose ≥99.0% in ddH<sub>2</sub>O. Sterilize by filtration. Do not autoclave (*see Note 1*).

## 2.2 Genomic DNA Extraction

1. Spheroplasting solution: 0.9 M sorbitol, 0.1 M ethylenediaminetetraacetic acid (EDTA), pH 7.5.
2. Zymolyase solution: dissolve 2 mg/mL Zymolyase 20T<sup>®</sup> (see **Note 2**) from *Arthrobacter luteus* (Nacalai Tesque) in spheroplasting solution + 14 mM  $\beta$ -mercaptoethanol.
3. 1 $\times$  TE: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, pH 7.4. Autoclave.
4. Lysis solution: 2.2% sodium dodecyl sulfate (SDS), 278 mM EDTA, pH 8.5, 445 mM Tris-base. Prepare the lysis solution just before adding it to the samples by mixing the appropriate amounts of 10% SDS, 0.5 M EDTA, pH 8.5, 2 M Tris-base stock solutions in a 15 mL tube.
5. 5 M potassium acetate: dissolve 5 M potassium acetate in ddH<sub>2</sub>O. Autoclave.
6. Ice-cold ethanol: 96% and 70%. Prepare a 70% ethanol solution by diluting 96% ethanol in ddH<sub>2</sub>O. Store at -20 °C aliquots of both 70% and 96% ethanol in glass bottles.
7. RNase solution: dissolve 10 mg/mL RNase A, DNase-free, in 10 mL 10 mM Tris-HCl, pH 7.5, 15 mM NaCl. Heat to 100 °C for 5 min and cool down at room temperature. Prepare small aliquots and store them at -20 °C.
8. 2-propanol anhydrous 99.5%.
9. 6 $\times$  DNA loading buffer: 30% glycerol, 0.25% bromophenol blue.
10. Regular agarose gel: melt 0.8% agarose in 1 $\times$  TAE buffer. Cool at approximately 60 °C and add 10 mg/mL ethidium bromide solution to a final concentration of 1  $\mu$ g/mL. Pour the gel into a gel tank and insert a comb.
11. 1 $\times$  TAE (Tris-acetate-EDTA) buffer: 40 mM Tris, 20 mM acetic acid, 1 mM EDTA. Prepare a 50 $\times$  TAE buffer: for 1 L dissolve 242 g of Tris-base in approximately 600 mL ddH<sub>2</sub>O. Add 57.1 mL glacial acetic acid and 100 mL 0.5 M EDTA, pH 8.0, and bring final volume to 1 L with ddH<sub>2</sub>O. Autoclave. Before use, dilute in ddH<sub>2</sub>O to a final concentration of 1 $\times$ .
12. Ethidium bromide solution: prepare a stock of 10 mg/mL ethidium bromide in ddH<sub>2</sub>O and store in light-tight containers (see **Note 3**).
13. UV lamp with camera.

## 2.3 Alkaline Denaturing Electrophoresis and Southern Blot

1. SspI restriction enzyme (20,000 U/mL; New England Biolabs) and buffer supplied from distributor.
2. 3 M sodium acetate, pH 5.2: dissolve 3 M sodium acetate in ddH<sub>2</sub>O. Adjust the pH to 5.2 with glacial acetic acid. Autoclave.

### 2.3.1 DNA Digestion and Denaturation

3. 0.5 M EDTA, pH 8.0: dissolve 0.5 M EDTA in ddH<sub>2</sub>O. Adjust the pH to 8.0 with sodium hydroxide (NaOH). Autoclave.
4. 1× alkaline loading buffer: 50 mM NaOH, 1 mM EDTA, pH 8.0, 2.5% ficoll (type 400) in ddH<sub>2</sub>O, 0.025% bromophenol blue.

### 2.3.2 Alkaline Denaturing Gel Electrophoresis and Transfer

1. Horizontal electrophoresis system with a large gel running chamber (gel size 25 × 20 cm) and 32-tooth comb (thickness 1.0 mm and width of teeth 4.0 mm).
2. 1× alkaline electrophoresis buffer: 50 mM NaOH, 1 mM EDTA, pH 8.5.
3. Glass plate that fits the gel.
4. 0.25 N hydrochloric acid (HCl): dilute HCl in ddH<sub>2</sub>O just before use.
5. 0.5 N NaOH, 1.5 M NaCl: dissolve in ddH<sub>2</sub>O just before use.
6. Nylon hybridization transfer membrane (GeneScreen<sup>®</sup> from Perkinelmer or equivalent).
7. 20× SSC buffer: 3 M NaCl, 300 mM sodium citrate. Adjust the pH to 7.0 with HCl. Autoclave.
8. Whatman 3 MM paper.
9. Parafilm from Bemis NA or equivalent.
10. Paper towels.
11. Neutralization solution: 0.5 M Tris-HCl pH 7.5, 1 M NaCl.
12. UV crosslinker.

### 2.3.3 Probe Labeling

1. Plasmid pML514 (available upon request), carrying part of the *MAT* locus downstream to the T7 bacteriophage promoter. Plasmid pML514 was constructed by inserting in the pGEM<sup>®</sup>-7Zf(+/-) (purchased from Promega) EcoRI site a 900-bp fragment of the *MAT* locus, obtained by PCR using yeast genomic DNA as a template and PRP643 (5'-CGG AAT TCC CTG GTT TTG GTT TTG TAG AGT GG-3') and PRP644 (5'-CGG AAT TCG AAA CAC CAA GGG AGA GAA GAC-3') as primers.
2. BamHI restriction enzyme (20,000 U/mL; New England Biolabs) and buffer supplied from distributor or equivalent (*see Note 4*).
3. In vitro transcription system Riboprobe System-T7 (Purchased from Promega and containing recombinant RNasin<sup>®</sup> RNase inhibitor, 10 mM rATP, 10 mM rCTP, 10 mM rGTP, 10 mM rUTP, 100 mM dithiothreitol (DTT), 5× transcription optimized buffer, T7 RNA polymerase, RQ1 RNase-free DNase, nuclease-free water, pGEM<sup>®</sup> Express positive control template), or equivalent.

4. rUTP- $\alpha^{32}\text{P}$ . Specific activity: 800 Ci/mmol.
5. Sephadex G-50 grade chromatography-columns (GE-Healthcare) or equivalent.

#### 2.3.4 Filter Hybridization

1. Hybridization oven and hybridization tubes.
2. Formamide hybridization buffer: 5 $\times$  SSPE, 50% formamide, 6% dextran sulfate sodium salt, 4 $\times$  Denhardt's solution (dilute 50 $\times$  Denhardt's solution from Sigma-Aldrich or equivalent), 100  $\mu\text{g}/\text{mL}$  deoxyribonucleic acid sodium salt from salmon testes, 200  $\mu\text{g}/\text{mL}$  yeast tRNA.
3. 20 $\times$  SSPE buffer: 3 M NaCl, 0.2 M  $\text{NaH}_2\text{PO}_4$ , 20 mM EDTA. Adjust the pH to 7.4 with NaOH. Autoclave.
4. 1 $\times$  SSPE, 0.1% SDS.
5. 0.1 $\times$  SSPE, 0.1% SDS.
6. 0.2 $\times$  SSPE, 0.1% SDS.
7. Autoradiography cassette with intensifying screens.
8. Autoradiography films.

---

## 3 Methods

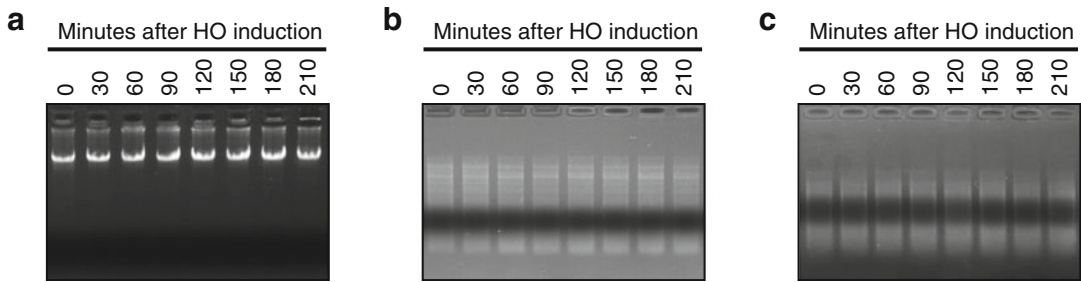
### 3.1 Double-Strand Break Induction in Yeast Cells

1. Inoculate cells in 50–100 mL YEPD medium.
2. Grow the cell culture 6–8 h at 26 °C until the exponentially growing cells reach a density of  $10^4$ – $10^6$  cells/mL.
3. Spin down the cells and wash with YEP + raffinose medium to remove glucose. Resuspend the cells in an equal or larger volume (depending on the number of samples to collect after HO induction) of YEP + raffinose medium (*see* **Notes 5** and **6**).
4. Grow the cell culture overnight at 26 °C.
5. When the cell culture has grown to a density between  $6 \times 10^6$  and  $10^7$  cells/mL draw a 50 mL sample for the uninduced control. Then add galactose from 30% solution to 3% final concentration to the remaining cell culture (*see* **Note 7**).
6. Grow the cell culture at 26 °C and take 50 mL samples at the desired time points after galactose addition. During the experiment the efficiency of DSB formation can be determined (*see* **Note 8**).

### 3.2 Genomic DNA Extraction

1. Pellet the cells by spinning 3 min at  $1600 \times g$  in 50 mL tubes.
2. Wash the cells in 1 mL spheroplasting solution and transfer the samples to 1.5 mL microcentrifuge tubes.
3. Spin 3 min at  $1600 \times g$  and completely remove the supernatant with a tip.

4. Freeze and store the cell pellets at  $-20^{\circ}\text{C}$ .
5. Thaw the samples at room temperature and resuspend the cell pellets in  $400\ \mu\text{L}$  spheroplasting solution +  $14\ \text{mM}$   $\beta$ -mercaptoethanol.
6. Add  $100\ \mu\text{L}$  Zymolyase solution to each sample and invert the tube four to six times. Incubate the samples at  $37^{\circ}\text{C}$ . After  $30\ \text{min}$  check the formation of spheroplasts under a light microscope (*see Note 9*).
7. When  $>95\%$  cells have become spheroplasts, spin  $30\ \text{s}$  at  $15,000 \times g$  and carefully remove the supernatant with a tip.
8. Gently resuspend spheroplasts in  $400\ \mu\text{L}$   $1\times$  TE (do not vortex).
9. Add  $90\ \mu\text{L}$  lysis solution (prepared just before use). Immediately shake vigorously the tube (a foam should form) and incubate the samples  $30\ \text{min}$  in a water-bath at  $65^{\circ}\text{C}$ .
10. Add  $80\ \mu\text{L}$   $5\ \text{M}$  potassium acetate and mix by inverting the tube several times. Place the tubes on ice for approximately  $1\ \text{h}$ .
11. Spin for  $20\ \text{min}$  at  $15,000 \times g$  at  $4^{\circ}\text{C}$ . Transfer the supernatant with nucleic acids to new  $1.5\ \text{mL}$  tubes. Discard the pellets.
12. Add  $1\ \text{mL}$  ice-cold  $96\%$  ethanol and mix by inverting the tube several times. A white cloudy precipitate should form. Place  $30\ \text{min}$  at  $-80^{\circ}\text{C}$  to facilitate precipitation.
13. Spin  $10\ \text{min}$  at  $15,000 \times g$  at  $4^{\circ}\text{C}$  and remove the supernatant.
14. Wash the pellet with  $1\ \text{mL}$  ice-cold  $70\%$  ethanol and immediately discard the ethanol.
15. Air-dry the pellet until it appears glassy.
16. Add  $500\ \mu\text{L}$   $1\times$  TE. Let tubes sit for  $30\ \text{min}$  at room temperature, then gently dissolve the pellet (do not vortex).
17. When all the pellets are completely dissolved and the solutions appear clear and transparent add  $2.5\ \mu\text{L}$  RNase solution to each sample and incubate for  $1\ \text{h}$  at  $37^{\circ}\text{C}$ .
18. Add  $500\ \mu\text{L}$   $2$ -propanol and invert the tube several times. A white DNA "clew" should form. Place  $30\ \text{min}$ –overnight at  $-80^{\circ}\text{C}$  to facilitate precipitation.
19. Spin for  $15$ – $30\ \text{min}$  at  $15,000 \times g$  at  $4^{\circ}\text{C}$  and remove the supernatant.
20. Wash the pellet with  $1\ \text{mL}$  ice-cold  $70\%$  ethanol and immediately discard the ethanol.
21. Air-dry until the pellet appears glassy.
22. Add  $30\ \mu\text{L}$   $1\times$  TE. Let tubes sit for  $30\ \text{min}$  at room temperature, then gently dissolve the DNA pellet (do not vortex).



**Fig. 2** Evaluation of the DNA quality during sample DNA preparation. (a–c) JKM139 cells exponentially growing in YEP + raffinose were transferred in YEP + raffinose + galactose to induce HO expression and DSB formation (*time zero*). Genomic DNA was prepared from samples taken at the indicated time points after galactose addition. (a) DNA extraction. 1  $\mu$ L from 30  $\mu$ L genomic DNA was visualized on regular agarose gel with ethidium bromide. (b) DNA digestion. Genomic DNA was digested with SspI. After 5 h at 37 °C, 2  $\mu$ L of each digestion reaction were analyzed on agarose gel with ethidium bromide. (c) DNA denaturation. SspI-digested genomic DNAs were dissolved in 18  $\mu$ L alkaline loading buffer. 1  $\mu$ L of each sample was visualized on agarose gel with ethidium bromide

23. Apply 1  $\mu$ L of each genomic DNA sample (added to 10  $\mu$ L of 1 $\times$  DNA loading buffer) on a 0.8% regular agarose gel with ethidium bromide and run in 1 $\times$  TAE buffer. Check the quality of the extracted DNA under an UV lamp (*see Note 10* and Fig. 2a).

### 3.3 Alkaline Denaturing Electrophoresis and Southern Blot

#### 3.3.1 DNA Digestion and Denaturation

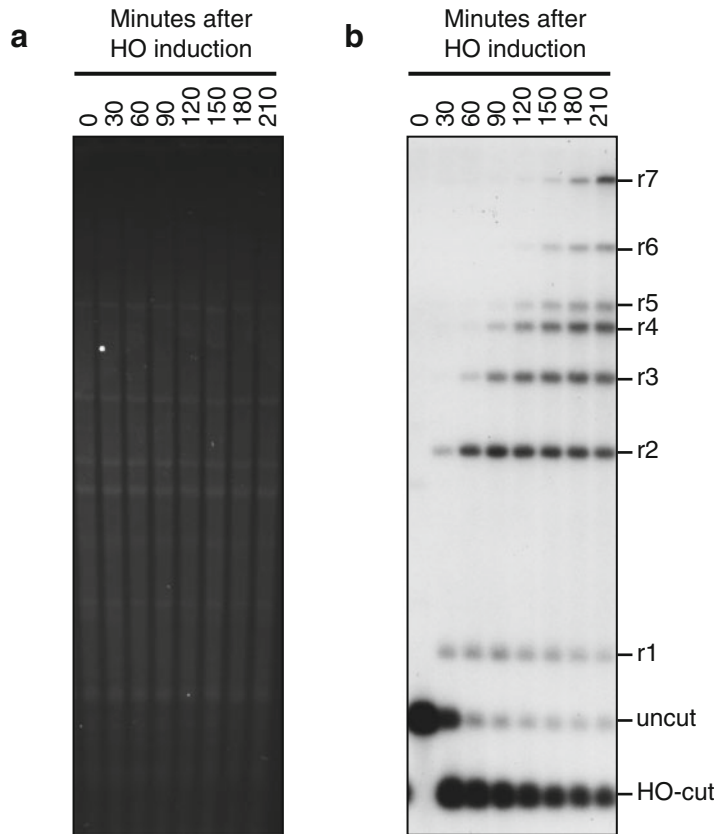
1. Digest each DNA sample (10–15  $\mu$ g DNA) with 10 U of SspI (New England Biolabs) or other restriction enzymes that cut double-stranded DNA (dsDNA) but not ssDNA (*see Notes 11* and *12*). Digest with 1 $\times$  enzyme buffer in a total volume of 70  $\mu$ L for 5–6 h at 37 °C.
2. Test 2  $\mu$ L of each digestion reaction on a 0.8% agarose gel with ethidium bromide and run in 1 $\times$  TAE buffer to check that all samples are digested (Fig. 2b).
3. Precipitate the digested DNA with 2 volumes 96% ethanol, 5 mM EDTA, pH 8.5, 0.3 M sodium acetate pH 5.2. Place overnight at –80 °C to facilitate precipitation.
4. Spin for 30 min at 15,000  $\times g$  at 4 °C and remove the supernatant.
5. Wash the pellet with 1 mL ice-cold 70% ethanol and immediately discard the ethanol.
6. Air-dry until the pellet appears glassy.
7. Add 18  $\mu$ L 1 $\times$  alkaline loading buffer. Let tubes sit for 30 min at room temperature, then gently dissolve the pellet (do not vortex). Let tubes sit for additional 1.5–2 h at room temperature by gently mixing every 15–30 min. The DNA should denature in single-stranded filaments. A 1 kb DNA ladder can

also be treated with  $1\times$  alkaline loading buffer and then used as marker for the size estimation of the DNA resection fragments.

8. Apply  $1\ \mu\text{L}$  denatured DNA to  $10\ \mu\text{L}$  alkaline loading buffer and load on a 0.8% agarose gel with ethidium bromide. Run in  $1\times$  TAE buffer to check whether the DNA is completely denatured (*see Note 13* and Fig. 2c) and to evaluate the amount of DNA in each digestion. Equilibrate the DNA concentration in the different tubes by adding an appropriate amount of  $1\times$  alkaline loading buffer.

### 3.3.2 Alkaline Denaturing Gel Electrophoresis and Transfer

1. Melt 0.8% agarose in 450 mL ddH<sub>2</sub>O, and pour into a gel tray.
2. When the gel is completely solidified, mount it in a large gel box and submerge the gel in  $1\times$  alkaline electrophoresis buffer. Allow the gel to equilibrate for 30 min or longer (*see Note 14*).
3. Load on the gel  $15\ \mu\text{L}$  of each sample dissolved and equilibrated in alkaline loading buffer.
4. Carry out electrophoresis by running the gel overnight at voltages  $<3\ \text{V}/\text{cm}$ . As bromophenol blue diffuses rapidly out of the gel into the alkaline electrophoresis buffer, place a glass plate directly on the top of the gel after the dye has migrated out of the loading slots.
5. After the DNA has migrated far enough (the dye has to migrate approximately 13–14 cm from the loading slots), remove the gel from the tank.
6. Stain the DNA with ethidium bromide by soaking the gel for 30 min–1 h in  $1\times$  TAE buffer with  $0.5\ \mu\text{g}/\text{mL}$  ethidium bromide (*see Note 15*).
7. Wash the gel 10 min in ddH<sub>2</sub>O.
8. Check the gel under an UV lamp (Fig. 3a).
9. Soak the gel 7 min with gentle agitation in 0.25 N HCl.
10. Rinse the gel with ddH<sub>2</sub>O.
11. Soak the gel 30 min with gentle agitation in 0.5 N NaOH, 1.5 M NaCl.
12. Rinse the gel with ddH<sub>2</sub>O.
13. Blot overnight the DNA from the gel to a nylon hybridization membrane by capillary transfer with  $10\times$  SSC buffer (*see Notes 16* and *17*). Fill a tray with  $10\times$  SSC buffer (approximately 1.5 L). Place a platform on the tray and create a bridge with a Whatman 3 MM paper onto the platform. Pour  $10\times$  SSC buffer over the bridge and remove air bubbles. Place the gel on the Whatman 3 MM paper bridge. Wet a membrane that fits with the size of the gel in  $10\times$  SSC and lay on the gel. Remove air bubbles. Wet three sheets of Whatman 3 MM paper in  $10\times$  SSC and lay on the membrane. To maintain capillary flow



**Fig. 3** DSB end resection at *MAT* locus. **(a)** Alkaline denaturing gel electrophoresis. *SspI*-cut genomic DNA was run overnight at 1.4 V/cm on an alkaline denaturing agarose gel and stained 30 min in  $1\times$  TAE buffer with 0.5  $\mu\text{g}/\text{mL}$  ethidium bromide. **(b)** Southern blot analysis. ssDNA fragments were transferred on a nylon membrane and hybridized with a *MAT* RNA probe. Before galactose addition the probe reveals the uncut band, which is converted in the smaller HO-cut band after DSB formation. The resection bands r1–r7 appear when resection proceeds and eliminates *SspI* sites, as depicted in Fig. 1

through the gel instead of around it, place strips of Parafilm on all the sides of the gel and create a barrier between the bridge and the paper towels. Build up a big stack of paper towels and add some weight on the top.

14. After transfer, soak the membrane 15 min in neutralization solution with gentle agitation.
15. Air-dry the filter.
16. Crosslink the DNA on the membrane with an UV crosslinker by following the instructions of the manufacturer.

### 3.3.3 Probe Labeling

1. Digest 5–10  $\mu\text{g}$  pML514 DNA with *Bam*HI at 37  $^{\circ}\text{C}$  to linearize the plasmid downstream the *MAT* insert.



2. Check whether the plasmid is completely linearized by loading 2  $\mu\text{L}$  of the digestion reaction and 1  $\mu\text{L}$  of the uncut plasmid on a 0.8% agarose gel with ethidium bromide (*see Note 18*).
3. Precipitate the linearized plasmid and suspend it in a small volume of nuclease-free water (*see Note 19*).
4. Prepare the labeled RNA probe with the Promega Riboprobe System-T7 kit or equivalent, according to the instructions of the manufacturer. If the Promega Riboprobe System-T7 is chosen, add the following components in the order listed. 4  $\mu\text{L}$  5 $\times$  transcription optimized buffer; 2  $\mu\text{L}$  100 mM DTT, 20 U recombinant RNasin<sup>®</sup> RNase inhibitor, 1  $\mu\text{L}$  10 mM rATP, 1  $\mu\text{L}$  10 mM rGTP, 1  $\mu\text{L}$  10 mM rCTP, 1–4  $\mu\text{L}$  linearized template DNA (0.2–1.0  $\mu\text{g}/\mu\text{L}$  in nuclease-free water), nuclease-free water to a 20  $\mu\text{L}$  final volume, 5  $\mu\text{L}$  rUTP- $\alpha^{32}\text{P}$  (100  $\mu\text{Ci}$ ), 15–20 U T7 RNA Polymerase.
5. Incubate for 45 min at 40 °C.
6. Add 1 U RQ1 RNase-free DNase.
7. Incubate for 15 min at 37 °C.
8. Add 180  $\mu\text{L}$  nuclease-free water.
9. Equilibrate a Sephadex G-50 grade chromatography-column with nuclease-free water.
10. Load the in vitro transcription solution on the chromatography-column.
11. Spin for 1 min at  $850 \times g$  and recover the eluate.

### 3.3.4 Filter Hybridization

1. Insert the filter in a hybridization tube and soak the filter with ddH<sub>2</sub>O.
2. Block the filter by incubating for 5 h in a hybridization oven at 42 °C in 25 mL formamide hybridization buffer (pre-hybridization).
3. Prepare the hybridization solution by adding the RNA probe obtained by in vitro transcription (almost 200  $\mu\text{L}$ ; *see Subheading 3.3.3*) to 25 mL of fresh formamide hybridization buffer. Replace the pre-hybridization solution with this hybridization solution.
4. Incubate overnight at 42 °C by gently rotating in the hybridization oven.
5. Remove the hybridization solution.
6. Wash the filter for 30 min at 42 °C in 5 $\times$  SSPE.
7. Wash for 30 min at 42 °C in 1 $\times$  SSPE, 0.1% SDS.
8. Wash for 30 min at 42 °C in 0.1 $\times$  SSPE, 0.1% SDS.
9. Wash for 15 min at 68 °C in 0.2 $\times$  SSPE, 0.1% SDS.

10. Wash for 5 min at room temperature in  $0.2\times$  SSPE.
11. Air-dry the filter.
12. Expose the filter to an autoradiography film in an autoradiography cassette with intensifying screens (*see* **Notes 20** and **21** and Fig. **3b**).

---

## 4 Notes

1. Galactose should not be autoclaved because it isomerizes at elevated temperatures.
2. There are two preparations of Zymolyase from Nacai Tesque, Zymolyase<sup>®</sup>-20T, and Zymolyase<sup>®</sup>-100T, having lytic activity of 20 U/g and 100 U/g, respectively. Zymolyase-100T can be used in place of Zymolyase-20T but the Zymolyase concentration in the spheroplasting solution and/or the time of incubation should be adjusted according to the different activity of the enzyme.
3. Alternatives to ethidium bromide are less-toxic dyes such as GelRed<sup>™</sup> or GelGreen<sup>™</sup>.
4. Templates used for *in vitro* transcription should either be blunt-ended or carry protruding 5' termini. As extraneous transcripts can appear in addition to the expected transcript when templates contain 3' overhangs [11], plasmids should not be linearized with any enzyme that leaves a 3' overhang.
5. Be sure to transfer in YEP + raffinose medium enough cells to reach a density of at least  $6 \times 10^6$  cells/mL the next day.
6. HO expression and DSB formation can also be induced in a strain carrying a plasmid. In this case, an appropriate selective medium should be used in place of YEPD to maintain the plasmid selection. However, cells grow poorly in selective media with raffinose as a carbon source. Thus, cells can be grown in glucose selective medium for approximately 8 h and then transferred in YEP + raffinose medium for the overnight growth to minimize the number of cell divisions in a rich medium.
7. As resection is stimulated by the activity of the cyclin-dependent kinases (Cdk1 in *S. cerevisiae*) and the kinetics of resection change in the different cell cycle phases [12, 13], sometimes it might be important to analyze the generation of ssDNA at DSB ends in cells arrested at a particular stage of the cell cycle. To monitor resection in G2/M phase, when Cdk1 activity is high, cells can be arrested in early metaphase before galactose addition by treatment with nocodazole, a drug that interferes with microtubule polymerization. Prepare a  $100\times$

nocodazole stock by dissolving 1.5 mg/mL nocodazole in 100% dimethyl sulfoxide (DMSO). Add the nocodazole stock to the exponentially growing cells in YEP + raffinose medium to a final concentration of  $1\times$  (final concentration of nocodazole 15  $\mu\text{g}/\text{mL}$ ). After 2.5–3 h at 26 °C, visualize cells under a light microscope. Add galactose to induce HO in G2/M-arrested cells when 90–100% cells accumulate as large-budded cells. Nocodazole-mediated arrest can be maintained for 8–10 h [7]. To monitor resection in G1 phase, when Cdk1 activity is low, *MAT $\alpha$*  haploid cells can be treated with the  $\alpha$ -factor pheromone. As these cells recover from the  $\alpha$ -factor block after 1.5–2 h due to  $\alpha$ -factor degradation by the Bar1 protease, a persistent G1 arrest with  $\alpha$ -factor can be induced in strains carrying the deletion of the *BAR1* gene [14]. Dissolve 1 mg/mL  $\alpha$ -factor in ddH<sub>2</sub>O and add this solution to the exponentially growing cells in YEP + raffinose to a final  $\alpha$ -factor concentration of 0.5  $\mu\text{g}/\text{mL}$ . After 2 h at 26 °C, visualize cells under a light microscope. Add galactose to induce HO in G1-arrested cells when 90–100% cells are unbudded.  $\alpha$ -factor-mediated arrest in *bar1 $\Delta$*  cells can be maintained for at least 8 h [15].

8. As a single DSB formation triggers the activation of the DNA damage checkpoint and a G2/M arrest [16], the effectiveness of DSB induction in checkpoint-proficient strains can be determined under a light microscope by evaluating the percentage of large-budded cells after 4–6 h of galactose induction. Alternatively, samples taken before HO induction and at the desired time points after galactose addition (2 and/or 4 h) can be diluted and 300–3000 cells can be plated on YEPD plates to test if galactose addition has efficiently induced the DSB formation. As in this strain the HO-induced DSB is irreparable and its generation leads to cell death [6], strain viability after HO induction can be used to determine the efficiency of DSB formation.
9. Spheroplasts display a rounded morphology in contrast to the ovoid shape of intact yeast cells. Furthermore, as spheroplasts lyse in SDS solution, cells should disappear after the addition of one drop of 10% SDS to 3  $\mu\text{L}$  of cells on a glass slide. We usually obtain >95% spheroplasts in 40–50 min.
10. If high molecular weight DNA molecules were extracted and completely dissolved, a single band should be detectable in the gel without smears or signals retained in the well.
11. Restriction endonucleases digest dsDNA by cleaving two phosphodiester bonds, one within each strand of the duplex DNA. ssDNA generally cannot be cut by restriction enzymes. A few restriction enzymes cleave ssDNA, although usually at low efficiency.

12. The conditions described here were set up for digestion with SspI restriction enzyme purchased from New England Biolabs. If a different enzyme and/or producer were chosen, consult the manufacturer's instructions for optimal digestion conditions.
13. When DNA is completely denatured, a smear is detectable on the agarose gel.
14. The agarose gel is equilibrated in alkaline electrophoresis buffer after solidification because the addition of NaOH to a warm agarose solution causes polysaccharide hydrolysis. Alternatively, the agarose can be melted in ddH<sub>2</sub>O and then cooled down to 60 °C, so that NaOH to 50 mM and EDTA pH 8.5 to 1 mM can be added just before pouring the gel.
15. Ethidium bromide is omitted from alkaline agarose gels because it does not bind to DNA at high pH. DNA can be stained with ethidium bromide after the electrophoresis. However, DNA will be faint because the ethidium bromide does not bind very well to ssDNA.
16. Alternatively, a vacuum blotter can be used.
17. DNA can be blotted onto either neutral or positively charged nylon membranes.
18. Plasmid DNA must be cleaved to completion, as trace amounts of uncut supercoiled plasmid DNA can give rise to long transcripts that include vector sequences. These transcripts may incorporate a fraction of the radiolabeled rNTP.
19. The linearized plasmid DNA should be highly concentrated because a small volume of template is required for the in vitro transcription reaction.
20. The film can be also developed in a Typhoon instrument or equivalent.
21. With freshly labeled rUTP- $\alpha^{32}\text{P}$ , we usually obtain a good signal after 4 h-overnight exposure at  $-80\text{ }^{\circ}\text{C}$ .

---

## Acknowledgments

We thank J. Haber (Brandeis University) for the JKM139 yeast strain, and M. P. Longhese and G. Lucchini for critical reading of the manuscript.

## References

1. Symington LS, Rothstein R, Lisby M (2014) Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* 198:795–835. doi:[10.1534/genetics.114.166140](https://doi.org/10.1534/genetics.114.166140)
2. Mehta A, Haber JE (2014) Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol* 6(9):a016428. doi:[10.1101/cshperspect.a016428](https://doi.org/10.1101/cshperspect.a016428)

3. Symington LS (2014) End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol* 6(8):a016436. doi:[10.1101/cshperspect.a016436](https://doi.org/10.1101/cshperspect.a016436)
4. Villa M, Cassani C, Gobbin E, Bonetti D, Longhese MP (2016) Coupling end resection with the checkpoint response at DNA double-strand breaks. *Cell Mol Life Sci* 73(19):3655–3663. (in press)
5. White CI, Haber JE (1990) Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *EMBO J* 9:663–673
6. Lee SE, Moore JK, Holmes A, Umez K, Kolodner RD, Haber JE (1998) *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* 94:399–409. doi:[10.1016/s0092-8674\(00\)81482-8](https://doi.org/10.1016/s0092-8674(00)81482-8)
7. Clerici M, Mantiero D, Lucchini G, Longhese MP (2005) The *Saccharomyces cerevisiae* Sae2 protein promotes resection and bridging of double strand break ends. *J Biol Chem* 280:38631–38638
8. Manfrini N, Guerini I, Citterio A, Lucchini G, Longhese MP (2010) Processing of meiotic DNA double strand breaks requires cyclin-dependent kinase and multiple nucleases. *J Biol Chem* 285:11628–11637. doi:[10.1074/jbc.M110.104083](https://doi.org/10.1074/jbc.M110.104083)
9. Bonetti D, Martina M, Clerici M, Lucchini G, Longhese MP (2009) Multiple pathways regulate 3' overhang generation at *S. cerevisiae* telomeres. *Mol Cell* 35:70–81. doi:[10.1016/j.molcel.2009.05.015](https://doi.org/10.1016/j.molcel.2009.05.015)
10. Lee CS, Haber JE (2015) Mating-type gene switching in *Saccharomyces cerevisiae*. *Microbiol Spectr* 3(2):MDNA3-0013-2014. doi:[10.1128/microbiolspec.MDNA3-0013-2014](https://doi.org/10.1128/microbiolspec.MDNA3-0013-2014)
11. Schenborn ET, Mierendorf RC (1985) A novel transcription property of SP6 and T7 RNA polymerases: dependence on template structure. *Nucleic Acids Res* 13:6223–6236
12. Aylon Y, Liefshitz B, Kupiec M (2004) The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. *EMBO J* 23:4868–4875
13. Ira G, Pelliccioli A, Balijja A, Wang X, Fiorani S, Carotenuto W, Liberi G, Bressan D, Wan L, Hollingsworth NM, Haber JE, Foiani M (2004) DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* 431:1011–1017
14. Chan RK, Otte CA (1982) Physiological characterization of *Saccharomyces cerevisiae* mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Mol Cell Biol* 2:21–29
15. Trovesi C, Falcattoni M, Lucchini G, Clerici M, Longhese MP (2011) Distinct Cdk1 requirements during single-strand annealing, noncrossover, and crossover recombination. *PLoS Genet* 7(8):e1002263. doi:[10.1371/journal.pgen.1002263](https://doi.org/10.1371/journal.pgen.1002263)
16. Pelliccioli A, Lee SE, Lucca C, Foiani M, Haber JE (2001) Regulation of *Saccharomyces* Rad53 checkpoint kinase during adaptation from DNA damage-induced G2/M arrest. *Mol Cell* 7:293–300

# Chapter 12

## Single Molecule Analysis of Resection Tracks

Pablo Huertas and Andrés Cruz-García

### Abstract

Homologous recombination is initiated by the so-called DNA end resection, the 5′–3′ nucleolytic degradation of a single strand of the DNA at each side of the break. The presence of resected DNA is an obligatory step for homologous recombination. Moreover, the amount of resected DNA modulates the prevalence of different recombination pathways. In different model organisms, there are several published ways to visualize and measure with more or less detail the amount of DNA resected. In human cells, however, technical constraints hampered the study of resection at high resolution. Some information might be gathered from the study of endonuclease-created DSBs, in which the resection of breaks at known sites can be followed by PCR or ChIP. In this chapter, we describe in detail a novel assay to study DNA end resection in breaks located on unknown positions. Here, we use ionizing radiation to induce double-strand breaks, but the same approach can be used to monitor resection induced by different DNA damaging agents. By modifying the DNA-combing technique, used for high-resolution replication analyses, we can measure resection progression at the level of individual DNA fibers. Thus, we named the method Single Molecule Analysis of Resection Tracks (SMART). We use human cells in culture as a model system, but in principle the same approach would be feasible to any model organism adjusting accordingly the DNA isolation part of the protocol.

**Key words** DNA resection, High-resolution resection assay, SMART, DNA combing, Fiber assay

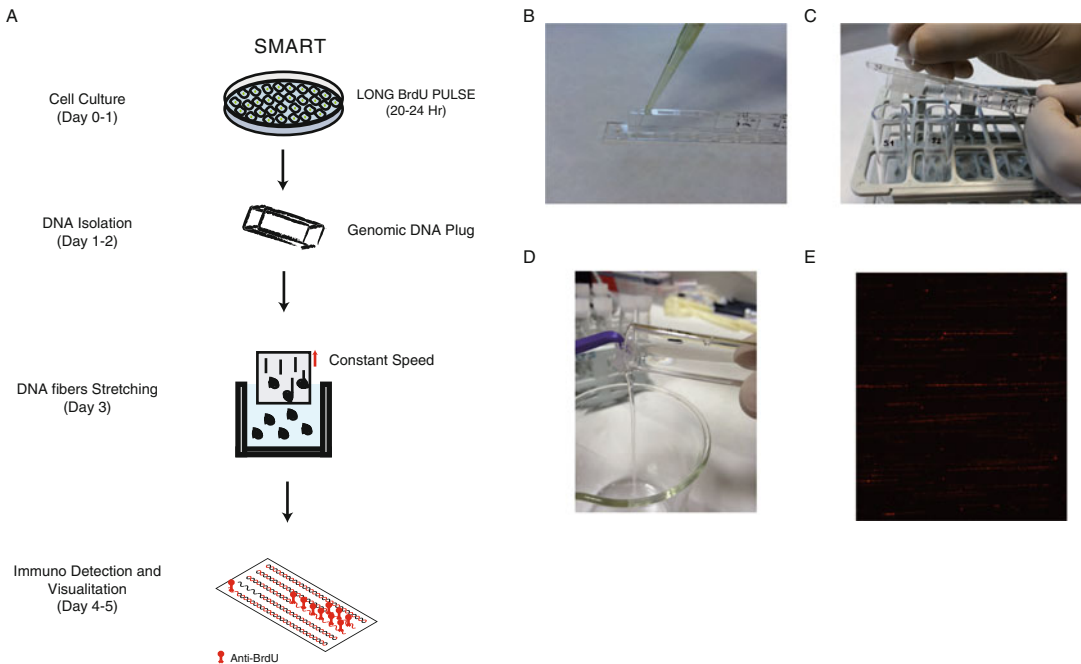
---

### 1 Introduction

The key event that controls the choice between different pathways to repair DNA Double-Strand Break (DSB) is DNA-end resection. This mechanism consists of a 5′-to-3′ degradation of one strand at each side of the break, yielding a long stretch of 3′-ended single-stranded DNA (ssDNA). Whereas non-homologous end-joining (NHEJ) is inhibited by DNA-end resection, the DNA-end once resected is committed to being repaired by either homologous recombination (HR) or microhomology-mediated end-joining (MMEJ) [1]. Once the break end is resected, the ssDNA tail that is generated is rapidly coated by RPA (replication protein A) [1], a heterotrimeric protein that prevents the formation of secondary structures and protects against degradation of the ssDNA [1].

DNA end resection has been studied by different methods depending on if the location of the DSB was known. For those DSB created randomly by DNA damaging agents such as ionizing radiation or several drugs, resection has been measured by the exposure of ssDNA-containing BrdU [2], by proxy using RPA accumulation (either by immunofluorescence or FACS; [2, 3]), by posttranslational modification of RPA [2] or modification of the apparent size of a circular chromosome [4]. Such studies can monitor large changes in resection, mainly in the number of breaks that are resected, but they all share a common problem: they are low-resolution techniques that cannot reflect changes on the length or speed of DNA resection. An alternative is to use breaks created at established position, like the ones created with nucleases such as HO, I-SceI, or AsiSI [4–10]. In these cases, high-resolution techniques based on chromatin IP, quantitative PCR, Southern blots, or sequencing are available [4–10]. However, with those techniques the study is restricted to one or several fixed loci.

Here, we describe a simple immunofluorescence-based technique to monitor DNA end resection of breaks created by any type of DNA-damaging agents. A schematic of the protocol can be seen in Fig. 1a. Briefly, the protocol relies on a long treatment with BrdU (one cell cycle) to label one strand of all DNA molecules,



**Fig. 1** Single Analysis of resection Tracks. (a) Schematic representation of the SMART protocol. (b) DNA plug formation with a casting mold. (c) Picture showing how to eject the DNA plugs onto the tubes. (d) To avoid losing the plugs during the washes, use a cell scraper to hold it into the tube. (e) A representative image of a SMART coverslip. Red signal corresponds to ssDNA visualized by BrdU immunofluorescence

followed by the isolation of DNA in gentle conditions using cells embedded in agarose plugs. Then, genomic DNA is stretched onto silanized coverslips to create parallel individual DNA fibers. Finally, the ssDNA is detected using an anti-BrdU antibody, taking advantage of the fact that the BrdU epitope is hidden on double-stranded but exposed in single-stranded DNA. The resected DNA can be visualized by fluorescence microscopy as fibers of BrdU/ssDNA strands whose length can serve as a metric to quantify resection in cells responding to DNA damage. Although we developed the assay to study DNA resection in mammalian cells, the same approach might be applied successfully to other model organisms by adjusting the BrdU labeling/DNA isolation part of the method using established DNA combing protocols.

---

## 2 Materials

### 2.1 Equipment

1. Agitator plate orbital.
2. Fluorescent microscope (with FITC and TX2 filters and 40× objective).
3. Hemocytometer cell counter or automated cell counter (i.e., Beckman Coulter).
4. Oven (42 °C and 65 °C).
5. Fiber Comb<sup>®</sup> Molecular Combing System. (Genomic Vision).

#### 2.1.1 Materials

1. Cell culture plates.
2. Coplin jar.
3. Cover slip (not silanized) (22 × 22 mm).
4. Cover slip silanized (22 × 22 mm) (GenomicVision).
5. Filter lab paper.
6. Gel plug mold (CHEF MAPPER PLUG MOLD, BIORAD).
7. Glass slide.
8. Masking tape.
9. Microtubes.
10. Pipettes.
11. Test tubes 12 ml round bottom.
12. Tips (white, yellow).

#### 2.1.2 Buffers

1. Blocking Buffer: 1% BSA in PBS/T.
2. MES 10×: 350 mM MES hydrate, 150 mM MES sodium salt.
3. PBS 10×: 81.9 g NaCl, 2.01 g KCl, 14.2 g Na<sub>2</sub>HPO<sub>4</sub>, 2.45 KH<sub>2</sub>PO<sub>4</sub> g per liter.
4. PBS/T: PBS 1×, 0.1% Triton X-100.



5. TE (1×): 10 mM Tris, 1 mM EDTA, pH 8.0.
6. TE50: 10 mM Tris, 50 mM EDTA, pH 8.0.

### 2.1.3 Stocks

1. 10% (W/V) N-lauryl-sarcosyl.
2. 10 mM BrdU: 100 mg of BrdU in 32.56 ml of distilled water. Aliquot in microtubes and store at  $-20\text{ }^{\circ}\text{C}$ .
3. 500 mM MES Hydrate: 9.762 g in 100 ml of distilled water.
4. 500 mM MES sodium salt: 10.861 g in 100 ml of distilled water.
5. Agarose 1% in PBS: 250 mg in 250 ml of PBS 1×. Aliquot in microtubes and store at  $-20\text{ }^{\circ}\text{C}$ .
6. Proteinase K 2 mg/ml: 2.5 g in 100 ml of TE50 buffer, 1% l-sarcosyl. Aliquot in microtubes and store at  $-20\text{ }^{\circ}\text{C}$ .
7. Yoyo-1 (Yoyo-1 iodide): 1 mM solution in DMSO (Molecular Probes).

---

## 3 Methods

A schematic of the protocol can be found in Fig. 1a.

### 3.1 Culture Growth (Day 0)

1. Seed in 100 mm plates,  $5 \times 10^5$  cells in 10 ml of proper cell culture medium (*see* **Notes 1** and **2**).
2. Incubate cells at  $37\text{ }^{\circ}\text{C}$  for 20–24 h.

### 3.2 Labeling of Mammalian Cells with BrdU (Day 1)

1. Remove old medium from cell culture. Add fresh medium supplemented with BrdU,  $10\text{ }\mu\text{M}$  final concentration.
2. Incubate cells at  $37\text{ }^{\circ}\text{C}$  for 20–24 h (one cell cycle).

### 3.3 Preparation of Genomic DNA Plugs (Day 2)

1. Heat agarose solution at  $65\text{ }^{\circ}\text{C}$  in a thermoblock for at least 15 min, until it is completely melted. Lower the temperature to  $42\text{ }^{\circ}\text{C}$  and let the agarose to temper.
2. Treat cells to generate the desired amount and type of DNA damage (*see* **Note 3**).
3. Incubate the cells for 1 h to allow resection to take place (*see* **Note 4**).
4. Rinse cells with cold PBS 1× ( $4\text{ }^{\circ}\text{C}$ ), and harvest cells with Accutase (eBiosciences) for 5 min (*see* **Notes 5** and **6**).
5. Resuspend cells gently in 5 ml of PBS 1×, spin 3 min at  $400 \times g$ . During centrifugation, label microtubes for each sample (*see* **Note 7**).
6. Aspirate the PBS and resuspend in 5 ml of ice-cold PBS 1×.
7. Count the number of cells and resuspend at a final concentration of 8,000,000 cells per 1 ml of PBS 1×.
8. Transfer cells solution (100  $\mu\text{l}$ ) to a pre-labeled microtube.

9. Briefly pre-warm cell suspension to 42 °C in a thermoblock.
10. To prepare 2 plugs add 100 µl of 1% LMT agarose (Bio-Rad) in PBS 1× to 100 µl of cell solution.
11. Mix very gently with a P200 pipetman (cut tip). This has to be done quickly; the melted agarose can become solid when the temperature is lower than 42 °C.
12. Transfer 100 µl of the agarose/cells mix per plug into a casting mold (Fig. 1b) (*see Note 8*).
13. Let plugs solidify for 25 min at room temperature and additionally for 5 min at 4 °C. While the solidification occurs, label 12 ml round-bottom tubes for each sample.
14. Eject plugs into 12 ml round-bottom tubes containing 500 µl of proteinase K in TE50 (Fig. 1c). Plugs for the same condition can be pooled in a single tube, scaling up the amount of buffer. Make sure that DNA agarose plugs are completely covered by the buffer solution (*see Note 9*).
15. Incubate overnight at 50 °C.

### 3.4 Melting of Genomic DNA Plug and YOYO-1 Staining (Day 3)

1. Gently, remove the liquid by blocking the agarose plug with a cell scraper (Fig. 1d).
2. Add 500 µl of proteinase K in TE<sub>50</sub> buffer per plug.
3. Incubate at 50 °C for at least another 6 h (*see Note 10*).
4. Gently remove buffer without damaging plugs, which should be now completely transparent.
5. Wash 4 × 10 min with 10 ml of TE<sub>50</sub> at RT with gently shaking (i.e., 300 rpm in Eppendorf thermomixer).  
**Optional:** Samples can be stored at this point at 4 °C in TE<sub>50</sub> until use.
6. Put each plug to be stained with YOYO-1 in a new 12 ml round-bottom tube (*see Note 11*).  
**Optional:** To keep long-term any extra plugs, add 10 ml of TE<sub>50</sub> to the tube and keep it at 4 °C.
7. Stain the plug with 100 µl of TE and add 1.5 µl of YOYO-1 solution (1 mM in DMSO). Incubate in the dark for 30 min at RT.
8. Wash the plugs with 4 × 10 ml of TE for 5 min with gently shaking.
9. Add 3 ml of MES 1× to each tube.
10. Incubate at 65 °C until the plugs are completely melted (15–20 min).
11. Let the temperature drop gradually to 42 °C. Add 3 units of beta Agarase (NEB) by adding 100 µl of MES 1× with 3 µl of β-Agarase per plug *see Note 12*.
12. Incubate at 42 °C overnight.

### 3.5 DNA Fibers Stretching (Day 4)

1. Incubate at 65 °C for 15 min to inactivate the Beta Agarase. Then, cool the sample until it reaches room temperature.
2. Perform the DNA fiber stretching in silanized coverslip using the Fiber Comb<sup>®</sup> Molecular Combing System (*see Note 13*).
3. Remove the silanized coverslips. Do not forget to mark the fiber orientation.
4. Put the coverslips on filter lab paper.
5. Bake the coverslips at 60 °C for at least 2 h.
6. Glue the coverslip to a microscope slide. Check the quality of the fibers by visualizing the YOYO-1 staining under the microscope. Only if good quality fibers are observed continue with the immunodetection.

### 3.6 Immunodetection

1. Put the slides in a coplin jar. Incubate for 15 min in a blocking solution (PBS/T, 1% BSA). Be sure that the coverslips are completely submerged in the buffer. Alternatively, commercial blocking reagents can be used.
2. Drain the slides and set them horizontally in a wet chamber. Add 18 µl of PBS/T containing anti BrdU (1:500) antibody to the coverslips and cover them with another coverslip to spread the solution. Incubate for 45 min at 37 °C (or at RT) in the wet chamber.
3. Remove the top coverslip. Wash five times for 2 min with PBS/T.
4. Incubate with secondary antibody (Dilution, 1:1000) in PBS/T for 30 min at 37 °C (or RT) in the wet chamber as described for the primary antibody.  
**Optional:** Spin fresh aliquots of secondary antibodies for 5 min at full speed with a microfuge to eliminate aggregates.
5. Wash for 5 × 2 min with PBS/T.
6. Dry slides and mount with 20 µl of Prolong Gold Antifade (Molecular Probes) using a micropipette with a cut tip. Let mounting reagent polymerize overnight at RT before proceeding with microscopy. Mounted coverslips are stable for months at −20 °C in the dark.

### 3.7 Visualization

1. Image the coverslips using a fluorescent microscope. We use 40× objectives for visualization (Fig. 1e).
2. Measure the length of resected DNA using the ruler tool in the Analysis menu of Photoshop Adobe Acrobat software. Other software with measurement tools can be used, i.e., MetaMorph.
3. Represent the dispersion of the resected DNA length and calculate the median.

4. Individual dispersion of different conditions can be compared using a Mann–Whitney test or the average of the median of three experiments for each condition can be plotted and compared using a *t*-test.

---

## 4 Notes

1. This protocol has been optimized for U2OS and HeLa cells. Other cell types might require a different seeding concentration or different growth conditions.
2. Generally for each plug 1 million cells are needed. The size of your cell culture might be scaled up depending on how many plugs will be required for the specific experiment. We recommend preparing extra plugs, especially the first few times, just in case.
3. In this specific example, we are using 10 Gy of Ionizing Radiation. Treatments with NCS, camptothecin, and etoposide have been used in the lab with good results. Note that the dose will depend on your expected results. We recommend avoiding low doses, as otherwise the amount of resected fiber might be too scarce to find a significant number.
4. The incubation time might be adjusted for the desired application. In our experience, 1 h is the best option to have tracks long enough to be measured confidently. For kinetic studies, several time points should be taken.
5. Warm up the Accutase at RT before use, do not add it at 37 °C. We usually use 700  $\mu$ l of Accutase and 4–5 min at 37 °C for 100 mm plate. The amount of Accutase and time should be adjusted for specific cell types.
6. To harvest cells trypsin can be used instead of Accutase, but the yield of DNA fibers is clearly lower.
7. In order to minimize the damage by manipulation of the cells, a P1000 can be used to resuspend the cells, adding PBS gently 1 ml at a time.
8. This step has to be done sample by sample to avoid solidification of the agarose.
9. If the ejected plug is deformed, perhaps the plug was not yet completely solidified. Put the mold back to the refrigerator and wait for another 15 min.
10. The order of proteinase K treatment can be reverted without impacting the final results. Thus, it is possible to start with a 6 h incubation at 50 °C followed by a change of buffer and an additional incubation of the samples overnight at 50 °C.

11. The YOYO-1 staining step is used to check the quality of the fibers before immunostaining. Therefore, it could be skipped, as the staining will anyway disappear during immunofluorescence.
12. Do not add the agarase before the sample is at 42 °C to avoid denaturalization of the enzyme.
13. Manual stretching can be achieved by letting a drop of sample to slide over a coverslip, but with low reproducibility between samples.

## References

1. Huertas P (2010) DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol* 17(1):11–16. doi:[10.1038/nsmb.1710](https://doi.org/10.1038/nsmb.1710). nsmb.1710 [pii]
2. Sartori AA, Lukas C, Coates J, Mistrik M, Fu S, Bartek J, Baer R, Lukas J, Jackson SP (2007) Human CtIP promotes DNA end resection. *Nature* 450(7169):509–514. doi:[10.1038/nature06337](https://doi.org/10.1038/nature06337). nature06337 [pii]
3. Forment JV, Jackson SP (2015) A flow cytometry-based method to simplify the analysis and quantification of protein association to chromatin in mammalian cells. *Nat Protoc* 10(9):1297–1307. doi:[10.1038/nprot.2015.066](https://doi.org/10.1038/nprot.2015.066)
4. Ma W, Westmoreland J, Nakai W, Malkova A, Resnick MA (2011) Characterizing resection at random and unique chromosome double-strand breaks and telomere ends. *Methods Mol Biol* 745:15–31. doi:[10.1007/978-1-61779-129-1\\_2](https://doi.org/10.1007/978-1-61779-129-1_2)
5. Bothmer A, Robbiani DF, Feldhahn N, Gazumyan A, Nussenzweig A, Nussenzweig MC (2010) 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J Exp Med* 207(4):855–865. doi:[10.1084/jem.20100244](https://doi.org/10.1084/jem.20100244)
6. Shroff R, Arbel-Eden A, Pilch D, Ira G, Bonner WM, Petrini JH, Haber JE, Lichten M (2004) Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. *Curr Biol* 14(19):1703–1711. doi:[10.1016/j.cub.2004.09.047](https://doi.org/10.1016/j.cub.2004.09.047)
7. Sugawara N, Haber JE (2006) Repair of DNA double strand breaks: in vivo biochemistry. *Methods Enzymol* 408:416–429. doi:[10.1016/S0076-6879\(06\)08026-8](https://doi.org/10.1016/S0076-6879(06)08026-8)
8. Westmoreland JW, Resnick MA (2013) Coincident resection at both ends of random, gamma-induced double-strand breaks requires MRX (MRN), Sae2 (Ctp1), and Mre11-nuclease. *PLoS Genet* 9(3):e1003420. doi:[10.1371/journal.pgen.1003420](https://doi.org/10.1371/journal.pgen.1003420)
9. Zhou Y, Paull TT (2015) Direct measurement of single-stranded DNA intermediates in mammalian cells by quantitative polymerase chain reaction. *Anal Biochem* 479:48–50. doi:[10.1016/j.ab.2015.03.025](https://doi.org/10.1016/j.ab.2015.03.025)
10. Zierhut C, Diffley JF (2008) Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. *EMBO J* 27(13):1875–1885. doi:[10.1038/emboj.2008.111](https://doi.org/10.1038/emboj.2008.111)

# Chapter 13

## Mapping DNA Breaks by Next-Generation Sequencing

Laura Baranello, Fedor Kouzine, Damian Wojtowicz, Kairong Cui, Keji Zhao, Teresa M. Przytycka, Giovanni Capranico, and David Levens

### Abstract

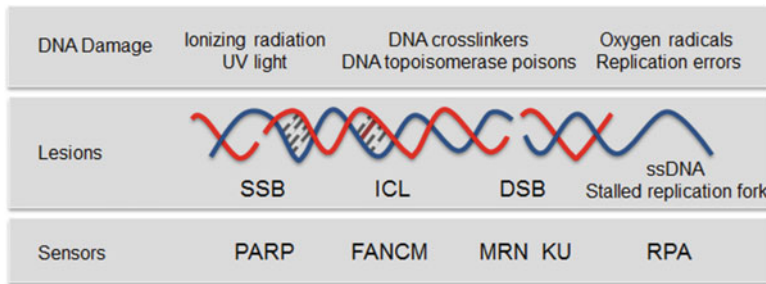
Here, we present two approaches to map DNA double-strand breaks (DSBs) and single-strand breaks (SSBs) in the genome of human cells. We named these methods respectively DSB-Seq and SSB-Seq. We tested the DSB and SSB-Seq in HCT1116, human colon cancer cells, and validated the results using the topoisomerase 2 (Top2)-poisoning agent etoposide (ETO). These methods are powerful tools for the direct detection of the physiological and pathological “breakome” of the DNA in human cells.

**Key words** DNA damage, Double-strand breaks (DSBs), Single-strand breaks (SSBs), Topoisomerase 2 (Top2), Etoposide (ETO)

---

### 1 Introduction

The cellular genome is constantly exposed to extracellular and intracellular DNA damaging agents that compromise DNA integrity and threaten genomic stability (Fig. 1). As a result, cellular mechanisms to cope with DNA damage developed early in evolution and are conserved in current animal and plant species. Understanding the sensitivity of the genome to various DNA insults is instrumental to implement effective preventive and treatment strategies of diseases. Most of the techniques developed to study DNA breaks are based on their indirect detection, by examining the localization of proteins involved in the repair of DNA damage such as the phosphorylated histone variant  $\gamma$ -H2AX [1] or the replication protein A (RPA), or by detecting the single-stranded DNA that transiently accumulates at DSB sites [2]. However, despite past work having provided comprehensive information on the pathways involved in detecting and repairing DNA breakage, our knowledge of how the genome breaks in response to various agents is incomplete and would be advanced by improved technology to detect DNA breaks.



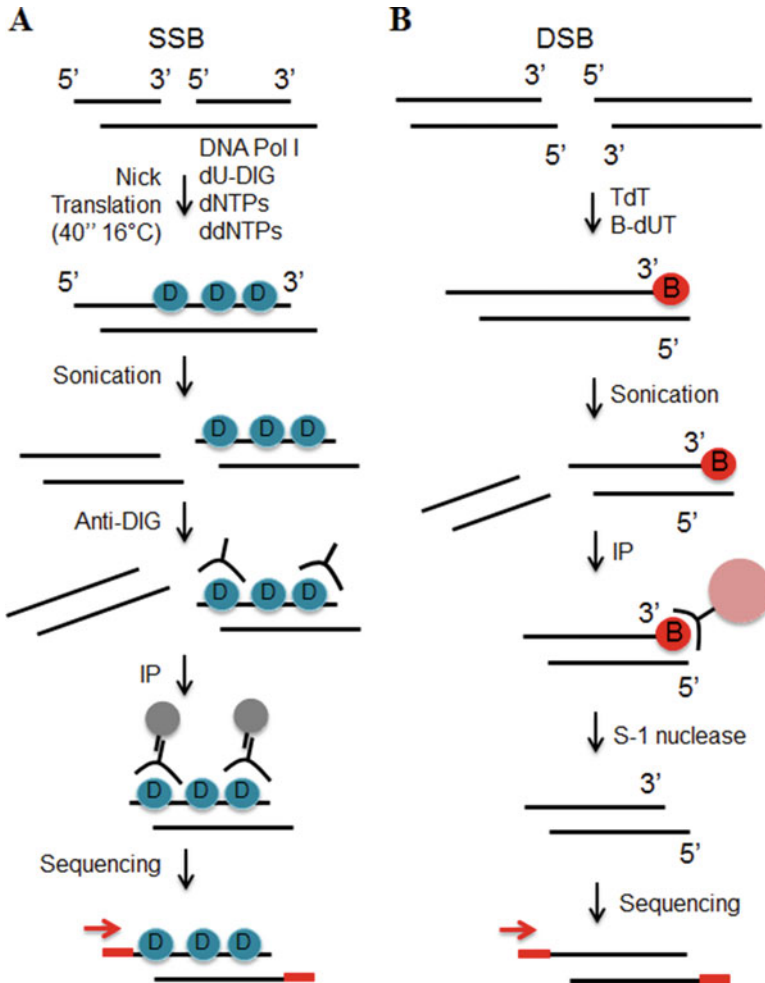
**Fig. 1** Types of DNA damage. Exogenous and endogenous DNA damaging agents generate various types of lesions including SSBs and DSBs. PARP predominantly acts as a sensor of SSB [17]. RPA binds to regions of single-stranded DNA (ssDNA) that are exposed to stalled replication forks or after DSB resection [18]. The multifunctional MRN complex and KU detect DSBs, FANCM is required for the DNA interstrand crosslink (ICL)-induced checkpoint response [19]. FANCM = Fanconi anemia complementation group M; ICL = interstrand crosslink; MRN = MRE11-RAD50-NBS1 complex; PARP = poly(ADP-ribose) polymerase; RPA = replication protein A

Several assays to detect DSBs were developed during recent years that either rely on the mapping of chromosomal translocation by using a “bait” DSB introduced into the genome [3–5] or by capturing the double-strand broken ends with oligodeoxynucleotides or integrase-defective lentiviral vectors. The junctions between the genomic DNA and the transfected DNA correspond to the location of the DSBs and can successively be sequenced [6, 7].

While these methods provide high resolution of DSB and low background, they are limited to the detection of only translocated breaks or integrated DNA. Consequently, they cannot be used to study highly dynamic genome instability events. Additionally, due to the necessity to introduce a “bait” DSB into the genome or double-stranded DNA into the nuclei of living cells, the use of these methods is now restricted to well-established cell lines, not primary cells or tissues. Real-time capture of genome damaging events is provided by recent methods based on direct in situ labeling of DSBs [8, 9]. However, the methods were not implemented to identify SSBs and so they do not allow a direct comparison between DSBs and SSBs in the same genome context [10].

Here, we present two comprehensive experimental and computational approaches to map DSBs and SSBs across the genome of human cells [11]. These are based on the direct labeling of breaks with two independent strategies. The approaches could be extended to any cell lines and tissue and even on previously isolated genomic DNA.

To detect SSBs (SSB-Seq, Fig. 2a) the high molecular weight genomic DNA isolated from HCT116 cells is subjected to nick translation. Nick translation is a tagging technique in which DNA polymerase I is used to replace some of the nucleotides of a DNA



**Fig. 2** DNA breaks mapping workflow. **(a)** SSBs are labeled during nick translation using nucleotides covalently linked to digoxigenin (*blue circle*). The DNA is subsequently purified, sonicated and incubated with anti-digoxigenin antibody (anti-DIG). The immuno-precipitated DNA is sequenced. **(b)** 3' tails of DSBs are ligated to biotinylated nucleotides (*red circle*). After sonication the labeled fragments are captured on streptavidin beads (*pink circle*). Tails are removed from released fragments and DNA is sequenced

sequence with their labeled analogues—digoxigenin-modified nucleotides—creating a tagged DNA fragment that can be immuno-precipitated with anti-digoxigenin antibody and sequenced [12]. To increase the resolution of mapping, we restricted the digoxigenin-labeling to a small patch of DNA by including in the reaction dideoxynucleotides (ddNTP), to inhibit excessive chain elongation by DNA polymerase I. As a control for the labeling, samples were also nick-translated without digoxigenin-labeled nucleotides.

To map DSBs (DSB-Seq, Fig. 2b) the double-stranded DNA ends were 3'-end tailed with terminal deoxynucleotidyl transferase



(TdT), a DNA polymerase I that catalyzes the addition of nucleotides to the 3' terminus of a DNA molecule. In the presence of TdT, biotinylated nucleotides were added to the region of DSB and after fragmentation, the biotinylated DNA was streptavidin-selected. In parallel, 3'-tailing was performed in the absence of biotinylated nucleotides, which constituted our negative control for labeling and selection. To remove the biotinylated-tails, samples were treated with S-1 nuclease and the resulting DNA was purified and sequenced (this technique was modified from our previous work [13]).

Our methods provide maps of the DNA break landscape, in various cell types and tissues, and in different experimental conditions.

---

## 2 Materials

### 2.1 Reagents

1. Proteinase K (Solution in water, 20 mg/ml).
2. Phenol, Tris saturated.
3. Phenol:Chloroform:Isoamyl Alcohol 25:24:1.
4. Ethanol 100%.
5. Ammonium acetate 7.5 M.
6. RNase, DNase-free (500 µg/ml).
7. SDS 10%.
8. dATP, dGTP, dCTP, dTTP.
9. Digoxigenin-11-dUTP.
10. ddATP, ddGTP, ddCTP, ddTTP.
11. *E. coli* DNA polymerase I.
12. EDTA 0.5 M, (pH 8.0).
13. Anti-digoxigenin antibody.
14. Protein G-Sepharose beads.
15. Terminal transferase (TdT).
16. Biotin-16-dUTP.
17. TTP.
18. Streptavidin-coated beads.
19. S-1 nuclease.
20. Klenow (exo-).
21. T4 DNA ligase.
22. Agarose (agarose gels are prepared in TAE buffer).
23. 2% precast agarose gel (e.g., E-Gel, Invitrogen).
24. Illumina adapter (Adaptor oligo mix).
25. Illumina primers (Fw: 5'-aca ctc ttt ccc tac acg acg c-3'/Rv: 5'-caa gca gaa gac ggc ata cga gc-3').

## 2.2 Buffers

1. Lysis buffer: 10 mM Tris-HCl (pH 8.0), 100 mM EDTA (pH 8.0), 0.5% SDS.
2. TE buffer: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0).
3. PBS buffer: 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KHPO<sub>4</sub> (pH 7.2).
4. NP-40 buffer: 20 mM Tris-HCl (pH 8.0), 137 mM NaCl, 10% Glycerol, 1% NP-40, 2 mM EDTA (pH 8.0).
5. TdT reaction buffer: 50 mM Potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate (pH 7.9).
6. DSB washing buffer: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), and 2.0 M NaCl.
7. DSB elution buffer: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 1 M NaCl, and 2 M β-mercaptoethanol.
8. TAE buffer: 40 mM Tris (pH 7.6), 20 mM acetic acid, 1 mM EDTA (pH 8.0).
9. End repair buffer: 33 mM Tris-acetate (pH 7.0), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT.

## 2.3 Equipment

1. Gel electrophoresis apparatus.
2. Spectrophotometer.
3. Thermoblock (a heating and cooling unit).
4. Sonicator.
5. Rocker.
6. E-Gel Precast Agarose Electrophoresis System (Invitrogen).
7. Genome sequencer.

## 2.4 Kits

1. QIAquick PCR Purification Kit (Qiagen).
2. Epicentre DNA END-Repair kit.
3. MinElute gel extraction kit (Qiagen).

---

## 3 Methods

### 3.1 Purification of High Molecular Weight DNA

1. Wash cells ( $1 \times 10^8$ ) twice with ice-cold PBS (*see Note 1*).
2. Lyse cells with 10 ml of lysis buffer (*see Note 2*).
3. Collect lysate by scraping and transfer the suspension in a 50 ml conical tube.
4. Digest the sample overnight with proteinase K (200 μg/ml) at 52 °C.
5. Purify DNA twice with phenol (1:1 v/v) and once with phenol-chloroform (1:1 v/v) (*see Note 3*).

6. Precipitate DNA with 2 volumes of ethanol 100% in the presence of 2 M ammonium acetate (*see Note 4*).
7. Centrifuge  $4,400 \times g$  for 30 min. Remove the supernatant, add ethanol 70%, mix by inverting the sample, centrifuge  $4,400 \times g$  for 25 min.
8. Air-dry the pellet (*see Note 5*).
9. Add to the pellet 500  $\mu$ l of TE buffer and incubate it for 12 h at room temperature with gentle rotation (*see Note 6*).
10. Incubate the sample with 5  $\mu$ g of pancreatic RNase for 1 h at 37°.
11. Adjust the sample to 0.5% SDS and incubate for 1 h at 55 °C with proteinase K (200  $\mu$ g/ml).
12. Bring the volume to 10 ml with TE buffer and extract the DNA twice with phenol-chloroform (1:1 v/v) (*see Note 5*).
13. Precipitate the DNA with 2 volumes of ethanol 100% in the presence of 2 M ammonium acetate (*see Note 6*).
14. Air-dry the pellet (*see Note 5*).
15. Add to the pellet 1 ml of TE buffer and incubate it for 12 h at room temperature with gentle rotation.
16. Run 10  $\mu$ l of DNA on a 0.6% agarose gel (*see Note 7*).
17. Determine DNA concentration with a spectrophotometer (*see Note 8*).

### 3.2 SSB-Seq

1. In a final volume of 1.5 ml, incubate with gentle mixing 500  $\mu$ g of DNA for 40 s at 16 °C with a mixture of 200  $\mu$ M of dATP, dGTP, dCTP and 20  $\mu$ M of digoxigenin-11-dUTP, 117  $\mu$ M of ddATP, ddGTP, ddCTP and 1000 units of *E. coli* DNA polymerase I. As a negative control for labeling, incubate 500  $\mu$ g of DNA with the same reagents except digoxigenin-11-dUTP that is substituted with 20  $\mu$ M of dTTP (*see Note 9*).
2. Stop the reaction with 50  $\mu$ M EDTA.
3. Extract DNA with phenol-chloroform (1:1 v/v).
4. Precipitate DNA in the presence of 2 volumes of 100% ethanol and 2 M ammonium acetate.
5. Centrifuge  $16,000 \times g$  for 30 min.
6. Air-dry pellet.
7. Add 1 ml of TE and mix by vortexing.
8. To help resuspension, incubate the sample at 45 °C for 15 min, mix by vortexing.
9. Precipitate DNA in the presence of 2 volumes of 100% ethanol and 2 M ammonium acetate.

10. Centrifuge  $16,000 \times g$  for 30 min. Remove the supernatant. Add 70% ethanol. Centrifuge  $16,000 \times g$  for 15 min (*see Note 10*). Remove the supernatant.
11. Air-dry pellet.
12. Add to the pellet 300  $\mu$ l of TE buffer.
13. Shear the DNA by sonication to an average fragment size of 250 bp (*see Note 11*).
14. Incubate the sample with 10  $\mu$ g of anti-digoxigenin antibody at 4 °C, overnight with gentle rotation.
15. To recover the immuno-complexes add 60  $\mu$ l of Protein G-Sepharose beads (*see Note 12*) and incubate for 4 h at 4 °C.
16. Wash the beads once with PBS buffer, three times with NP-40 buffer; twice with TE buffer; and finally add to the beads 200  $\mu$ l of TE buffer (*see Note 13*).
17. Adjust the sample to 0.5% SDS and digest with proteinase K (200  $\mu$ g/ml) at 65 °C overnight.
18. Purify DNA using QIAquick PCR Purification Kit according to the manufacturer's instructions and quantify the eluate (*see Note 14*).

### 3.3 DSB-Seq

1. Incubate 500  $\mu$ g of DNA in 3 ml of TdT buffer with 24,000 U TdT, 0.5 mM dCTP and 5 mM CoCl<sub>2</sub> at 37 °C for 5 min.
2. Add 0.02 mM Biotin-16-dUTP and incubate at 37 °C for 30 min. As a control of labeling, incubate 500  $\mu$ g of DNA with the same reagents substituting TTP for Biotin-16-dUTP.
3. Stop the reaction by adding EDTA to a final concentration of 20  $\mu$ M.
4. Extract the sample with phenol-chloroform (1:1 v/v).
5. Precipitate DNA in the presence of 2 volumes of 100% ethanol and 2 M ammonium acetate.
6. Centrifuge  $16,000 \times g$  for 30 min.
7. Air-dry pellet.
8. Add 1 ml of TE and mix by vortexing.
9. To help resuspension, incubate the sample at 45 °C for 15 min, mix by vortexing.
10. Precipitate DNA in the presence of 2 volumes of 100% ethanol and 2 M ammonium acetate.
11. Centrifuge  $16,000 \times g$  for 30 min. Remove the supernatant. Add ethanol 70%. Centrifuge  $16,000 \times g$  for 15 min (*see Note 15*). Remove the supernatant.
12. Air-dry the pellet and dissolve it in 200 ml of TE buffer.

13. Sonicate the biotinylated DNA to generate 200–400 bp DNA fragments (*see Note 11*).
14. Capture the biotinylated fragments by mixing the DNA with 200  $\mu$ l of streptavidin-coated beads (*see Note 16*).
15. Incubate for 3 h at room temperature with agitation.
16. Wash the beads four times with DSB washing buffer, incubating the sample at 50 °C (*see Note 17*).
17. Wash the beads four times with DSB washing buffer, incubating the sample at room temperature (*see Note 18*).
18. To disrupt biotin-streptavidin complexes incubate the sample in 200  $\mu$ l DSB elution buffer at 75 °C for 4 h (*see Note 19*).
19. Purify free DNA fragments with a QIAquick PCR Purification Kit.
20. To remove the biotinylated tails from DNA, incubate the sample with 30 U of S-I nuclease in 110  $\mu$ l of recommended buffer for 30 min at 37 °C.
21. Purify DNA with a QIAquick PCR Purification Kit.
22. Quantify the recovered DNA.

### **3.4 Template Preparation for Sequencing Analysis**

The DNA recovered from the immuno-precipitation/biotin-step-avidin selection as well as 10  $\mu$ l of the Input DNA—genomic DNA not subjected to immuno-precipitation—will be subjected to library preparation and sequencing (*see Note 20*).

1. To generate blunt-ended DNA, incubate the DNA for 45 min at room temperature in the 25  $\mu$ l reaction with a mixture of End repair buffer, 0.25 mM of each dNTPs, 1 mM ATP, and 1  $\mu$ l End-Repair Enzyme mix (T4 DNA polymerase + T4 PNK) (*see Note 21*).
2. Purify the DNA with a MinElute Reaction Cleanup Kit (*see Note 22*).
3. In the 25  $\mu$ l reaction, treat the blunt-ended DNA with 15 units of Klenow(exo-) for 30 min at 37 °C in the presence of 0.2 mM dATP to generate a protruding 3'A base used for adaptor ligation.
4. Purify the DNA with a MinElute Reaction Cleanup Kit.
5. In the 20  $\mu$ l reaction, ligate Illumina adapter to the end of DNA fragments by incubating with 0.1  $\mu$ l Adaptor oligo mix and 1000 units of T4 DNA ligase at room temperature for 30 min.
6. Purify DNA using MinElute Reaction Cleanup Kit.
7. To size-select the adapter ligated DNA, run the sample through 2% E-Gel electrophoresis.
8. Excise the gel slice, around the 200–400 bp region.

9. Purify DNA from the gel using the MinElute gel extraction kit and elute in a final volume of 12  $\mu$ l elution buffer.
10. Amplify the DNA for 18 cycles using Illumina primers (Fw: 5'-aca ctc ttt ccc tac acg acg c-3'/Rv: 5'-caa gca gaa gac ggc ata cga gc-3') according to the following protocol: 98 °C for 30 s; 65 °C for 30 s; 72 °C for 30 s.
11. Run the PCR product through 2.5% agarose gel and excise the gel slice around 220 bps–500 bps.
12. Purify the DNA from the gel using MinElute gel extraction kit.
13. The purified DNA is used directly for cluster generation and sequencing analysis using the Illumina Genome Analyzer following the manufacturer's protocols.

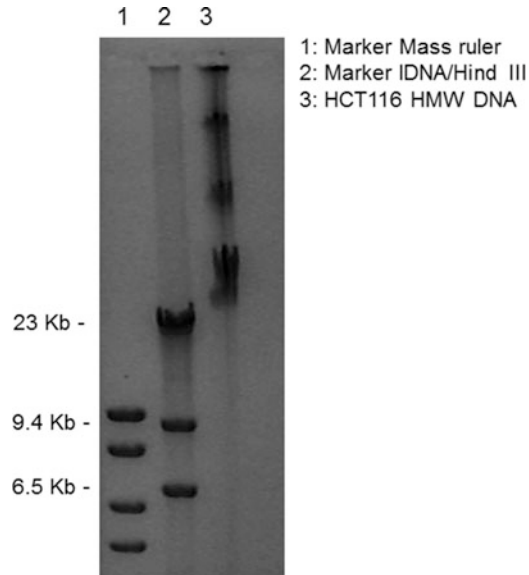
### 3.5 Processing of Sequencing Data

1. Process sequencing data from SSB-Seq and DSB-Seq protocols using Illumina Analysis Pipeline (image analysis and base calling).
2. Check quality of high-throughput sequencing data with the FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).
3. Align short sequencing reads of length 36 bp to the reference human genome using the Bowtie 2 tool (version 2.2.2) with default parameters [14].
4. Remove redundant reads from the datasets, to minimize potential PCR bias, using Samtools package (<http://www.htslib.org/doc/samtools.html>).
5. Generate a read density visualization of the aligned sequencing data that can be viewed in most genome browsers (e.g., UCSC Genome Browser <http://genome.ucsc.edu/cgi-bin/hgGateway>). For wiggle track format (wig), extend the reads to the average length of the genomic fragments, count the number of reads at each position in the genome, and normalize the library size to 1 million reads [15]. For faster upload and display of wiggle file it can be compressed to bigWig format with the wigToBigWig tool (<http://hgdownload.cse.ucsc.edu/admin/exe/>).

---

## 4 Notes

1. This protocol is designed for adherent cells. If suspension cells are used, centrifuge cells  $140 \times g$  for 4 min, at 4 °C.
2. Lysis buffer should be kept at room temperature.
3. In general, during the steps preceding nick-translation DNA should be handled very gently. After the addition of phenol or phenol-chloroform mix the phases by inverting the tube for a



**Fig. 3** Representative example of High Molecular Weight (HMW) DNA after the purification steps described in Subheading 3.1. In lanes 1 and 2 we run two different markers. The numbers on the left refer to the molecular weight of the marker in lane 2

few minutes. The aqueous phase should be purified from the organic phase with a 25 ml pipette, sucking the liquid very slowly. Do not vortex the sample.

4. Mix by gently inverting the sample for 3 min until a white puff forms.
5. The sample should not be completely dry, otherwise resuspension will be difficult.
6. Do not resuspend the pellet by pipetting.
7. This step is performed to check the quality of DNA, which should run around 23 kb or higher molecular weight. The sample is viscous; therefore, the tip should be cut before transferring the aliquot for gel electrophoresis. *See Fig. 3* as a representative example.
8.  $1 \times 10^8$  HCT116 cells give approximately 1 mg of DNA. The 260/280 ratio should be around 1.8 and 1.9 and the 260/230 ratio between 2.0 and 2.2.
9. It is important to control the time of nick translation. Thus, it is suggested to incubate one sample at the time.
10. The double precipitation with ammonium acetate is necessary to remove free digoxigenin-11-dUTP.
11. Check the DNA fragment sizes by running a 1% agarose gel. In our procedure sonication was performed with an ultrasonic

sonicator (Bioruptor, Diagenode) at medium power, by pulsing 30 times for 30 s and incubating on ice for 30 s between each pulse.

12. To prepare the beads, mix the slurry, take 60  $\mu$ l of beads per immuno-precipitation, wash the beads three times with ice-cold PBS buffer. Perform each wash by adding 900  $\mu$ l of ice-cold PBS buffer. Incubate the beads for 10 min by rocking at 4 °C. Centrifuge the sample 4 min at 1500  $\times g$ , 4 °C. Remove the supernatant.
13. Perform each wash by adding 900  $\mu$ l of ice-cold washing buffer. Incubate the immuno-complex for 10 min by rocking, centrifuge the sample for 4 min at 1500  $\times g$ , 4 °C. Remove the supernatant.
14. This protocol is suitable for mapping SSBs with a free 3' end. SSBs generated during topoisomerase 1 catalytic cycle will be labeled only after the treatment of DNA with tyrosyl-DNA phosphodiesterase 1 (TDP1) [16].
15. The double precipitation with ammonium acetate is necessary to remove free Biotin-16-dUTP.
16. Available from Invitrogen (Dynabeads kilobase BINDER Kit, DYNAL). To prepare the beads, use the magnet to separate beads from the supernatant. Add to the beads 200  $\mu$ l of binding buffer (provided with the beads). Mix for 5 min. Remove the supernatant. Repeat the wash.
17. To perform each wash, add 900  $\mu$ l of washing solution. Incubate at 50° for 5 min with agitation. Use the magnet to separate the beads from the washing buffer. Add a new washing solution.
18. To perform each wash, add 900  $\mu$ l of the washing solution. Incubate at room temperature for 5 min, with agitation. Use the magnet to separate the beads from the washing buffer. Add a new washing solution.
19. For a better yield, add 100  $\mu$ l of DSB elution buffer, incubate at 75 °C for 2 h, use the magnet to separate the supernatant from the beads. Keep the supernatant. Add to the beads 100  $\mu$ l of new DSB elution buffer, incubate at 75 °C for 2 h, use the magnet to separate the supernatant from the beads. Pool the supernatants together.
20. The Input DNA is sonicated.
21. The Epicentre DNA END-Repair kit is available at Epicentre Biotechnologies.
22. This kit is available at QIAGEN.



## References

1. Iacovoni JS et al (2010) High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J* 29(8):1446–1457
2. Blitzblau HG, Hochwagen A (2011) Genome-wide detection of meiotic DNA double-strand break hotspots using single-stranded DNA. *Methods Mol Biol* 745:47–63
3. Hu J et al (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc* 11(5):853–871
4. Klein IA et al (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* 147(1):95–106
5. Chiarle R et al (2011) Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 147(1):107–119
6. Tsai SQ et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33(2):187–197
7. Wang XL et al (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 33(2):175–178
8. Crosetto N et al (2013) Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods* 10(4):361–365
9. Canela A et al (2016) DNA breaks and end resection measured genome-wide by end sequencing. *Mol Cell* 63(5):898–911
10. Aguilera A, Garcia-Muse T (2013) Causes of genome instability. *Annu Rev Genet* 47:1–32
11. Baranello L et al (2014) DNA break mapping reveals topoisomerase II activity genome-wide. *Int J Mol Sci* 15(7):13111–13122
12. Rigby PW et al (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J Mol Biol* 113(1):237–251
13. Kouzine F et al (2013) Global regulation of promoter melting in naive lymphocytes. *Cell* 153(5):988–999
14. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* 9(4):357–359
15. Bardet AF et al (2012) A computational pipeline for comparative ChIP-seq analyses. *Nat Protoc* 7(1):45–61
16. Caldecott KW (2008) Single-strand break repair and genetic disease. *Nat Rev Genet* 9(8):619–631
17. Barnes DE, Lindahl T (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* 38:445–476
18. Sartori AA et al (2007) Human CtIP promotes DNA end resection. *Nature* 450(7169):509–514
19. Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* 45:247–271

## Genome-Wide Profiling of DNA Double-Strand Breaks by the BLESS and BLISS Methods

Reza Mirzazadeh, Tomasz Kallas, Magda Bienko, and Nicola Crosetto

### Abstract

DNA double-strand breaks (DSBs) are major DNA lesions that are constantly formed during physiological processes such as DNA replication, transcription, and recombination, or as a result of exogenous agents such as ionizing radiation, radiomimetic drugs, and genome editing nucleases. Unrepaired DSBs threaten genomic stability by leading to the formation of potentially oncogenic rearrangements such as translocations. In past few years, several methods based on next-generation sequencing (NGS) have been developed to study the genome-wide distribution of DSBs or their conversion to translocation events. We developed Breaks Labeling, Enrichment on Streptavidin, and Sequencing (BLESS), which was the first method for direct labeling of DSBs in situ followed by their genome-wide mapping at nucleotide resolution (Crosetto et al., *Nat Methods* 10:361–365, 2013). Recently, we have further expanded the quantitative nature, applicability, and scalability of BLESS by developing Breaks Labeling In Situ and Sequencing (BLISS) (Yan et al., *Nat Commun* 8:15058, 2017). Here, we first present an overview of existing methods for genome-wide localization of DSBs, and then focus on the BLESS and BLISS methods, discussing different assay design options depending on the sample type and application.

**Key words** DNA double-strand breaks, Genome instability, Next-generation sequencing, Genome editing, BLESS, BLISS

---

### 1 Introduction

The integrity of DNA is constantly challenged by exogenous and endogenous damaging agents. In multicellular organisms, individual cells are thought to undergo thousands of lesion events per day, most of which are promptly repaired by an intricate network of signaling pathways in a process known as DNA damage response [1]. Among different lesion types, DNA double-strand breaks (DSBs) pose a major threat to genomic stability and organismal homeostasis, since unrepaired or mis-repaired DSBs can result in a variety of mutations—chromosomal rearrangements as well as small insertions/deletions (indels) due to the error-prone activity of the Non-Homologous End Joining (NHEJ) repair process—that can result in cell death or potentially initiate carcinogenesis.

DSBs arise as a consequence of exposure to exogenous damaging agents, such as ionizing radiation and several chemotherapeutic drugs, but the vast majority of DSBs is formed during physiological cellular processes, including DNA recombination, replication, and transcription [2, 3]. Another source of DSBs that has recently become under the spotlight are endonucleases used for genome editing purposes, including zinc-finger nucleases [4], transcription activator-like effector nucleases (TALENs) [5], and endonucleases belonging to the clustered, regularly interspaced, short palindromic repeats (CRISPR) system, such as Cas9 [6] and Cpf1 [7]. While these enzymatic activities mainly induce DSBs at a specific genomic location where the desired genome-editing event is planned, the formation of unspecific, off-target DSBs is not a completely unlikely event and represents an important concern particularly in the context of medical applications [8]. Thus, systematic genome-wide characterization of the off-target activity of designer endonucleases is necessary before they can fulfill their therapeutic potential.

In addition to helping characterize the safety profile of genome-editing endonucleases, the ability to accurately map the location of DSBs along the genome is essential to understand the role of DNA sequence and chromatin structure in the formation, processing, and repair of these lesions. In recent years, several methods have been developed to localize DSBs genome-wide by harnessing the high-throughput capacity of next-generation sequencing (NGS). A summary of these methods and their main characteristics is presented in Table 1. Chromatin immunoprecipitation sequencing (ChIP-seq) targeting the phosphorylated histone variant H2A.X ( $\gamma$ H2A.X) around DSBs was initially used to identify fragile sites in yeast [9]. Subsequently, ChIP-seq against  $\gamma$ H2A.X or other repair factors recruited at DSB sites, such as 53BP1, was applied to investigate the process of repair of DSBs induced by the AsiSI endonuclease in mammalian cells [10]. Recently, ChIP-seq against Topoisomerase II $\beta$  in neuronal stem/progenitor cells showed that neuronal activity stimulation results in the formation of DSBs in the promoters of early-response genes, which is essential to their transcription [11], highlighting the importance of mapping DSBs at genomic scale in order to gain new insights into fundamental biological processes.

A major limitation of ChIP-seq, however, is that  $\gamma$ H2A.X and other DSB markers extend around DSBs for thousands or even millions bases, thus preventing the identification of fragile hotspots at nucleotide resolution. In order to circumvent this problem, we pioneered BLESS, which was the first method enabling direct labeling of DSBs and identification of their genomic location by NGS [12]. Using BLESS, we identified numerous aphidicolin-sensitive regions that recurrently break upon replication stress, and found that certain repeat element classes, such as alpha satellites, are particularly prone to break under replication stress. Subsequently, BLESS

**Table 1**  
**Comparison between different methods for DSBs sequencing**

Method	Detection	Main features	Sample (input)	Reported applications
BLESS	Direct	In situ DSB blunting and ligation of biotinylated adapters. DSB capture on streptavidin	Fixed cells (at least $10^6$ cells)	Replication stress-induced DSBs in mammalian cells [12], Cas9 specificity [13, 14]
DSBCapture	Direct	In situ DSB blunting and A-tailing. Modified BLESS adapters containing Illumina adapter sequences	Fixed cells (at least $10^6$ cells)	DSBs at G-quadruplex-rich sites, active genes and transcription start sites [15]
END-seq	Direct	In vivo DSB blunting and A-tailing in agarose plugs. Modified BLESS adapters containing Illumina adapter sequences	Live cells (at least $10^7$ cells)	AsiSI-induced DSBs resection mapping, RAG endonuclease specificity [16]
BLISS	Direct	In situ DSB blunting, A-tailing, and ligation of adapters containing the T7 promoter sequence and UMIs. Selective DSB amplification by in vitro transcription	Fixed cells, tissue sections (at least $10^3$ cells)	Transcription-associated DSBs in cells and tissue and Cas9 and Cpf1 specificity [17]
Digenome-seq	Direct	In vitro nuclease digestion of purified genomic DNA and detection of DSBs by whole-genome sequencing	Purified DNA	Cas9 and Cpf1 specificity [18, 19]
ChIP-seq	Indirect	Capture of chromatin containing DSBs markers such as $\gamma$ H2A.X	Fixed cells (at least $10^7$ cells)	Replication stress-induced DSBs in yeast [9], AsiSI-induced DSBs processing in mammalian cells [10], transcription-associated DSBs in neuronal cells [14]
GUIDE-seq	Indirect	In vivo DSB labeling by incorporation of dsDNA oligos through NHEJ-mediated repair	Transfected live cells	Cas9 and Cpf1 specificity [20, 21]
IDLV capture	Indirect	In vivo DSB labeling by random incorporation of integration defective lentiviral vectors through NHEJ-mediated repair	Transduced live cells	Cas9 and TALENs specificity [22]
LAM-HTGTS	Indirect	In vivo induction of DSBs and sequencing of translocation products originated from NHEJ-mediated repair	Live cells treated to induce translocations	Cas9 specificity [23], transcription-associated DSBs in neuronal cells [24]

was also applied to profile the off-target activity of CRISPR/Cas9 on a genome-wide scale [13, 14]. Recently, two new methods, DSB-Capture [15] and END-seq [16], extended the BLESS design by introducing an A-tailing step, improving the adapter design, and implementing DSB labeling in vivo based on the use of agarose plugs (in END-seq). An important limitation of BLESS and its recent modifications, however, is that a substantial amount of starting material is required (typically, more than  $10^7$  cells), thus challenging their application to low-input samples and tissue specimens. Moreover, these methods are labor-intensive and thus difficult to scale up to large numbers of samples. To overcome these limitations, we recently improved the BLESS concept by developing DNA Double-Strand Breaks Labeling In Situ and Sequencing (BLISS). In BLISS, DSBs are in situ ligated to a dsDNA oligonucleotide adapter that contains the T7 promoter sequence enabling selective linear amplification of the labeled DSBs by in vitro transcription (IVT). In addition, unique molecular identifiers (UMIs) [25] and sample-specific barcodes are incorporated at the site of in situ DSB ligation, allowing for multiplexing and quantitative mapping of DSBs even in low-input samples. A side-by-side comparison and detailed description of the BLESS and BLISS methods is presented in Fig. 1.

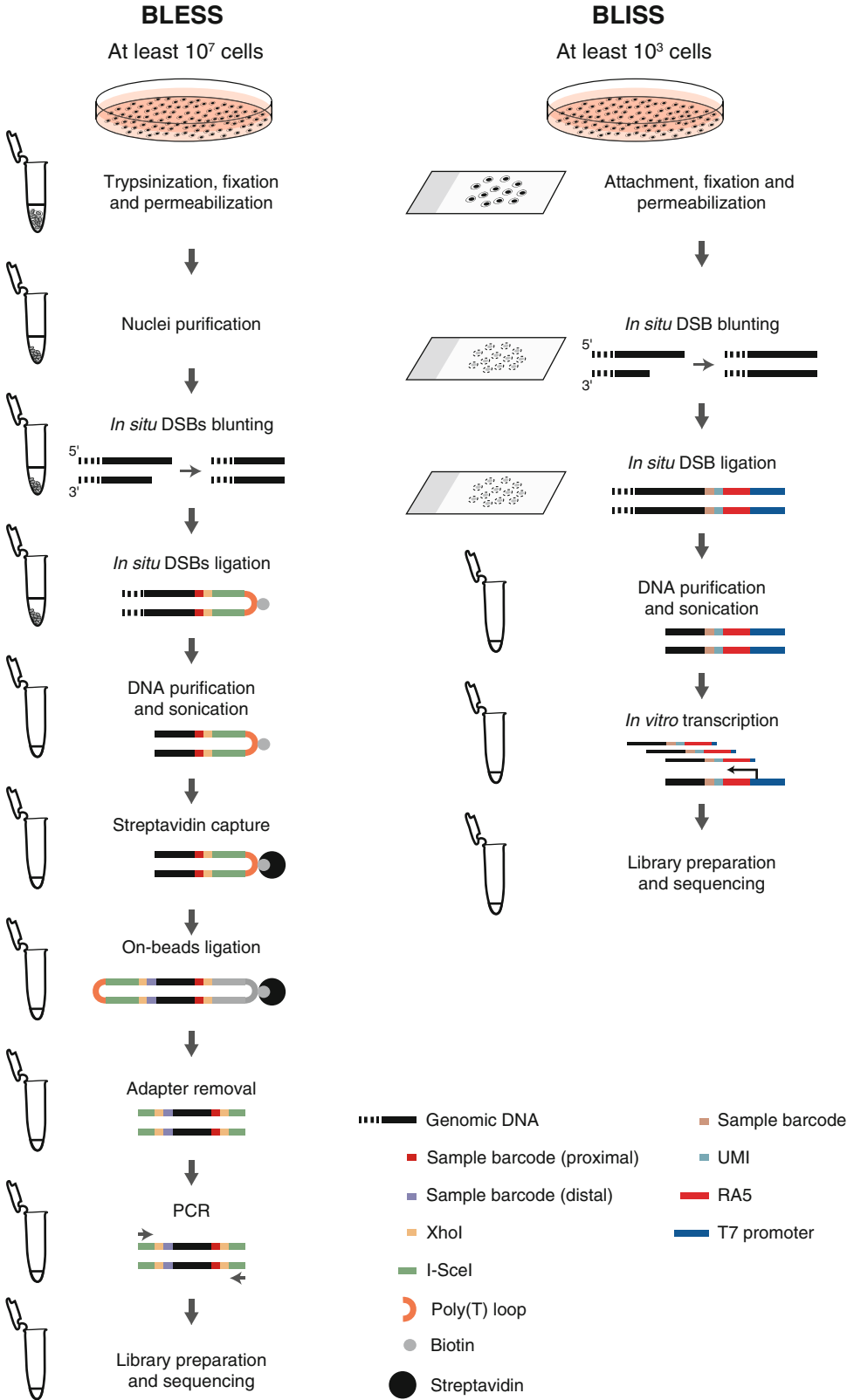
In parallel to the BLESS and its derivative methods, other genome-wide DSB localization approaches have recently emerged, particularly with the purpose of characterizing the specificity profile of genome-editing endonucleases. These methods include in vitro digested whole-genome sequencing (Digenome-seq) [18], GUIDE-seq [20], integration-deficient lentiviral vector (IDLV) capture [22], and high-throughput, genome-wide translocation sequencing (HTGTS) [23]. An overview of these methods and their comparison to BLESS and BLISS is presented in Table 1. The main distinctive feature of these methods in comparison to BLESS and its derivatives is that DSBs are not directly labeled (except for Digenome-seq, in which purified DNA is digested in vitro). GUIDEseq, IDLV incorporation, and HTGTS methods require the activity of NHEJ to label DSBs, thus potentially missing DSBs that are not repaired through this pathway. Furthermore, exogenous DNA introduction in the case of GUIDEseq and IDLV integration can be challenging in primary cells and in tissues, where transfection efficiency and toxicity in the former, and delivery in the latter may be limiting factors. Here, we focus on the BLESS and BLISS methods, and provide a flexible step-by-step protocol that can be adapted to various applications.

---

## 2 Materials

### 2.1 Common Reagents for BLESS and BLISS

1. Methanol-free paraformaldehyde (PFA) 16%.
2. Lysis buffer 1 (LB1): Tris-HCl 10 mM, NaCl 10 mM, EDTA 1 mM, Triton X-100 0.2%, pH 8 at 4 °C.



**Fig. 1** Side-by-side comparison of BLESS and BLISS workflows. Both the procedures start by fixing cells in order to stabilize chromatin and prevent the formation of artificial breaks during subsequent steps. Next, cells

3. Lysis buffer 2 (LB2): Tris-HCl 10 mM, NaCl 150 mM, EDTA 1 mM, SDS 0.3%, pH 8 at 25 °C.
4. Nuclease-free water.
5. Nuclease-free Phosphate-Buffered Saline (10×), pH 7.4.
6. CutSmart<sup>®</sup> buffer (NEB, cat. no. B7204S).
7. Quick Blunting<sup>™</sup> Kit (NEB, cat. no. E1201L).
8. NEBNext<sup>®</sup> dA-Tailing Module (NEB, cat. no. E6053L), OPTIONAL.
9. BSA (50 mg/ml) (e.g., Thermo, cat. no. AM2616).
10. ATP Solution (100 mM).
11. T4 DNA Ligase (e.g., NEB, cat. no. M0202M).
12. High-salt wash buffer (HSW): Tris-HCl 10 mM, NaCl 2 M, EDTA 2 mM, Triton X-100 0.5%, pH 8 at 25 °C.
13. Proteinase K, Molecular Biology Grade (NEB, cat. no. P8107S).
14. Nuclease-free TE buffer.
15. NEBNext<sup>®</sup> High-Fidelity 2× PCR Master Mix (NEB, cat. no. M0541L).



**Fig. 1** (continued) are permeabilized in order for in situ DSB labeling reactions to take place in the crowded environment of cross-linked chromatin. In BLESS, but not in BLISS, a short incubation in the presence of Proteinase K is used to purify nuclei and render them more accessible for in situ reactions (*see Note 1*). After permeabilization, in situ blunting is performed in order to convert the DSB ends that contain an overhang (such as the intermediates of end-resection formed during homologous recombination repair) to a ligatable configuration. Blunt DSBs are then ligated in situ using oligo adapters with a different design in BLESS and BLISS (*see Note 2* and Table 2). A key difference between BLESS and BLISS is that, while in BLESS permeabilization and in situ reactions are performed in suspension, in BLISS all these steps are done on a solid surface (a microscope slide or coverslip). This avoids the requirement for multiple centrifugations (which cause progressive sample loss and increased turnaround time) and enables safe processing even of very low-input samples consisting of few thousand cells, such as rare cell populations and precious clinical specimens. In addition, BLISS may be performed directly on fixed tissue sections mounted on a microscope coverslip or slide, opening the possibility to study DSBs in a wide range of clinically relevant samples. After in situ labeling of DSBs, DNA can be safely extracted and purified (in BLISS, cells or tissue sections are either scraped off the slide/coverslip or they are captured using Laser Capture Microdissection or less resolved methods such as the Pinpoint Slide DNA Isolation System<sup>™</sup> [Zymo Research, cat. no. D3001]). In BLESS, the genomic sequence surrounding the DSB ends is selectively captured using affinity purification on streptavidin beads. In BLISS, DSBs are selectively amplified using IVT driven by the T7 RNA polymerase bound to the T7 promoter sequence ligated to the DSB ends. Finally, preparation of sequencing libraries using the captured/amplified material is achieved using standard Illumina technology (*see Notes 18* and *20*). Major advantages of BLISS over BLESS include: (1) lower input requirement (few million cells in BLESS, as little as 10<sup>3</sup> cells in BLISS); (2) faster turnaround (4 days from fixation to library for BLISS, at least 10 days for BLESS); (3) simpler workflow (especially due to the avoidance of multiple centrifugations and Proteinase K incubation); (4) improved adapter design, enabling more quantitative data analysis thanks to UMIs; (5) easy scalability by performing all in situ reactions in multi-well plates

**2.2 Reagents  
for BLESS**

1. NEBuffer 2 (NEB, cat. no. B7002S).
2. Dynabeads<sup>®</sup> MyOne<sup>™</sup> Streptavidin C1 (Thermo, cat. no. 65001).
3. I-SceI (NEB, cat. no. R0694L).
4. XhoI (NEB, cat. no. R0146L).
5. NEBNext<sup>®</sup> DNA Library Prep Master Mix Set for Illumina (NEB, cat. no. E6040L).

**2.3 Reagents  
for BLISS**

1. Poly-L-lysine (PLL) solution (Sigma, cat. no. P8920-100ML).
2. Nucleus Suspension Buffer (NSB): NaCl 146 mM, Tris-HCl 10 mM, CaCl<sub>2</sub> 1 mM, MgCl<sub>2</sub> 21 mM, BSA 0.05%, Nonidet P-40 0.2%, pH 7.8. OPTIONAL (for tissue-BLISS).
3. T4 Polynucleotide Kinase (e.g., NEB, cat. no. M0201L).
4. MEGAscript<sup>®</sup> T7 Transcription Kit (Thermo, cat. no. AM1334).
5. DNase I, RNase-free.
6. RiboSafe RNase Inhibitor (Bioline, cat. no. 65027).
7. AMPure XP (Beckman Coulter, cat. no. A63881).
8. Agencourt RNAClean XP (Beckman Coulter, cat. no. A63987).
9. T4 RNA Ligase 2, truncated (NEB, cat. no. M0242L).
10. RNaseOUT<sup>™</sup> Recombinant Ribonuclease Inhibitor (Thermo, cat. no. 10777019).
11. SuperScript<sup>®</sup> IV First-Strand Synthesis System (Thermo, cat. no. 18091050).
12. TruSeq Small RNA Sample Preparation Kit (Illumina, cat. no. RS-200-0012, RS-200-0024).

**2.4 Consumables**

1. Coverslips (e.g., VWR, cat. no. 631-0148).
2. Secure-Seal<sup>™</sup> Hybridization Chambers (EMS, cat. no. 70333-10).
3. Microcentrifuge tubes 0.5 ml (e.g., Eppendorf RNA/DNA LoBind).
4. Microcentrifuge tubes 1.5 ml (e.g., Eppendorf RNA/DNA LoBind).
5. Sapphire Filter tips, low retention (Greiner Bio-One, cat. no. 771265, 773265, 738265, 750265).
6. Cell scrapers.
7. Qubit<sup>®</sup> dsDNA HS Assay Kit (Thermo, cat. no. Q32851).
8. High Sensitivity DNA Kit (Agilent, cat. no. 5067-4626).
9. RNA 6000 Pico Kit (Agilent, cat. no. 5067-1513).



## 2.5 Equipment

1. Cell counter (e.g., Countess II FL Automated Cell Counter, Thermo).
2. Cooling incubator (e.g., Binder incubator, Model KB 53).
3. Tabletop centrifuge.
4. Thermoshaker (e.g., Eppendorf® Thermomixer Compact).
5. PCR cycler.
6. Sonication device (e.g., Bioruptor® Plus, Diagenode, cat. no. B01020001).
7. DynaMag™-2 Magnet (Thermo, cat. no. 12321D).
8. Qubit® 2.0 Fluorometer (Thermo, cat. no. Q32866).
9. Bioanalyzer 2100 (Agilent, cat. no. G2943CA).

---

## 3 Methods

In both BLESS and BLISS, suspension or adherent cells are grown inside cell culture dishes or flasks and, depending on the experiment, they are treated with the desired DSB-inducing protocol (for example, a DSB-inducing drug or gene manipulation using siRNA, zinc-finger nucleases, TALENs, or CRISPR). In BLISS, adherent cells are typically grown and treated on coverslips placed in a cell culture dish, while suspension cells are grown and treated in cell culture flasks, and then spotted onto PLL-coated coverslips before fixation. For BLISS in large numbers of samples (for example, to assess the specificity of multiple CRISPR guide RNAs) cells are grown, treated, and processed directly inside multi-well plates.

### 3.1 BLISS Adapters

BLISS adapters are generated by annealing a forward oligo with a reverse complementary oligo, each purified by standard desalting (*see* Table 2 and **Note 2**).

1. Prepare the phosphorylation mix (volumes for 100  $\mu$ l, to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	58 $\mu$ l
(b) Upper oligo 100 $\mu$ M	10 $\mu$ l
(c) 10 $\times$ T4 PNK buffer	10 $\mu$ l
(d) ATP 10 mM	10 $\mu$ l
(e) T4 polynucleotide kinase 10 U/ $\mu$ l	2 $\mu$ l

2. Incubate the sample for 1 h at 37 °C.
3. Add 10  $\mu$ l of bottom oligo at 100  $\mu$ M and mix well.

**Table 2**  
**Sequence of adapters and primers used in BLESS and BLISS**

Oligo name	Sequence (5'–3')	Structure
BLESS P adapter <sup>a</sup>	TACTAC_CTCGAG_AGTTACGCTAGGGATAACAG GGTAATATAG_TTT/Bio/TTTT_CTATATTACCCTG TTATCCCTAGCGTAACT_CTCGAG_GTAGTA	Barcode_XhoI_ISceI ISceI_XhoI_Barcode
BLESS D adapter <sup>b</sup>	CGTCGT_CTCGAG_AGTTACGCTAGGGATAACAGG GTAATATAG_TTTTTTTT_CTATATTACCCTGTTAT CCCTAGCGTAACT_CTCGAG_ACGACG	Barcode_XhoI_ISceI ISceI_XhoI_Barcode
BLESS PCR primer P1 <sup>c</sup>	CCCTAGCGTAACT_CTCGAG_GTAGTA	ISceI_XhoI_Barcode
BLESS PCR primer P2 <sup>d</sup>	CTAGCGTAACT_CTCGAG_ACGACG	ISceI_XhoI_Barcode
BLISS adapter <sup>e</sup>	Upper oligo GCGTGATG_NNNNNNNN_GATCGTCGG ACTGTAGA ACTCTGAAC_CCCTATAGTGA GTCGTATTACCGCCTCAATCG_AA Bottom oligo CGATTGAGGCCGGTAATACGACTCA CTATAGGG_GTTCAGAGTTCTACAG TCCGACGATC_NNNNNNNN_CATCACGC	Barcode_UMI_RA5_ T7_Overhang T7_RA5_UMI_ Barcode
BLISS RA3 adapter <sup>f</sup>	TGGAATTCTCGGGTGCCAAGG	
BLISS RTP primer <sup>g</sup>	GCCTTGGCACCCGAGAATTCCA	
BLISS universal PCR primer <sup>h</sup>	AATGATACGGCGACCACCGAGATCTACACG TTCAGAGTTCTACAGTCCGA	
BLISS indexing PCR primer <sup>i</sup>	CAAGCAGAAGACGGCATACGAGAT <b>CGTGATGTGACTGGAGTTCCCT</b> TGGCACCCGAGAATTCCA	Bold:Index sequence

<sup>a</sup>Same as linker L1 linker in **Suppl. Table 5** in [12]. As discussed in **Note 2**, this adapter may be modified by substituting the I-SceI sequence with the Illumina P5 adapter sequence, removing the XhoI site, and replacing the barcode with a longer or different sequence, if multiplexing is needed. A UMI sequence may be added before the barcode to avoid problems with cluster calling during sequencing, and enable removal of PCR duplicates as in BLISS. In addition, instead of synthesizing one long self-annealing oligo, one upper and one bottom reverse complementary oligo may be annealed instead, as during the preparation of BLISS adapters.

<sup>b</sup>Same as linker L3 in **Suppl. Table 5** in [12]. The same considerations exposed above for the proximal adapter also apply for the distal adapter. However, in this case, the I-SceI sequence should be replaced with the Illumina P3 adapter sequence. Moreover, adding a UMI in this adapter is not required.

<sup>c</sup>Same as primer P1 in **Suppl. Table 5** in [12].

<sup>d</sup>Same as primer P3 in **Suppl. Table 5** in [12].

<sup>e</sup>As explained in **Note 2**, BLISS adapters are prepared by annealing two complementary oligos containing the T7 promoter sequence, the sequence of the 5' adapter (RA5) contained in Illumina's TruSeq Small RNA Sample Preparation Kit, a UMI and a sample barcode. The upper oligos terminates with an A dinucleotide overhang to prevent the formation of head-to-tail concatemers during in situ ligation. We have successfully used 8 nt and 12 nt UMIs, as well as multiple barcode sequences [17]. One barcode sequence is shown as an example.

<sup>f</sup>Same as RA5 adapter in Illumina's TruSeq Small RNA Sample Preparation Kit.

<sup>g</sup>Same as RTP primer in Illumina's TruSeq Small RNA Sample Preparation Kit.

<sup>h</sup>Same as RP1 primer in Illumina's TruSeq Small RNA Sample Preparation Kit.

<sup>i</sup>Same as RPI1 primer in Illumina's TruSeq Small RNA Sample Preparation Kit. Different RPI primers are available for multiplexing when multiple libraries need to be sequenced in the same run.

4. In a PCR thermo-cycler, perform denaturation for 5 min at 95 °C, then ramp down to 25 °C over a period of 45 min (approx. -1.55 °C/min).
5. Store the adapter at -20 °C.

### 3.2 PLL Coating of Coverslips for BLISS

1. Place the coverslips in a 10 cm dish, cover them with 5 ml PLL solution, and then gently shake the dishes for 15 min (make sure that coverslips stay covered by liquid).
2. Aspirate the solution and transfer it into a 15 ml tube (it can be then used up to three times).
3. Wash the coverslips three times with 1 × PBS at room temperature (rt) and once with ethanol (EtOH) 70%.
4. Air-dry the coverslips and proceed with cell spotting and fixation (if not used immediately, store the coverslips in EtOH 70% at 4 °C and air-dry before use).

### 3.3 Attachment and Fixation of Suspension Cells for BLISS (See Note 3)

1. Place the desired number of PLL-coated coverslips into a 24-well cell culture plate.
2. In each well, dispense a cell suspension freshly prepared in 1 × PBS (we usually spot up to  $3 \times 10^5$  cells onto a 13 mm coverslip).
3. Let the cells sediment onto the coverslip for 10 min at rt.
4. Slowly add one volume of PFA 8% in 1 × PBS equal to the volume of cell suspension added before onto the coverslip.
5. Incubate for 10 min at rt.
6. Rinse the coverslips twice with 1 × PBS at rt.
7. Store the samples in 1 × PBS at 4 °C or proceed to permeabilization (*see* Note 4).

### 3.4 Cell Permeabilization (See Note 5)

1. Incubate the samples in LB1 for 1 h at 4 °C.
2. Equilibrate the samples at rt. and then incubate in LB2 for 1 h at 37 °C.
3. Wash the samples (*see* Note 6) and proceed to in situ blunting.

### 3.5 In Situ DSB Blunting (See Note 7)

1. Equilibrate the samples once with 1 × blunting buffer.
2. Prepare the following blunting mix using the Quick Blunting™ Kit (volumes for 100 µl, to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	75 µl
(b) 10 × blunting buffer	10 µl
(c) dNTPs 1 mM	10 µl
(d) BSA 10 mg/ml	1 µl
(e) Blunting enzyme mix	4 µl

3. Incubate the samples for 1 h at rt.

**3.6 OPTIONAL:  
In Situ A-Tailing  
(See Note 8)**

1. Wash the samples twice, 5 min each (*see Note 6*).
2. Prepare the following A-tailing mix using the NEBNext dA-Tailing Module (volumes for 100  $\mu\text{l}$ , to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	84 $\mu\text{l}$
(b) NEBNext dA-tailing reaction buffer (10 $\times$ )	10 $\mu\text{l}$
(c) Klenow fragment (3' $\rightarrow$ 5' $\text{exo}^-$ )	6 $\mu\text{l}$

3. Incubate the samples for 30 min at 37  $^{\circ}\text{C}$ .

**3.7 In Situ DSB  
Ligation (See Note 9)**

1. Wash the samples twice, 5 min each (*see Note 6*).
2. Equilibrate the samples once with 1 $\times$  T4 ligase buffer.
3. Prepare the following ligation mix (volumes for 100  $\mu\text{l}$ , to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	75 $\mu\text{l}$
(b) T4 ligase buffer	10 $\mu\text{l}$
(c) ATP 10 mM	8 $\mu\text{l}$
(d) BSA 50 mg/ml	2 $\mu\text{l}$
(e) Proximal adapter 10 $\mu\text{M}$ ( <i>see Table 2 and Note 2</i> )	4 $\mu\text{l}$
(f) T4 ligase 2000 U/ $\mu\text{l}$	1 $\mu\text{l}$

4. Incubate the samples for 16–18 h at 16  $^{\circ}\text{C}$ .

**3.8 Removal  
of Unligated Adapters  
(See Note 10)**

1. Wash the samples with HSW for three times, 1 h each, at 37  $^{\circ}\text{C}$ .
2. Quickly rinse the samples with 1 $\times$  PBS at rt.

**3.9 Extraction  
and Purification  
of Genomic DNA  
(See Note 11)**

1. Extract and purify genomic DNA either using silica-based columns or standard methods based on Proteinase K and alcohol extraction.
2. Dissolve the purified DNA in nuclease-free TE buffer.
3. Measure DNA concentration by Nanodrop<sup>TM</sup> or Qubit<sup>®</sup>, and dilute DNA in TE buffer to the concentration recommended for the sonication instrument used.

**3.10 DNA  
Fragmentation  
by Sonication  
(See Note 12)**

1. Prepare samples for sonication according to the manufacturer's instructions for the system used.
2. Sonicate DNA aiming to achieve a mean fragment size of 300–500 bp. An example of settings that we routinely use is shown in Table 3.

**Table 3**  
**Sonication options for DNA shearing in BLESS and BLISS**

Instrument	Sample	Target size (bp)	Settings	Volume, buffer
Bioruptor® Plus	Cells	350	30 s. on/90 s. off, high-mode, 20 cycles	100 µl, TE
	Tissue	350	30 s. on/90 s. off, high-mode, 40 cycles	100 µl, TE
Covaris S-series	Cells	350	Duty 10%, intensity 4, time 30 s., 4 cycle/ burst 200	50 µl, TE

- Depending on the application, concentrate the sonicated DNA using AMPure XP beads or silica-based columns.
  - Checkpoint: check DNA fragment size by agarose gel electrophoresis or Bioanalyzer.

**3.11 Affinity Capture of DSBs in BLESS (See Note 13)**

- Aliquot 5 µl of Dynabeads® MyOne™ Streptavidin C1 suspension into a 1.5 ml tube, and wash the beads two times with 500 µl of HSW.
- Dispense a volume of sonicated genomic DNA corresponding to 20 µg, and then bring the volume up to 600 µl with HSW buffer freshly supplemented with Triton X-100 0.1%.
- Incubate the samples for 30 min at 4 °C, rotating.
- Wash the beads three times, each with 600 µl of HSW buffer freshly supplemented with Triton X-100 0.1%.
- Resuspend the samples in nuclease-free water.
- Store the samples on ice before proceeding to the ligation of the distal adapter.

**3.12 Ligation of Distal Adapters in BLESS (See Note 14)**

- Prepare the following blunting mix using the Quick Blunting™ Kit (volumes for 100 µl, to be adjusted proportionally depending on the sample volume):

(a) Beads suspension	75 µl
(b) 10× blunting buffer	5 µl
(c) dNTPs 1 mM	5 µl
(d) BSA 10 mg/ml	1 µl
(e) Blunting enzyme mix	4 µl

- Incubate for 1 h at rt.
- Wash the samples twice in 600 µl of HSW buffer freshly supplemented with Triton X-100 0.1%.

4. Resuspend the beads in the following ligation mix (volumes for 100  $\mu$ l, to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	75 $\mu$ l
(b) T4 ligase buffer	10 $\mu$ l
(c) ATP 10 mM	8 $\mu$ l
(d) BSA 50 mg/ml	2 $\mu$ l
(e) D adapter 10 $\mu$ M (Table 2)	4 $\mu$ l
(f) T4 ligase 2000 U/ $\mu$ l	1 $\mu$ l

5. Incubate the samples for 16–18 h at 16  $^{\circ}$ C.

### 3.13 Removal of Adapter Loops in BLESS (See Note 15)

1. Wash the beads three times, each with 600  $\mu$ l of HSW buffer freshly supplemented with Triton X-100 0.1%.
2. Resuspend the beads in the following mix (volumes for 100  $\mu$ l, to be adjusted proportionally depending on the sample volume):

(a) Nuclease-free water	85 $\mu$ l
(b) 10 $\times$ I-SceI buffer	10 $\mu$ l
(c) BSA 10 mg/ml	1 $\mu$ l
(d) I-SceI	4 $\mu$ l

3. Incubate for 4 h at 37  $^{\circ}$ C.
4. Spin the beads for 5 min at 15–20,000  $\times g$ .
5. Transfer as much supernatant as possible to a new 1.5 ml tube.
6. Store the samples at –20  $^{\circ}$ C if not immediately used for PCR.
1. For each sample, prepare the following PCR mix (volumes for 50  $\mu$ l, to be adjusted proportionally, depending on the sample volume):

### 3.14 PCR Amplification of Captured DSBs in BLESS (See Note 16)

(a) I-SceI digested DNA	23 $\mu$ l
(b) Primer P1 and P2 10 $\mu$ M (Table 2)	1 $\mu$ l each
(c) NEBNext <sup>®</sup> high-Fidelity 2 $\times$ PCR master mix	25 $\mu$ l

2. In a fast-cycling PCR thermo-cycler perform the following steps:
  - (a) 98  $^{\circ}$ C, 30 s.
  - (b) 98  $^{\circ}$ C, 10 s.
  - (c) 60  $^{\circ}$ C, 30 s.

- (d) 72 °C, 30 s.
  - (e) 4 °C, pause.
3. Repeat **steps (b)–(d)** for 18–20 times depending on the initial sample's size.
  4. Purify PCR products using silica-based columns or alcohol precipitation.
    - Checkpoint: check PCR product size by agarose gel electrophoresis or Bioanalyzer.

**3.15 Adapter Cleavage and Library Preparation in BLESS (See Note 17)**

1. Digest the purified PCR products in the following mix (volumes for 100 µl, to be adjusted proportionally depending on the sample volume):

(a) Purified PCR	85 µl
(b) 10× CutSmart® buffer	10 µl
(c) XhoI 10 U/µl	5 µl

2. Incubate the samples 16–18 h at 37 °C.
3. Purify the PCR products using silica-based columns or alcohol precipitation.
4. Measure DNA concentration and prepare a sequencing library using the NEBNext DNA Library Prep Master Mix Set for Illumina, according to the manufacturer's instructions.

**3.16 Selective Linear Amplification of DSBs in BLISS (See Note 18)**

1. After sonication, concentrate the samples using AMPure XP beads and elute DNA in 10 µl of nuclease-free water.
  - Checkpoint: run 1 µl of sample on Bioanalyzer using a High Sensitivity DNA chip.
2. Prepare the following IVT mix using the MEGAscript® T7 Transcription Kit:

(a) Purified DNA	7.5 µl
(b) rNTPs (premixed in equal volumes)	8 µl
(c) 10× T7 polymerase buffer	2 µl
(d) T7 RNA polymerase	2 µl
(e) RiboSafe RNase inhibitor 40 U/µl	0.5 µl

3. Incubate for 14–16 h at 37 °C.

**3.17 Removal of Genomic DNA and Amplified RNA (aRNA) Purification in BLISS (See Note 19)**

1. To each sample add 1 µl of DNaseI.
2. Incubate for 15 min at 37 °C.
3. Purify the aRNA using Agencourt RNAClean XP beads or silica-based columns.

4. Elute the aRNA in 6  $\mu\text{l}$  nuclease-free water.
  - Checkpoint: run 1  $\mu\text{l}$  of sample on Bioanalyzer using a RNA 6000 Pico chip.

### 3.18 Library

#### Preparation in BLISS

(See Note 20)

1. 3' Illumina adapter ligation.
  - (a) On ice, dilute the RA3 adapter 1:5 in nuclease-free water.
  - (b) Add 1  $\mu\text{l}$  of diluted RA3 to 5  $\mu\text{l}$  of purified aRNA.
  - (c) In a PCR thermo-cycler, incubate the sample for 2 min at 70 °C, and then immediately place it on ice.
  - (d) Add 4  $\mu\text{l}$  of the following mix:

10 $\times$ RNA ligase buffer	2 $\mu\text{l}$
RNaseOUT™	1 $\mu\text{l}$
T4 RNA ligase truncated	1 $\mu\text{l}$

- (e) In a PCR thermo-cycler with the lid open, incubate the sample for 1 h at 28 °C.
  - (f) Place the sample on ice, and add 3.5  $\mu\text{l}$  of nuclease-free water.
2. Reverse transcription (1<sup>st</sup> strand synthesis).

- (a) On ice, prepare a dilution of dNTPs at 12.5 mM in nuclease-free water.
- (b) Add 1  $\mu\text{l}$  of RTP primer to the RNA sample.
- (c) In a PCR thermo-cycler, incubate the sample for 2 min at 70 °C and then immediately bring back the sample on ice.
- (d) Add 5.5  $\mu\text{l}$  of the following mix prepared using the SuperScript IV First-Strand Synthesis System:

1 <sup>st</sup> strand buffer 5 $\times$	2 $\mu\text{l}$
dNTPs 12.5 mM	0.5 $\mu\text{l}$
DTT 100 mM	1 $\mu\text{l}$
RNaseOUT™	1 $\mu\text{l}$
SuperScript® IV	1 $\mu\text{l}$

- (e) In a PCR thermo-cycler with the lid at 50 °C, incubate the sample for 60 min at 50 °C.
3. Library indexing and amplification.
  - (a) Transfer 20  $\mu\text{l}$  of cDNA sample into a PCR tube on ice.
  - (b) Add 2  $\mu\text{l}$  of the desired RPI indexing primer (Table 2), and then 28  $\mu\text{l}$  of the following mix:

Nuclease-free water	1 $\mu\text{l}$
RPI primer	2 $\mu\text{l}$
NEBNext 2 $\times$ PCR master mix	25 $\mu\text{l}$



(c) In a fast-cycling PCR thermo-cycler perform the following steps:

98 °C, 30 s.

98 °C, 10 s.

60 °C, 30 s.

72 °C, 30 s.

4 °C, pause.

(d) Repeat **steps (b)–(d)** for 14–16 times depending on the initial sample's size.

#### 4. Library purification.

(a) Transfer the PCR sample into a 0.5 ml LoBind tube.

(b) Add 40 µl of AMPure XP bead suspension (pre-warmed at rt. for 30 min), and then mix thoroughly by pipetting up-down five to six times.

(c) Incubate the sample for 7 min at rt.

(d) Place the sample onto a magnetic stand and wait 2–3 min until all the beads have attached to the magnet.

(e) With a 200 µl pipette, aspirate the supernatant.

(f) Wash the beads once with 200 µl of ice-cold EtOH 80%, with the sample still on the magnetic stand.

(g) Aspirate the supernatant.

(h) Repeat the wash once.

(i) Air-dry the beads for 7 min at rt.

(j) Remove the sample from the magnetic stand, and resuspend the beads in 20 µl of nuclease-free water.

(k) Incubate for 5 min at rt.

(l) Place the sample on the magnetic stand, and wait 2–3 min until all the beads have attached to the magnet.

(m) Transfer 18 µl of the cleared solution into a 1.5 ml LoBind tube.

(n) Store the library at –20 °C.

- Checkpoint: run 1 µl of sample on Bioanalyzer using a High Sensitivity DNA chip.

---

## 4 Notes

1. **Using Proteinase K to purify nuclei.** This step requires optimization for every cell type in order to avoid over-digestion. For example, in our original paper we found that incubation of HeLa cells for 7–8 min at 37 °C in the presence of Proteinase K yields purified nuclei that retain their morphology and form a visible pellet upon centrifugation. Other cell types (such as

U2OS and mouse lymphocytes) required shorter incubation times (< 5 min). However, since this renders downstream sample manipulation challenging (the nuclei pellet becomes almost transparent after Proteinase K digestion), and since we have found that it does not add a significant benefit (unpublished observations), this step can be safely omitted in BLESS. We also note that in BLISS, since permeabilization and in situ reactions are performed in cells attached onto a microscope slide or coverslip, incubation with Proteinase K is skipped in order to avoid the detachment of cells from the glass surface.

- 2. Design and preparation of BLESS and BLISS adapters.** A scheme of BLESS and BLISS adapters is shown in Table 2. BLESS adapters consist of single, self-annealing oligos that form a hairpin-like structure to prevent the formation of head-to-tail concatemers. The proximal (P) oligo (i.e., the one that is in situ ligated to DSBs) is synthesized to have a single biotin group attached to the loop and is purified using HPLC, while the distal (D) oligo is purified by standard desalting. Recently, a modified adapter design consisting of a hairpin-forming oligo conjugated with two biotin moieties was used in END-seq [16]. Self-annealing is carried out by incubating the oligo at a concentration of 10  $\mu$ M in 1 $\times$  T4 ligase buffer for 5 min at 95  $^{\circ}$ C (typically, in a PCR thermo-cycler), followed by immediate cooling down by placing the sample on ice. In the original design, the proximal BLESS adapter contained a barcode sequence tagging the site of ligation, an XhoI cutting site, and an I-SceI recognition site [12]. The latter was used to release the captured biotinylated DSB fragments from streptavidin beads, while XhoI was used to remove the I-SceI sequence before the final library preparation step. A disadvantage of this design is that, after XhoI cleavage, the remaining portion of the proximal adapter (i.e., where the DSB was ligated) can be ligated to either the 5' or the 3' Illumina adapter. Paired-end sequencing is then preferably used in order to maximize the number of DSB ends detected given a certain sequencing depth. This limitation was recently mitigated in DSBCapture [15] using an improved BLESS adapter design. In DSBCapture Y-shaped adapters similar to the ones routinely used DNA library preparation for whole-genome sequencing is used instead, enabling sequencing of only the DSB ends by single-end sequencing. In contrast to BLESS, where the proximal adapter is used to selectively capture the genomic DNA sequence surrounding the DSBs using affinity purification on streptavidin beads, BLISS adapters have been designed to enable selective DSB amplification using in vitro transcription by the T7 RNA polymerase/promoter system. In addition to the T7 promoter sequence, another special feature of BLISS

over BLESS adapters is that they contain a sequence of random nucleotides (so far we have successfully used adaptors containing either 8 or 12 random bases) that is used as a UMI during processing of sequencing reads. In addition, the adapter contains the same RA5 adapter sequence included in Illumina TruSeq Small RNA library preparation kit, therefore making single-end sequencing sufficient to identify the original in situ ligation location. BLISS adapters are obtained by annealing a forward oligo with a reverse complementary oligo, each purified by standard desalting (Table 2).

3. **Attachment of cells and fixation.** In BLESS, adherent cell cultures are first trypsinized and washed in PBS, and then fixation is performed in suspension inside 15 ml tubes (for up to  $10^7$  cells) rotating at room temperature. In the original BLESS protocol, cells are fixed in formaldehyde 2% in culture medium for 30 min at rt., followed by a quenching step of 5 min in glycine 125 mM. However, we have now shortened the fixation procedure and only use methanol-free PFA instead of formaldehyde. Fixation in PFA 4% for 10 min at rt. is sufficient for both BLESS and BLISS to stabilize chromatin while enabling DSB accessibility for in situ reactions. For BLISS in tissues, we have developed two approaches: (1) Tissue cryopreservation and sectioning. Freshly extracted tissue biopsies are first fixed in PFA 4% for 1 h at rt., followed by three washes in  $1 \times$  PBS, immersion first in sucrose 15% and then in sucrose 30% until the tissue sinks. Biopsies are finally embedded in OCT medium and quickly frozen at  $-20^\circ\text{C}$  before sectioning. OCT blocks are cryosectioned into 30  $\mu\text{m}$ -thick slices, which are then mounted onto microscope slides, dried for 60 min at rt., and then stored at  $4^\circ\text{C}$  before further processing. To reduce reagent use and facilitate sample processing, all steps from lysis until in situ ligation are performed inside Secure-Seal™ Hybridization Chambers mounted onto the slide. (2) Preparation of nuclei suspensions (adapted from [26]). In the second approach, fresh tissue biopsies are cut into small pieces and transferred into a volume of NSB inside a 1.5–2 ml tube (note that the volume of NSB buffer must be adjusted depending on the biopsy size in order to cover all the tissue fragments). The sample is incubated for 15–40 min until the tissue fragments become transparent. At this point, examining a small volume of solution under a bright-field microscope shall reveal the presence of nuclei in suspension (DAPI staining and examination under a fluorescence microscope can also be done to further assess the quality of the nuclei suspension). Nuclei are then pelleted by centrifugation for 5 min at  $500 \times g$  and resuspended in 200–500  $\mu\text{l}$  of  $1 \times$  PBS. 100  $\mu\text{l}$  of nuclei suspension are dispensed onto a 13 mm PLL-coated coverslip and

incubated for 10 min at rt. 100  $\mu$ l of PFA 8% 1 $\times$  PBS are then gently added and incubated for 10 min at rt., followed by 2 washes in 1 $\times$  PBS at rt. The samples can be stored in 1 $\times$  PBS at 4 °C up to 1 month before performing BLISS.

4. **Storage of fixed samples.** Fixed cells can be stored (in tubes in BLESS or on coverslips in BLISS) in 1 $\times$  PBS at 4 °C for several days up to a month. To avoid bacterial contamination, NaN<sub>3</sub> can be added to the solution to 0.05% final concentration.
5. **Cell permeabilization.** In BLESS, all the permeabilization steps are typically done in rotating tubes (max 10<sup>6</sup> cells per 15 ml tube), and then the cell pellet is transferred into a 1.5–2 ml tube for washings before in situ blunting. Equilibration at rt. before shifting from buffer LB1 to LB2 is needed to avoid precipitation of SDS contained in LB2. In the original BLESS protocol cells were incubated in LB1 for 1 h and in LB2 for 45 min. However, we now routinely perform 1 h incubations in LB1 and LB2 buffers in both BLESS and BLISS. In BLISS, cells fixed on coverslips are typically permeabilized and washed inside 6- or 12-well plates, depending on their size (for 6-well plates, use at least 1.5 ml of buffer per well).
6. **Washes.** Following permeabilization and in between subsequent in situ reactions, both the BLESS and BLISS samples should be washed with a buffer that will not interfere with the next reaction, if not completely removed. In the original BLESS protocol NEBuffer 2 is used after permeabilization, digestion with Proteinase K, and in situ blunting steps. In addition, in situ blunting and ligation were each preceded by one equilibration step in 1 $\times$  blunting and 1 $\times$  T4 ligase buffer, respectively. In BLISS, NEBuffer 2 has been substituted with CutSmart<sup>®</sup> buffer, which was meanwhile introduced by New England Biolabs as a universal buffer for many of their enzymes. Moreover, the equilibration step in blunting buffer has been removed in BLISS. Washing volumes depend on the assay type and number of cells. In BLESS, a pellet containing up to 5  $\times$  10<sup>6</sup> cells is washed in a 1.5–2 ml tube with 200  $\mu$ l of buffer per 10<sup>6</sup> cells. In BLISS, approximately 3  $\times$  10<sup>5</sup> cells on a 13 mm circular coverslip can be washed inside a 6-well plate containing 2 ml of 1 $\times$  CutSmart<sup>®</sup> of buffer per well, and shaking to avoid cell drying. For BLISS in tissue sections, the washes are done by pipetting in and out a volume of buffer equal to the volume of the Secure-Seal<sup>™</sup> chamber covering each section.
7. **In situ DSB blunting.** In both BLESS and BLISS, the blunting step is needed to convert the DSBs that have been resected during homologous recombination repair as well as the breaks that lack a 5' phosphate group, into a configuration that can be ligated by a universal oligo adapter. In BLESS, a pellet containing

up to  $5 \times 10^6$  cells is incubated with 100  $\mu$ l of blunting mix. In contrast, for BLISS on coverslips, the in situ blunting procedure is performed by dispensing 50  $\mu$ l of blunting mix per sample (for a 13 mm circular coverslip) on a piece of Parafilm placed inside a cell culture dish containing a piece of wipe tissue presoaked in distilled water and wrapped all around the dish edge to prevent drying. Coverslips are then placed upside-down onto the blunting mix droplet, and the dish is sealed with Parafilm during the time of incubation. Alternatively, in situ blunting can also be performed directly in a multi-well plate (with cells either on coverslips or directly attached to the plastic) by adjusting the volumes depending on the well size. Although more costly, this approach is particularly suited when a large number of samples need to be processed in parallel (for example, when screening for off-target DSBs of multiple CRISPR guide RNAs).

8. **In situ A-tailing.** Addition of a 3' A-tail to blunted DSBs is expected to increase the efficiency of ligation by providing short sticky ends. We have performed comparison experiments using BLISS to detect Cas9-induced DSBs, and demonstrated that A-tailing increases the number of off-targets that are detected at the same sequencing depth. A-tailing step is also included in END-seq and DSBCapture [17, 18]. In BLESS, an optional A-tailing step may also be included after in situ blunting by incubating a pellet of up to  $5 \times 10^6$  nuclei with 100  $\mu$ l of A-tailing mix.
9. **In situ DSB ligation.** Similar to blunting, in situ ligation may be performed in tubes, on Parafilm or directly in multi-well plates depending on the assay type. In BLESS, a pellet containing up to  $10^6$  nuclei is incubated with 50  $\mu$ l of ligation mix. For BLISS on coverslips, the in situ ligation procedure is performed by dispensing 50  $\mu$ l of ligation mix (for a 13 mm circular coverslip) on a piece of Parafilm placed inside a cell culture dish containing a piece of wipe tissue presoaked in distilled water and wrapped all around the dish edge to prevent drying. Compared to the original BLESS protocol, in both BLESS and BLISS we now complement the ligation mix with extra BSA and ATP to increase the efficiency of in situ, similar to the ligation step in HiC chromosome conformation capture [27]. In all cases, the reaction is carried out at 16 °C overnight, with the goal of ligating as many DSBs as possible. However, we have not determined the effect of shorter incubation times and higher temperature on the efficiency of in situ ligation.
10. **Removal of unligated adapters.** In both BLESS and BLISS, removal of adapters that have not been ligated to DSBs and of head-to-head adapter dimers is critical to avoid contamination of the final library by short adapter sequences, which results in poor mappability of sequenced reads. In the original BLESS protocol, the nucleus pellet remaining after in situ ligation is

washed three times with HSW buffer containing NaCl 1 M, by pipetting the nucleus suspension (200  $\mu$ l of HSW buffer per  $10^6$  nuclei) thoroughly up and down for five to six times at room temperature. In BLISS, washing cells on coverslips or in culture plates with the same NaCl concentration and duration is insufficient, likely because many adapters stick to the positively charged surface of glass and plastic (especially when PLL coating is used). Instead, cells are thoroughly washed over a period of 3 h in HSW buffer containing NaCl 2 M at 37 °C, while shaking (at least three washes, 1 h each). This should remove most of the unligated adapters and adapter dimers. However, the washing efficiency needs to be determined empirically for each surface and coating type. For BLISS in tissue sections, after the in situ ligation step, the hybridization chamber covering the tissue is removed, and the slide is washed with HSW in a washing jar.

11. **Extraction and purification of genomic DNA.** Once DSBs have been ligated in situ, genomic DNA can be extracted even though the extraction procedure causes artificial breaks. In the original BLESS protocol, DNA is extracted by digesting washed nuclei in the presence of Proteinase K followed by the precipitation of genomic DNA with isopropanol. Alternatively, commercial kits for genomic DNA extraction can be used. In BLISS, since cells are adherent onto a solid surface, they have to be manually scraped off and transferred into tubes for convenient downstream processing. Typically, coverslips are quickly rinsed with nuclease-free water, air-dried, and transferred onto a piece of Parafilm with cells facing upward. For a coverslip of 13 mm, a digestion mix containing 5  $\mu$ l of Proteinase K in 100  $\mu$ l of digestion buffer (SDS 1%, NaCl 100 mM, EDTA 50 mM, Tris-HCl 10 mM, pH 8) is deposited onto it, and the cells are manually scraped off using a sterile cell scraper. The solution is then transferred into a 1.5–2 ml tube and incubated for 16–18 h at 55 °C in a thermo-shaker. For samples consisting of a few thousand cells, it is advisable to use Lo-bind tubes to reduce DNA adsorption to the tube's walls. DNA can then be extracted with phenol-chloroform and precipitated with isopropanol or ethanol in the presence of glycogen (0.05–1  $\mu$ g/ $\mu$ l final conc.), or column-based purification kits. When cells are grown directly in multi-well plates, the digestion mix is dispensed and the cells are scraped directly inside each well. In both BLESS and BLISS, purified genomic DNA should be dissolved in nuclease-free TE buffer.
12. **DNA sonication.** In both BLESS and BLISS, genomic DNA must be fragmented to prevent biases due to different fragment size during enrichment of the sequences surrounding DSBs by affinity purification (in BLESS) or IVT (in BLISS). In the

original BLESS protocol, we had to apply both enzymatic digestion with HaeIII and sonication, since with the available sonication device we could not generate reproducible fragmentation patterns. However, this approach is suboptimal, not only because more steps and reagents are needed, but also because the genomic distribution of HaeIII cut sites can interfere with the detection of DSBs at locations with high density of HaeIII cut sites (for example, inside gene promoter regions) or where HaeIII sites and DSBs would be too close (for instance, in CRISPR applications HaeIII cannot be used since the recognition site CCGG also identifies the PAM sequence to which the endonuclease binds). With the widespread diffusion of programmable sonication devices such as Bioruptor<sup>®</sup> Plus and Covaris, we have now abandoned the use of HaeIII both in BLESS and BLISS, and only apply sonication. Table 3 shows examples of sonication settings that we routinely use. Following sonication, the sample is concentrated to allow the highest amount possible of DNA to be used as input for streptavidin capture in BLESS or IVT in BLISS (*see* Notes 14 and 19).

13. **Affinity capture of DSBs on streptavidin beads.** In BLESS, streptavidin beads are used to enrich for the genomic sequences surrounding the DSBs that were previously in situ ligated with biotinylated adapters. During the development of the method, we tested different types of commercially available streptavidin beads (Dynabeads<sup>®</sup> M-270, M-280, MyOne<sup>™</sup> C1), and found that C1 beads were very specific and had low unspecific binding to non-biotinylated DNA. Per 5  $\mu$ l of beads, we typically use 20  $\mu$ g of genomic DNA as input (corresponding to approx. 3.3 million diploid genome equivalents), but this amount can be lowered depending on initial sample size.
14. **Ligation of distal adapters in BLESS.** In the original BLESS protocol, a distal adapter (D, *see* Table 2) is ligated to the free extremity of captured DSB fragments attached onto streptavidin beads. However, as already discussed in Note 2, this procedure requires downstream paired-end sequencing to maximize the number of ligated DSB ends that can be retrieved. We now recommend the Illumina 3' adapter sequence to be incorporated in the D adapter by replacing the I-SceI sequence with it.
15. **Removal of adapter loops in BLESS.** In the original BLESS protocol, after distal adapter ligation DSB fragments are released from streptavidin beads using the I-SceI homing endonuclease before PCR is performed (*see* Note 17). Alternatively, PCR can be performed directly on the beads, even though based on our experience this procedure is less efficient.
16. **PCR amplification of captured DSBs in BLESS.** In BLESS, DSB fragments captured by streptavidin must be amplified to

yield a quantity of DNA sufficient to prepare a sequencing library. In the original protocol, we subjected samples derived from 5–10 million cells to 18–20 cycles of amplification in order for the PCR product to be detectable by agarose gel electrophoresis and ethidium bromide staining. The number of cycles can be reduced to minimize the amount of PCR duplicates, by using more sensitive assays such as Agilent's Bioanalyzer. However, the optimal number of cycles will still vary depending on the sample type and amount, and needs to be empirically adjusted.

17. **Adapter cleavage and library preparation in BLESS.** In the original BLESS protocol PCR-amplified DSB fragments are digested with XhoI to cleave off both the proximal and distal adapter sequences (with the exception of barcodes marking the proximal and distal ligation position), and the resulting DNA is finally used to prepare a sequencing library. However, this poses two major limitations. First, the enzyme may cut inside certain fragments, producing DNA fragments that only contain either the proximal or distal barcode sequence (this will reduce the number of DSB ends effectively identified at a given sequencing depth). Second, as already discussed in **Note 2**, since sequencing adapters are introduced after PCR, the original directionality of ligation is not preserved, and paired-end sequencing must be performed to maximize the number of proximal adapter ligation sites that are sequenced. This increases costs and renders the procedure very time-consuming. Thus, as recommended above, the original BLESS adapter design should be modified to include the P5 and P3 Illumina adapter sequences.
18. **Selective linear amplification of DSBs in BLISS.** In BLISS, the DNA sequence surrounding the labeled DSBs is selectively amplified using IVT driven by the T7 RNA polymerase. Independently of the amount of input, we usually perform IVT for 14–16 h, although the incubation time may be shortened if the input is large enough ( $> 10^5$  genome equivalents). For multiplexing, DNA purified from multiple samples that have been in situ labeled with different barcodes can be premixed and used as input for a single IVT reaction. Depending on the concentration of input DNA and the level of multiplexing, the final IVT reaction volume may be adjusted by proportionally scaling all the volumes.
19. **Removal of genomic DNA and purification of amplified RNA (aRNA) in BLISS.** After IVT, genomic DNA is digested with DNaseI to avoid carry-over of DNA into the final library. However, even without this step we have managed to obtain high-quality libraries, and thus we recommend it only in the case of high DNA inputs to the IVT reaction ( $> 10^5$  genome equivalents). Purification of aRNA generated by IVT can be done using



Agencourt RNAClean XP beads or silica-based columns (for example, RNeasy<sup>®</sup> MinElute<sup>®</sup> Cleanup kit, Qiagen, cat. no. 74204).

20. **Library preparation in BLISS.** A custom-modified protocol of Illumina TruSeq Small RNA Sample Preparation Kit is used to prepare BLISS libraries. The protocol has been adapted to account for low input as well as for the fact that the RA5 adapter sequence is already present in the aRNA. Purchase of the kit is not necessary, since all the needed components (RA3 adapter and primers, *see* Table 2) can be purchased separately (for example, we used Integrated DNA Technologies Inc., and purify the RA3 adapter by HPLC, and RTP, RPI, and RPI primers by standard desalting). T4 RNA ligase, RNaseOUT<sup>™</sup>, SuperScript<sup>®</sup> IV are not part of the TruSeq Small RNA Sample Preparation Kit and are purchased separately (*see* **Reagents**). As in BLESS, the optimal number of PCR cycles depends on the input size (initial amount of cells or number of genome equivalents used as input in the IVT reaction) and expected number of DSBs per cell. Typically, for samples of  $10^3$ – $10^4$  cells we perform 12–16 amplification cycles, but this number needs to be empirically adjusted to limit the number of PCR duplicates.

---

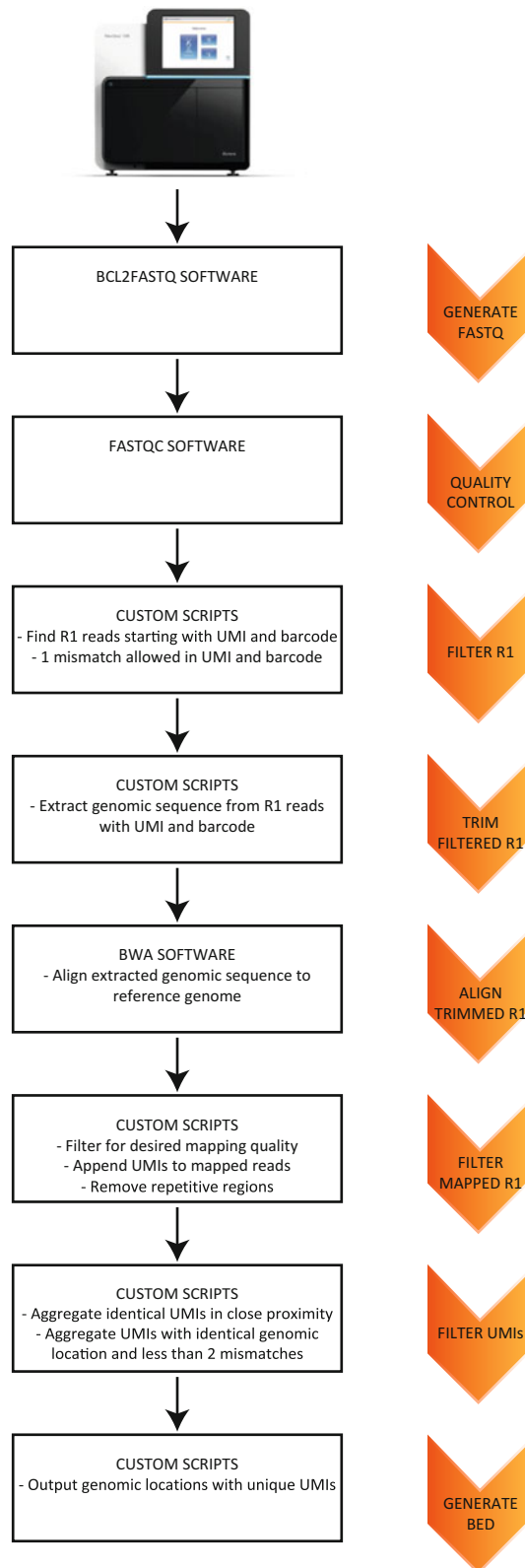
## 5 Additional Considerations

### 5.1 Choice of Sequencing Platform and Sequencing Depth

Both the BLESS and BLISS protocols described here terminate with the preparation of sequencing libraries that can be sequenced on several Illumina platforms, including MiSeq, NextSeq, and HiSeq (Fig. 1). In the original BLESS paper [12], we also used the Roche 454 system, but since this platform is no longer commercially available and because of the widespread availability of Illumina platforms, we now only use the latter. The platform choice depends on the desired sequencing throughput, which ultimately will be dictated by the sample size (number of cells or genome equivalents) and by the number of DSBs per cell. Based on our experience, for a sample of approx.  $10^5$  human or mouse cells processed by BLISS, a sequencing depth of 40–50 million reads per sample is typically sufficient to identify genomic regions spanning 50–100 kb with significant enrichment in DSBs probability, by comparing a sample treated with a DSBs-inducing condition with a control. Certain applications, such as identifying rare CRISPR off-target events, might require a higher depth, and in general the optimal sequencing depth must be empirically determined based on the resolution needed.

### 5.2 Data Analysis

There is currently no ready-to-use software package for the analysis of BLESS and BLISS data, and pipelines similar to the one which we have developed for BLISS (Fig. 2) must be developed in house.



**Fig. 2** Sequence reads processing workflow in BLISS. The BCL file obtained after sequencing on an Illumina platform (MiSeq, NextSeq or HiSeq) is first converted into a FASTQ file and checked for quality using available

Following read filtering and alignment to the reference genome, genomic regions with higher breaking probability (compared to the rest of the genome or to the same region in a different sample) can be identified using different approaches. In the original BLESS paper [12], we applied the hypergeometric probability distribution to compute the statistical significance of differences in DSBs counts between corresponding sliding genomic windows in a control sample versus a sample treated with aphidicolin to induce replication stress. However, depending on the size of the sliding window, this approach is computationally intensive. Furthermore, since additional methods such as the Bonferroni method [28] must be implemented to correct for multiple hypothesis testing, this approach is rather time-consuming. In alternative, nearest neighbor clustering algorithms can be used to identify DSBs hotspots. However, while this approach can robustly identify DSBs clusters of a given expected size and shape, such as those induced by CRISPR/Cas9 [13, 14], tailoring it to identify clusters of mixed DSBs forming through different processes (e.g., replication stress, transcription, etc.) and repaired by different mechanisms (NHEJ vs. homologous recombination) is not trivial. Peak calling algorithms can also be applied, although they were originally developed for ChIP-seq analysis and are not tailored for the type of data obtained with BLESS or BLISS. For example, MACS2 [29] was recently used to analyze DSBapture data, while HOMER [30] and SICER [31] were used in END-seq. Finally, a key advantage in BLISS data analysis is that, thanks to the use of UMIs, single nucleotide locations that break recurrently in multiple cells can be robustly identified without the need for clustering algorithms or peak calling tools. Indeed, simply by filtering genomic locations based on the number of UMIs associated with each location, we have been able to identify etoposide-induced DSBs hotspots as well as recurrent breaking sites in untreated cells [17].



**Fig. 2** (continued) software. Afterward, a series of custom scripts written in Unix are used to select the R1 reads that start with the correct pattern of 8 nt UMI followed by the proper barcode. At this point, the genomic sequence downstream of the barcode is aligned to a reference genome using BWA. Afterward, again using custom scripts in Unix, the original UMI is associated with the corresponding aligned read (after removing reads mapping to repetitive regions), and a filter is applied to remove reads mapped to the same or very near genomic location and having the same UMI (IVT and PCR duplicates). Finally, a BED file is constructed, containing information about the genomic coordinate of each DSB event labeled by a unique UMI tag

## References

1. Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. *Nature* 461:1071–1078
2. van Gent DC, Hoeijmakers JH, Kanaar R (2001) Chromosomal stability and the DNA double-strandbreak connection. *Nat Rev Genet* 2:196–206
3. McKinnon PJ, Caldecott KW (2007) DNA strand break repair and human genetic disease. *Annu Rev Genomics Hum Genet* 8:37–55
4. Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. *Nat Biotechnol* 23:967–973
5. Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* 14:49–55
6. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157:1262–1278
7. Zetsche B, Gootenberg JS, Abudayyeh OO et al (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163:759–771
8. Tsai SQ, Keith Joung J (2016) Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* 17:300–312
9. Szilard RK, Jacques P-E, Laramée L et al (2010) Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. *Nat Struct Mol Biol* 17:299–305
10. Iacovoni JS, Caron P, Lassadi I et al (2010) High-resolution profiling of gamma-H2AX around DNA double strand breaks in the mammalian genome. *EMBO J* 29:1446–1457
11. Madabhushi R, Gao F, Pfenning AR et al (2015) Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* 161:1592–1605
12. Crosetto N, Mitra A, Silva MJ et al (2013) Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods* 10:361–365
13. Ran FA, Cong L, Yan WX et al (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520:186–191
14. Slaymaker IM, Gao L, Zetsche B et al (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
15. Lensing SV, Marsico G, Hänsel-Hertsch R et al (2016) DSBcapture: in situ capture and sequencing of DNA breaks. *Nat Methods* 13:855–857
16. Canela A, Sridharan S, Sciascia N et al (2016) DNA breaks and end resection measured genome-wide by end sequencing. *Mol Cell* 63:898–911
17. Yan WX, Mirzazadeh R, Garnerone S et al (2017) BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat Commun* 8:15058
18. Kim D, Bae S, Park J et al (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* 12:237–243. 1 p following 243
19. Kim D, Kim J, Hur JK et al (2016) Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol* 34:863–868
20. Tsai SQ, Zheng Z, Nguyen NT et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33:187–197
21. Kleinstiver BP, Tsai SQ, Prew MS et al (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat Biotechnol* 34:869–874
22. Wang X, Wang Y, Wu X et al (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 33:175–178
23. Frock RL, Hu J, Meyers RM et al (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* 33:179–186
24. Wei P-C, Chang AN, Kao J et al (2016) Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. *Cell* 164:644–655
25. Kivioja T, Vähärautio A, Karlsson K et al (2012) Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods* 9:72–74
26. Baslan T, Kendall J, Rodgers L et al (2012) Genome-wide copy number analysis of single cells. *Nat Protoc* 7:1024–1041
27. Naumova N, Imakaev M, Fudenberg G et al (2013) Organization of the mitotic chromosome. *Science* 342:948–953
28. Dunnett CW (1955) A multiple comparison procedure for comparing several treatments with a control. *J Am Stat Assoc* 50:1096–1121
29. Zhang Y, Liu T, Meyer CA et al (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137

30. Heinz S, Benner C, Spann N et al (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38:576–589
31. Zang C, Schones DE, Zeng C et al (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* 25:1952–1958

## DNA Replication Profiling Using Deep Sequencing

Xanita Saayman, Cristina Ramos-Pérez, and Grant W. Brown

### Abstract

Profiling of DNA replication during progression through S phase allows a quantitative snap-shot of replication origin usage and DNA replication fork progression. We present a method for using deep sequencing data to profile DNA replication in *S. cerevisiae*.

**Key words** DNA replication, Replication forks, Deep sequencing, Replication fork rate, Replication origins

---

### 1 Introduction

DNA replication in eukaryotes is initiated at multiple origins of replications in a highly defined two-step mechanism [1–4]. Briefly, origins are “licensed” in G1 by the formation of pre-replicative (pre-RC) protein complexes. DNA synthesis is then initiated as a subset of these licensed origins are “fired” in a defined temporal order throughout S phase [5]. The resulting replication forks travel bidirectionally from each origin until encountering newly replicated regions emanating from adjacent replication origins, or until the replisome dissociates from the DNA strands. The latter case is typically a result of aberrant DNA structures or limiting replication factors under replication stress conditions.

Due to the inherently mutagenic nature of DNA replication, the rate and timing at which replication is completed can influence somatic mutation rates in both yeast and human cells [6, 7]. This implies that temporal regulation of DNA replication is important in preserving genomic integrity. Global replication rates are a function of both origin firing patterns and fork progression rate. While it is evident that eukaryotes show remarkable plasticity in both parameters in response to conditions such as DNA damage [8],

---

Xanita Saayman and Cristina Ramos-Pérez contributed equally to this work.

relatively little is known about how origin activation and fork progression are temporally regulated [4, 9].

Previous techniques for assessing replication dynamics include density transfer methods, flow cytometry, chromatin-immunoprecipitation (ChIP) of replication proteins or nucleotide analogs, single-stranded DNA accumulation, and single molecule DNA fiber analysis [10–14]. Additionally, high-density oligonucleotide microarrays used in DNA “copy number variation” (CNV) procedures have been adapted to map replication origins and monitor fork progression [5, 15]. However, microarray-based CNV introduces intrinsic resolution constraints due to probe size, and signal disruption due to nonlinear hybridization signal artifacts [16]. CNV by deep sequencing offers an increasingly accessible alternative for investigating genome-wide replication patterns on a population level, and can generate replication profiles with remarkable spatial and temporal resolution [17].

Here, we describe a method for generating DNA replication profiles with CNV analysis from deep sequencing of synchronous *S. cerevisiae* cultures to investigate replication features (e.g., origin firing, replication fork progression). To this end, logarithmically growing cultures are arrested in G<sub>1</sub> with mating pheromone and synchronously released into S phase in the presence of DNA damaging agents. A variety of alternative cell synchronization techniques can be applied including centrifugal elutriation, conditional mutants, or drug treatment [18, 19]. Flow cytometric analysis is used to assess synchronization quality and to determine median sample DNA content for downstream normalization. Genomic DNA is extracted from samples taken at different times after release from G<sub>1</sub>, and sequenced with Illumina technology.

We also present a computational method adapted from Muller et al. [17] for data extraction from sequencing reads, data processing and normalization, and replication profile visualization. Unique reads are first aligned to a reference *S. cerevisiae* genome and scaled according to the sample DNA content as determined by peak-aligned flow cytometry profiles. This normalization technique is applied to establish a baseline of 1.0 (as described in [17]) for an intra-sample comparison. The number of uniquely mapped sequences at each replicating time point is compared to that of the non-replicating sample, and the resulting enrichment ratio is smoothed by a least-squares polynomial digital filter to preserve higher moments and prevent signal distortion [20]. Overlapping replication profiles are generated to visualize origin firing and replisome progression.

---

## 2 Materials

### 2.1 Specialized Equipment

1. Qubit Fluorometer (Invitrogen).
2. Magnetic Particle Concentrator (e.g., Invitrogen Dynal MPCTM-S).
3. Thermal Cycler (e.g., BioRad C1000).
4. Real-Time PCR Detection System (e.g., BioRad CFX96).
5. NextSeq 500 High-Throughput Sequencing System (Illumina).
6. 2100 Bioanalyzer (Agilent).

### 2.2 Kits

1. Yeast DNA Purification Kit (Epicentre MasterPure).
2. Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific).
3. Nextera Amplicon DNA Sample Preparation Kit (Illumina).
4. Nextera XT Index Kit (24 indices—96 samples) (Illumina).
5. AMPure XP beads (Beckman Coulter).
6. KAPA Library Quantification Kit Illumina platforms (Complete Kit - KK4824, KAPA Biosystems).
7. High Sensitivity DNA Kit (Agilent).

### 2.3 Plasticware and Reagents

1. 1.5 mL microfuge tubes (e.g., Eppendorf DNA LoBind).
2. Multiplate Low-Profile 96-Well, Unskirted PCR Plates.
3. Adhesive Seals for 96-well PCR plates.
4. Alpha factor: 5 mg/mL in 95% ethanol, store at  $-20^{\circ}\text{C}$ .
5. Pronase: 10 mg/mL in double-distilled  $\text{H}_2\text{O}$ , always prepare fresh.
6. 10% (w/v) Sodium Azide.
7.  $1\times$  TE buffer: 10 mM Tris pH 8.0, 1 mM EDTA.
8. 80% (v/v) Ethanol.
9. Buffer P1 Resuspension Buffer: 50 mM Tris-HCl, 10 mM EDTA, adjust pH to 8.0, 100  $\mu\text{g}/\text{mL}$  RNase A final, store at  $2-8^{\circ}\text{C}$ .
10.  $1\times$  Library Dilution Buffer: 10 mM Tris-HCl pH 8.0, 0.05% Tween 20.
11. FACS Buffer: 200 mM Tris-Cl pH 7.5, 200 mM NaCl, 78  $\text{MgCl}_2$ .
12. 5 mL Polystyrene Round-Bottom tube (e.g., Falcon 352052).
13. 2X SYBR solution: 1/5000 SYBR green in 50 mM Tris-Cl pH 7.5.

### 2.4 Data Analysis Software

1. FlowJo (<http://www.flowjo.com>).
2. DNA sequencing files (FASTQ-formatted files).



3. FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).
4. Bowtie (<http://bowtie-bio.sourceforge.net/index.shtml>).
5. matplotlib version 1.5.1 (<http://matplotlib.org>).
6. NumPy (<http://www.numpy.org>).
7. Python (<http://www.python.org>).
8. SciPy (<http://www.scipy.org>).
9. SAMtools version 1.3 (<https://github.com/samtools/samtools>).
10. deepTools version 2.3.4 (<https://github.com/fidelram/deepTools>).
11. Plotly (<https://plot.ly>).
12. bigWigToWig (<https://genome.ucsc.edu/goldenpath/help/bigWig.html>).
13. wig2bed from BEDOPS (<http://bedops.readthedocs.io/en/latest/index.html>).

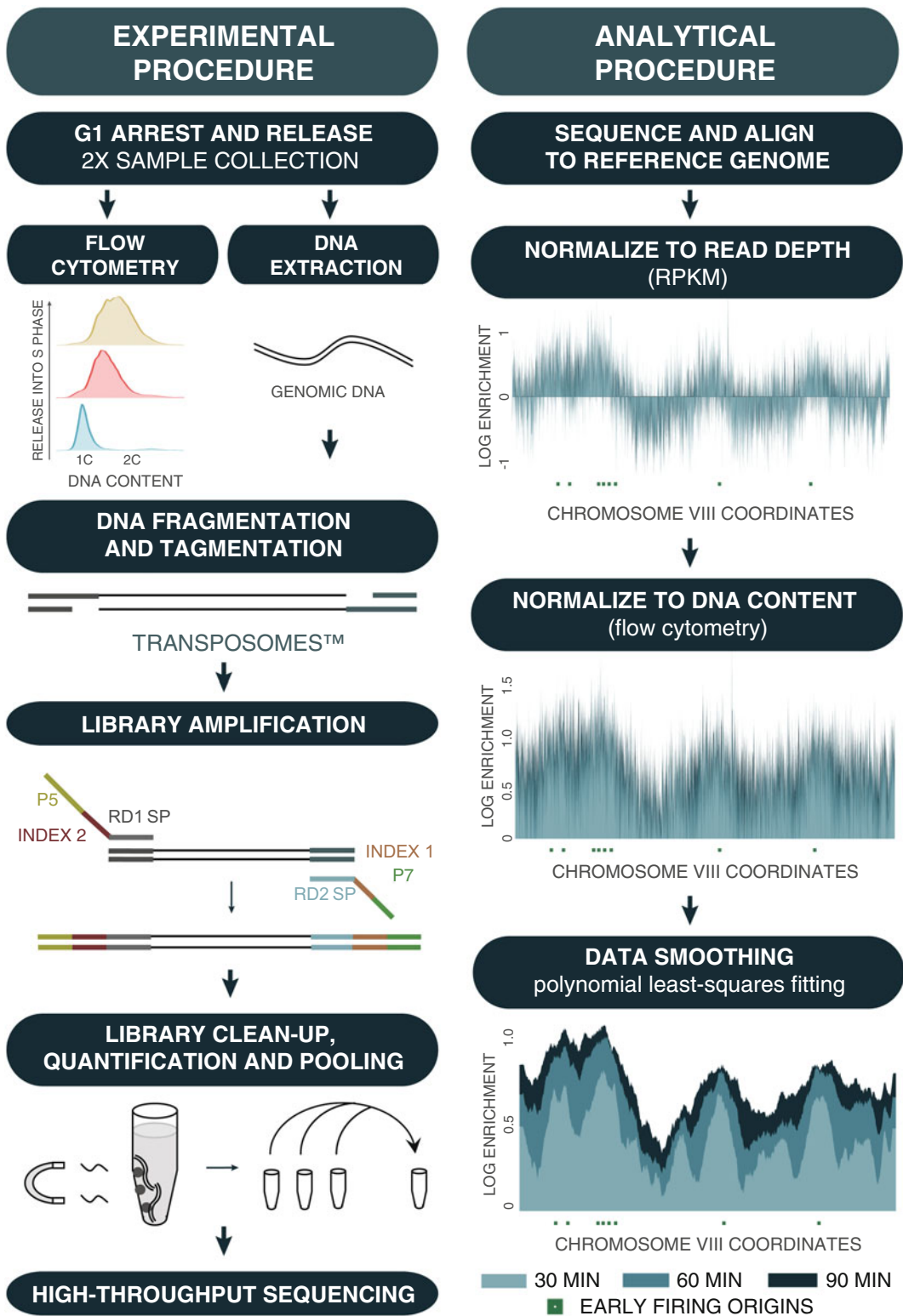
---

### 3 Methods

#### 3.1 Cell Synchronization and DNA Preparation

The experimental and analytical procedures of this protocol are summarized in Fig. 1.

1. Grow cells overnight in 50 mL of liquid medium to early log phase (OD<sub>600</sub> of 0.2–0.4) in a water bath shaker under preferred conditions (e.g., 30 °C, 200 rpm) (*see Note 1*).
2. Arrest the cells in G1 by adding alpha factor to 2 µg/mL final concentration, and incubate for 1 h. Add an additional 1 µg/mL alpha-factor and incubate for 1 h or until >90% of the cells have the “shmoo” morphology (*see Note 2*). Before releasing the cells from G1 take a sample of at least 2 OD, add sodium azide to 0.1% (w/v) final concentration, and leave on ice for at least 15 min (*see Note 3*).
3. Release the cells from G1 by adding pronase to 100 µg/mL final concentration (*see Note 4*).
4. For flow cytometry, harvest at least 0.5 OD<sub>600</sub> at the desired time points, and immediately resuspend in 1 mL 70% EtOH. Incubate at room temperature for 15 min or 4 °C until further processing (*see Subheading 3.2*).
5. For library preparations, harvest at least 2 OD<sub>600</sub> of the cells at the desired time points, immediately mix with sodium azide to 0.1% (w/v) final concentration, and leave on ice for at least 15 min.



**Fig. 1** Workflow of the experimental and analytical procedures. Asynchronous yeast cultures are arrested in G1 and released into S phase. DNA is extracted from G1-arrested samples as well as at the desired times

6. Extract genomic DNA with MasterPure Yeast DNA Purification Kit from Epicentre or similar, according to the manufacturer's instructions. The protocol can be downloaded from <http://www.epibio.com/docs/default-source/protocols/masterpure-yeast-dna-purification-kit.pdf>. Perform the RNase A treatment included in the kit (*see Note 5*).
7. Resuspend the samples in 35  $\mu\text{L}$  TE and determine DNA concentration using a Qubit Fluorometer and the Qubit dsDNA BR Assay Kit, or similar. Proceed to DNA tagmentation or store the DNA at  $-20\text{ }^{\circ}\text{C}$  (*see Note 6*).

### 3.2 Flow Cytometric Analysis

Flow cytometric analysis measures bulk cellular DNA content at each time sampled for copy number analysis, to verify synchronization quality and for downstream data normalization.

1. Starting with 1 mL of ethanol-fixed cells from Subheading 3.1.4, centrifuge at  $1200 \times g$  at room temperature for 2 min.
2. Remove the supernatant and wash cell pellet with 1 mL  $\text{dH}_2\text{O}$ . Centrifuge again at  $1200 \times g$  at room temperature for 1 min.
3. Remove the supernatant and resuspend cell pellet in 0.5 mL 50 mM Tris-Cl pH 7.5 containing 10 mg/mL freshly prepared proteinase K. Incubate for 40 min at  $50\text{ }^{\circ}\text{C}$ .
4. Centrifuge at  $1200 \times g$  at room temperature for 1 min. Remove the supernatant and resuspend cell pellet in 0.5 mL FACS buffer.
5. Samples can now be stored at  $4\text{ }^{\circ}\text{C}$  for up to a week.
6. Transfer 100  $\mu\text{L}$  samples to a 5 mL round-bottom tube.
7. Add 0.5 mL  $2\times$  SYBR solution (diluted in 50 mM Tris-Cl pH 7.5).
8. Sonicate each sample at low intensity for 1–3 s.
9. Analyze on flow cytometer, measuring SYBR green fluorescence for at least 10,000 cells.

### 3.3 DNA Tagmentation

This step uses Illumina's Nextera Amplicon DNA Sample Preparation Kit to fragment the DNA and add adapter sequences to the ends.

---

**Fig. 1** (continued) following release into S phase. Libraries for each sample are prepared by the addition of the Transposome complex, which fragments the DNA and attaches transposons to the DNA 5' ends. The libraries are amplified and barcoded with a unique combination of indexed adapters. The samples are then purified, quantified, and normalized. Finally, the libraries are pooled together and submitted for high-throughput sequencing. For data analysis, reads are aligned to a reference *S. cerevisiae* genome and normalized to the reads per kilobase per million (RPKM). Each sample is scaled according to the median DNA content, as determined by flow cytometry. Read counts are then compared to that of the G1-arrested (non-replicating) sample. Resultant replication profiles are smoothed by the Savitzky-Golay filter and visualized using graphing software

1. To prevent contamination of the samples, clean the bench thoroughly with bleach or similar and always wear gloves.
2. In a PCR tube for each sample, add 5  $\mu\text{L}$  of the extracted DNA at 0.2 ng/ $\mu\text{L}$  (1 ng in total), 10  $\mu\text{L}$  Tagment DNA buffer (TD buffer), and 5  $\mu\text{L}$  Amplicon Tagmentation Mix (ATM). Mix gently.
3. Centrifuge at  $1000 \times g$  at room temperature for 1 min.
4. Incubate the reaction in a thermal cycler at 55 °C for 5 min and then hold at 10 °C.
5. Add 5  $\mu\text{L}$  Neutralize Tagment buffer (NT buffer) to each sample and mix gently.
6. Centrifuge at  $1000 \times g$  at room temperature for 1 min.
7. Incubate the reaction at room temperature for 5 min.

### 3.4 Library Amplification

This step uses Illumina's Nextera XT Index Kit to amplify the tagmented DNA by PCR, by adding unique combinations of the adapters Index 1 (i7) and Index 2 (i5) to each library.

1. In a PCR tube, add to each sample 15  $\mu\text{L}$  of Nextera PCR Master Mix and 5  $\mu\text{L}$  of each appropriate index primer (i5 and i7 respectively) (*see Note 7*). If multiple libraries will be pooled for sequencing, then each library must have a unique i7 primer (for single-end reads) or a unique index primer combination (for paired-end reads).
2. Centrifuge at  $280 \times g$  for 1 min.
3. Place samples in a thermal cycler and heat at 72 °C for 3 min, followed by 30 s at 95 °C. Next, run 12 cycles of 95 °C denaturation for 10 s, 55 °C annealing for 30 s and 72 °C extension for 30 s. Allow a final extension of 72 °C for 5 min and hold at 10 °C (*see Note 8*).
4. At this point the library can be left at 2–8 °C for up to 2 days, or proceed to library cleanup.

### 3.5 Library Cleanup

1. Centrifuge PCR products at  $280 \times g$  for 1 min at room temperature. Transfer volume (50  $\mu\text{L}$ ) to a LoBind eppendorf tube, or similar.
2. Bring the AMPure XP beads to room temperature and vortex for 30 s to ensure that they are evenly dispersed. Add 30  $\mu\text{L}$  of beads to each PCR product sample ( $0.6 \times$  volume). Mix well.
3. Incubate at room temperature without shaking for 5 min. Place on a magnetic particle concentrator (e.g., Invitrogen Dynal MPC™-S Magnetic Particle Concentrator) for 2 min or until the supernatant has cleared. Discard the supernatant.

4. Wash beads with 200  $\mu\text{L}$  fresh 80% ethanol without resuspending the beads. Incubate on a stand for 30 s or until the supernatant has cleared. Discard the supernatant.
5. Repeat washing step with 200  $\mu\text{L}$  80% ethanol. Remove excess ethanol with a P10 pipette and allow beads to air-dry for 15 min.
6. Remove samples from the magnetic stand and add 52.5  $\mu\text{L}$  of P1 Resuspension buffer. Mix well and incubate at room temperature for 2 min.
7. Place on a magnetic stand for 2 min or until the supernatant has cleared.
8. Transfer 48  $\mu\text{L}$  of the supernatant to a fresh LoBind eppendorf tube.
9. The samples can be stored at  $-20\text{ }^{\circ}\text{C}$  for up to 7 days before sequencing. Library quantification is recommended to reduce sequencing cost (*see* Subheading 3.6).

### **3.6 Library Quantification and Pooling**

Accurate library quantification ensures equal representation of each library when they are pooled prior to sequencing. In this protocol, DNA quantification is performed by qPCR analysis using the KAPA Library Quantification Kit for Illumina platforms, which includes 2 $\times$  KAPA SYBR FAST qPCR Master Mix, 10 $\times$  Illumina Primer Premix, and Illumina DNA Standards 1–6.

1. Make three separate 1:1000 dilutions of each library in Library Dilution Buffer.
2. Prepare the PCR reaction mix taking into account that both the samples and the DNA standards must have three replicates each. Aliquot 16  $\mu\text{L}$  of the PCR Reaction Mix into each well. Add 4  $\mu\text{L}$  of the corresponding DNA Standard or 4  $\mu\text{L}$  of the 1:1000 dilution of each library into the appropriate well of the 96-well plate.
3. Seal the 96-well plate with an adhesive seal and centrifuge at  $280 \times g$  for 1 min at room temperature.
4. Incubate in the real-time thermocycler for 5 min at  $95\text{ }^{\circ}\text{C}$  and then apply 35 cycles of denaturation at  $95\text{ }^{\circ}\text{C}$  for 30 s and extension at  $60\text{ }^{\circ}\text{C}$  for 60 s.
5. Generate the standard curve by plotting the quantification cycle values (Cq) to the DNA concentrations of the standards. It should have an efficiency of amplification close to 100% and  $R^2 > 0.99$ . Calculate the concentration of each library using the standard curve as a reference.
6. Analyze the samples on an Agilent 2100 Bioanalyzer using the High Sensitivity DNA Kit to determine the size of each library, which will be used to adjust the real concentration of each

sample. The protocol can be downloaded from [http://www.agilent.com/cs/library/usermanuals/Public/G2938-90321\\_SensitivityDNA\\_KG\\_EN.pdf](http://www.agilent.com/cs/library/usermanuals/Public/G2938-90321_SensitivityDNA_KG_EN.pdf).

7. With the concentration corrected, normalize the libraries to 30 nM and pool them, taking into account that the lowest volume added of the most concentrated library should be at least 2  $\mu$ L. We pool between 24 and 32 libraries, depending on the number of reads we are anticipating from the sequencer.
8. As a quality control, the pool must be analyzed on Bioanalyzer. Finally, perform a qPCR to determine the concentration of the final pool, using the KAPA Library Quantification Kit.
9. Submit the pooled library for high-throughput sequencing on the Illumina NextSeq 500 System. At least  $10^7$  reads per library for yeast are necessary for robust analysis. Samples can be multiplexed in a Single-Read Illumina flow cell as long as each carries a unique Index 1 (i7) primer. Unique Index 2 (i5) primers can also be used to increase the number of samples per flow cell, using Paired-End flow cell sequencing. Single-end reads of 50 bp are sufficient for this application.

### **3.7 Measuring Median DNA Content from Flow Cytometry Data**

As described in Muller et al. [17], normalization by flow cytometry is applied to establish a baseline of  $\log_2 = 1.0$  for samples taken at different time points.

1. Data processing. Using any available package for flow cytometry analysis (i.e., FlowJo), upload flow-cytometry standard format (FCS) files and gate appropriately to exclude debris, aggregates, and cell doublets. FCS files can then be exported for downstream analysis using any statistical software package (i.e., SciPy, R).
2. Peak normalization of flow cytometry data. To control for fluorescence readout shifts from experimental variation, align G1 and G2 asynchronous peaks of each time sample strain to the wild-type control. Apply the same scaling factor to all time points for each strain.
3. Scale factor determination. Determine the median DNA content of each time sample relative to the median DNA content of the G1-arrested cells.

### **3.8 Quantifying and Visualizing Genomic Replication**

Data processing steps are conducted in a Unix shell environment. Downstream data analysis is conducted in Python, but similar packages are available to run analyses through alternative software (i.e., MATLAB, R). Input files are expected to be compressed FASTQ files, containing all raw sequence reads.

1. Quality assessment. Retrieve demultiplexed FASTQ-formatted sequencing files from the Illumina sequencing run. Perform quality assessment using FastQC or similar software (*see Note 9*).
2. Read alignment. Index the reference genome in FASTA format with the bowtie “bowtie-build” indexer function. Align FASTQ format cells to the reference genome with the bowtie “aligner” function. Specify output in SAM file format (*see Note 10*).
3. Read mapping and processing.
  - (a) Convert SAM files into binary compressed BAM files using SAMtools “view” function.
  - (b) Sort and index BAM files using SAMtools “sort” and “index” functions. This should be done on a per-chromosome basis to generate individual chromosome maps if desired. *See* SAMtools documentation for specifying genomic regions.
  - (c) Determine number of uniquely mapped reads per specified bin size using the “bamCoverage” function of deepTools. Normalize to Reads Per Kilobase per Million mapped reads (RPKM) to normalize for read depth. Extend reads to read fragment size, as specified by Illumina sequencing conditions (*see Note 11*).
  - (d) Compare output bigWig file of each time sample to output bigWig file of G1-arrested sample using “bigwigCompare” function of deepTools. Specify “scaleFactors” according to factors determined by flow cytometry in Subheading 3.7 (*see Note 12*).
  - (e) For easier data manipulation, convert to BED files using bigWigToWig executable followed by wig2bed. BED files can still be uploaded to most genome browsers for manual inspection or be read by statistical analysis software (i.e., Python, R).

### **3.9 Replication Profile Visualization**

1. For more readable replication profile visualization, import data from BED files and apply a polynomial data smoothing filter (i.e., Savitsky-Golay digital filter) to reduce noise. This can be done using the “savgol\_filter” function of SciPy in Python or similar software (*see Note 13*).
2. Generate overlapping chromosomal maps of smoothed data from BED files by chromosomal coordinates using Plotly, matplotlib, or similar software.

---

## 4 Notes

1. If the culture is too dense (i.e.,  $OD_{600} > 0.7$ ), dilute to an  $OD_{600}$  of 0.1–0.2 and allow at least two cell doublings before adding alpha factor.
2. It is recommended that the cell cycle arrest and release be confirmed by flow cytometry, taking samples at **steps 1, 2, and 4**.
3. Multiple samples can be left on ice for up to 2 h. The amount of cells needed will depend on the yield of the DNA extraction and purification, if different methods are chosen. Minimum DNA final concentration required for the library preparation is 0.2 ng/ $\mu$ L.
4. At this point, conditions can be customized, for example by adding a drug treatment or changing the temperature.
5. Some kits specify RNase treatment during the DNA extraction. If RNase treatment is done after the DNA extraction, an additional extraction of the DNA with phenol:chloroform followed by ethanol precipitation is required to remove the RNase A.
6. The DNA concentration should be determined not longer than 1 day before the library preparation, and is typically 3–4  $\mu$ g. It is important to use a fluorescence-based method rather than UV absorbance.
7. Use the recommended combination of the i5 and i7 adapters as suggested by the Nextera XT DNA Library Prep Reference Guide.
8. Use the recommended number of PCR cycles to ensure high-quality results.
9. FastQC will generate an HTML file (readable with a web browser) containing information on quality including potential contaminations or technical biases. Full descriptions of analysis modules as well as common reasons for errors are available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/>.
10. Reference genomes in FASTA format are available for download at <http://hgdownload.cse.ucsc.edu/downloads.html>. If paired-end sequencing was used, this should be specified at this point. An alternative to this is uploading sample files to BWA Aligner, a BaseSpace Lab App developed by Illumina. This software will index and align FASTQ files against any given reference file and fix read pairing information and flags. SAM files are converted to BAM files which are then sorted, merged, and indexed. If selected, the BWA Aligner app will calculate alignment statistics.



11. Bin size can be modified according to desired resolution. The output file is in bigWig Track format (compressed binary versions of WIG files) and can be visualized independently in most genome browsers (i.e., UCSC Genome Browser, IGV). Visualization at this stage may identify any experimental or analytical issues up to this point.
12. Chromosome maps can be visualized at this point by uploading output bigWig files to UCSC Genome Browser, IGV or similar software to compare replication profiles of various time point samples.
13. Filter width can be specified but should be maintained below width of smooth line profiles. Non-polynomial least-square filters can be applied with loss of peak resolution.

---

## Acknowledgments

This work was supported by the Canadian Institutes of Health Research and the Canadian Cancer Society Research Institute. We thank the Donnelly Sequencing Centre for their deep sequencing expertise.

## References

1. Huberman JA, Riggs AD (1968) On the mechanism of DNA replication in mammalian chromosomes. *J Mol Biol* 32(2):327–341
2. Kelly TJ, Brown GW (2000) Regulation of chromosome replication. *Annu Rev Biochem* 69:829–880
3. Bell SP, Dutta A (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* 71:333–374
4. Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M (2010) Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem* 79:89–130. doi:10.1146/annurev.biochem.052308.103205
5. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, Lockhart DJ, Davis RW, Brewer BJ, Fangman WL (2001) Replication dynamics of the yeast genome. *Science* 294(5540):115–121
6. Agier N, Fischer G (2012) The mutational profile of the yeast genome is shaped by replication. *Mol Biol Evol* 29(3):905–913. doi:10.1093/molbev/msr280
7. Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR (2009) Human mutation rate associated with DNA replication timing. *Nat Genet* 41(4):393–395. doi:10.1038/ng.363
8. Seiler JA, Conti C, Syed A, Aladjem MI, Pomnier Y (2007) The intra-S-phase checkpoint affects both DNA replication initiation and elongation: single-cell and -DNA fiber analyses. *Mol Cell Biol* 27(16):5806–5818. doi:10.1128/MCB.02278-06
9. Yamazaki S, Hayano M, Masai H (2013) Replication timing regulation of eukaryotic replisomes: Rif1 as a global regulator of replication timing. *Trends Genet: TIG* 29(8):449–460. doi:10.1016/j.tig.2013.05.001
10. van Brabant AJ, Raghuraman MK (2002) Assaying replication fork direction and migration rates. *Methods Enzymol* 351:539–568
11. Tercero JA, Diffley JF (2001) Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature* 412(6846):553–557
12. Katou Y, Kanoh Y, Bando M, Noguchi H, Tanaka H, Ashikari T, Sugimoto K, Shirahige K (2003) S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature* 424(6952):1078–1083
13. Feng W, Collingwood D, Boeck ME, Fox LA, Alvino GM, Fangman WL, Raghuraman MK, Brewer BJ (2006) Genomic mapping of single-stranded DNA in hydroxyurea-challenged

- yeasts identifies origins of replication. *Nat Cell Biol* 8(2):148–155
14. Lengronne A, Pasero P, Bensimon A, Schwob E (2001) Monitoring S phase progression globally and locally using BrdU incorporation in TK(+) yeast strains. *Nucleic Acids Res* 29 (7):1433–1442
  15. Yabuki N, Terashima H, Kitada K (2002) Mapping of early firing origins on a replication profile of budding yeast. *Genes Cells* 7 (8):781–789. doi:559 [pii]
  16. Gilbert DM (2010) Evaluating genome-scale approaches to eukaryotic DNA replication. *Nat Rev Genet* 11(10):673–684. doi:10.1038/nrg2830
  17. Muller CA, Hawkins M, Retkute R, Malla S, Wilson R, Blythe MJ, Nakato R, Komata M, Shirahige K, de Moura AP, Nieduszynski CA (2014) The dynamics of genome replication using deep sequencing. *Nucleic Acids Res* 42 (1):e3. doi:10.1093/nar/gkt878
  18. Futcher B (1999) Cell cycle synchronization. *Methods Cell Sci* 21(2–3):79–86
  19. Smith J, Manukyan A, Hua H, Dungrawala H, Schneider BL (2017) Synchronization of yeast. *Methods Mol Biol* 1524:215–242. doi:10.1007/978-1-4939-6603-5\_14
  20. Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least square procedures. *Analytical Chem* 36(8):1627–1639

# Chapter 16

## Quantitative Bromodeoxyuridine Immunoprecipitation Analyzed by High-Throughput Sequencing (qBrdU-Seq or QBU)

Joanna E. Haye-Bertolozzi and Oscar M. Aparicio

### Abstract

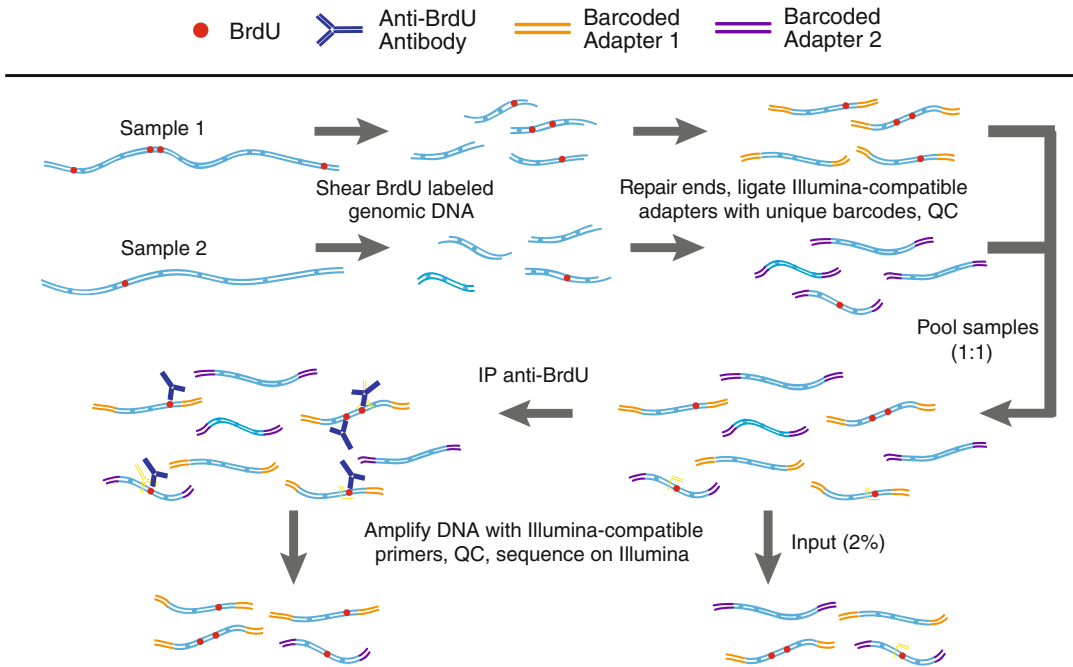
Incorporation into DNA of nucleoside analogs like 5-bromo-2'-deoxyuridine (BrdU) is a powerful tool for *in vivo* studies of DNA synthesis during replication and repair. Immunoprecipitation of BrdU-labeled DNA analyzed by DNA sequencing (BrdU-IP-seq) allows for genome-wide, sequence-specific tracking of replication origin and replication fork dynamics under different conditions, such as DNA damage and replication stress, and in mutant strains. We have recently developed a quantitative method for BrdU-IP-seq (qBrdU-seq) involving DNA barcoding to enable quantitative analysis of multiple experimental samples subjected to BrdU-IP-seq. After initial barcoding of multiple, individually BrdU-labeled genomic DNA samples, a pooling strategy is used for all subsequent steps including immunoprecipitation, amplification, and sequencing, which eliminates sample-to-sample variability in these steps. Parallel processing of an aliquot of the pooled input sample provides a direct control for the normalization of the data and yields results that allow quantitative comparisons of the experimental samples. Though developed for the analysis of *S. cerevisiae*, this method should be directly adaptable to other model systems.

**Key words** 5-Bromo-2'-deoxyuridine, Nucleoside analog, DNA replication, DNA barcoding, Data normalization between experiments

---

### 1 Introduction

Incorporation of nucleoside analogs such as 5-Bromo-2'-deoxyuridine (BrdU) has been used for decades to study DNA replication *in vivo* [1]. Such studies in yeasts have been facilitated by the development of vectors expressing thymidine kinase and a nucleoside transporter to reconstruct the thymidine salvage pathway, which is lacking in budding and fission yeasts [2–7]. BrdU incorporation is now routinely used in studies of replication origin and replication fork dynamics under normal and stressed conditions, which are known to affect both origin and fork functions. BrdU incorporation into DNA may be analyzed by several means, such as immunofluorescence, which allows for cytogenetic analysis [8],



**Fig. 1** Scheme of quantitative BrdU-IP-seq analysis. BrdU-labeled genomic DNA from each sample is barcoded by end-ligation of Illumina-compatible linkers. Samples are pooled, a small fraction of this pool is set aside as “Input” and the remainder is subjected to immunoprecipitation (IP) with anti-BrdU antibody. The IP and Input samples are PCR-amplified with indexed primers and sequenced. IP sample reads are normalized against Input sample reads. Adapted from [14] under Creative Commons Attribution-NonCommercial 4.0 International License (CC-BY-NC)

DNA fiber labeling, which allows for the analysis of the replication of individual molecules, though typically without DNA sequence information [9], and immunoprecipitation (IP), which allows determination of specific DNA sequences incorporating BrdU (BrdU-IP-chip or BrdU-IP-seq), but is normally carried out as a population measure [10–12]. Nevertheless, we note that there is no inherent reason BrdU-IP-seq cannot be performed on individual cells; indeed, BrdU-IP is more sensitive than ChIP, for example, which is amenable to single cell analysis [13].

We have recently developed a quantitative BrdU-IP-seq (qBrdU-seq or QBU) that allows for direct quantitative comparisons of BrdU-incorporation amongst independently generated experimental samples (Fig. 1) [14]. The key features that distinguish the QBU procedure from BrdU-IP-seq are: initial barcoding of individual BrdU-labeled samples; pooling of multiple experimental samples, including replicates, for IP, amplification, and sequencing; and parallel analysis of an aliquot of the pooled input for quantitative normalization of the data. Not only does this approach eliminate technical variability due to individual sample processing at each of the above steps and the purification steps in

between, but it reduces labor and expense of multiple, individual sample preparations. QBU is frequently performed on up to 12 samples concurrently, but should be scalable.

Because BrdU-IP experiments are performed on populations of cells and events related to DNA replication and repair occur in a spatio-temporal manner (e.g., differential origin initiation timing or replication fork stalling/progression), meaningful analysis typically requires culture synchronization. In the protocol below, we describe synchronization of *S. cerevisiae* cells in late G1 phase with  $\alpha$ -factor followed by release into S-phase in the presence of BrdU with or without hydroxyurea, which arrests replication in early S phase. Hydroxyurea is often used to provide an early replication profile from a single time-point, which can reveal deregulation of normal replication timing patterns or intra-S checkpoint control of (late and dormant) origin firing. Cells are harvested at the appropriate time(s) depending on experimental design and genomic DNA is isolated. The DNA from each sample is uniquely identified by barcoded linker addition to the DNA ends (Fig. 1). The samples are pooled, a small aliquot of this “Input” is set aside, and the pool is denatured and subjected to IP with antibody against BrdU. The IP DNA and Input DNA are amplified in parallel with distinctly indexed primers and sequenced. The resulting reads are demultiplexed and mapped to the genome and IP read counts are divided by the Input read counts for the corresponding sample. This normalization step corrects for any differences in the initial amounts of the individual samples pooled together, as well as possible differences in efficiency of the adapter ligation/barcoding step. Thus, the normalized data are suitable for quantitative comparisons among the pooled samples. The protocol includes quality control steps to confirm efficient end-linker addition and library amplification, and outlines a computational pipeline for initial analysis of the sequence data.

---

## 2 Materials

All solutions should be prepared with deionized, nuclease-free H<sub>2</sub>O, except growth medium, which should be prepared with pure, deionized, or distilled H<sub>2</sub>O. All tips and tubes should be sterile and nuclease-free. Reagents, kits, and solutions should be stored according to the manufacturer’s guidelines, or as indicated below.

### 2.1 Growth, Synchronization, and Harvesting of Cells

1. YEPD broth: 1%(w/v) Yeast Extract, 2%(w/v) Bacto-Peptone, 2%(w/v) Dextrose.
2. 1 mg/mL  $\alpha$ -factor (e.g., Sigma T6901) (in 1% DMSO), store at  $-20^{\circ}\text{C}$ .

3. Hydroxyurea (e.g., Sigma H8627).
4. 40 mg/mL 5-Bromo-2'-deoxyuridine, (in 100 mM Tris-HCl pH 7.6), store at  $-20^{\circ}\text{C}$ , protected from light.
5. 20 mg/mL Pronase (e.g., Sigma P5147), store at  $-20^{\circ}\text{C}$ .
6. Probe sonicator, for disaggregation of cells.
7. 10%(w/v) Sodium azide, store at  $4^{\circ}\text{C}$ .
8. 50 mL sterile polypropylene tubes.
9. 2 mL flat- or round-bottom tubes with gasket-sealed screw-caps.
10. Tris-buffered saline (TBS): 100 mM Tris-HCL pH 7.6, 150 mM NaCl, store at  $4^{\circ}\text{C}$ .

**2.2 Cell Lysis,  
Genomic DNA  
Isolation, and  
Chromatin  
Fragmentation**

1. Glass beads, 425–600  $\mu\text{m}$  (sterile).
2. Genomic lysis buffer: 100 mM Tris-HCl pH 8.0, 50 mM EDTA, 1% SDS, store at room temperature.
3. 5 M NaCl, store at room temperature.
4. FastPrep 24 Instrument (MP Biomedicals). Vortexers or other vigorous agitators may be used instead.
5. 26G hypodermic needles.
6. 5 mL polypropylene snap-cap tubes; caps not required.
7. 1.7 mL microcentrifuge tubes.
8. 2 mL Phase-lock gel tubes (Quanta Biosciences 2302830). Optional: The use of Phase-lock gel is recommended for safety, reproducibility, and effective separation of contaminants.
9. PCI (Phenol/Chloroform/Isoamyl alcohol, 25:24:1, buffered for DNA), store at  $4^{\circ}\text{C}$ .
10. TE: 10 mM Tris-HCl pH 7.6, 1 mM EDTA, store at room temperature.
11. 20 mg/mL RNase A (DNase-free), store at  $-20^{\circ}\text{C}$ .
12. 10 $\times$  SDS-PK: 1 mg/mL Proteinase K (in 200 mM Tris-HCl pH 7.6, 200 mM EDTA pH 8, 2% SDS, 1 mM  $\text{CaCl}_2$ ), store at  $-20^{\circ}\text{C}$ .
13. Qiagen QIAquick PCR purification kit (includes buffers PB, PE, and EB) (*see Note 1*).
14. Nanodrop spectrophotometer (Cole-Palmer) or alternative instrument for measurement of DNA concentration in a low volume.
15. Covaris S2 Sonicator (*see Note 2*).
16. Covaris microTUBE AFA Fiber Pre-slit, Snap-Cap (Covaris 520045).

### 2.3 DNA Barcoding/ Adapter Ligation

1. KAPABiosystems Hyper prep kit, KR0961-v3.15 (*see Note 3*).
2. Qiagen QIAquick PCR purification kit (includes buffers PB, PE, and EB) (*see Note 1*).
3. Barcoded adapters (*see Table 1* for sequences, and **Note 4** for annealing protocol).
4. Nanodrop spectrophotometer (Cole-Palmer) or alternative instrument for measurement of DNA concentration in a low volume.

**Table 1**  
Oligonucleotides sequences for barcoded adapters. All sequences are given 5' to 3'; P indicates 5'-phosphate. Sequences are taken from [23].

Oligonucleotide	Sequence
Barcode 1a	P-CAGTAGATCGGAAGAGCACACGTCT
Barcode 1b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGT
Barcode 2a	P-GCATAGATCGGAAGAGCACACGTCT
Barcode 2b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGCT
Barcode 3a	P-AGCTAGATCGGAAGAGCACACGTCT
Barcode 3b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTT
Barcode 4a	P-CGTAAGATCGGAAGAGCACACGTCT
Barcode 4b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGT
Barcode 5a	P-TCGAAGATCGGAAGAGCACACGTCT
Barcode 5b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
Barcode 6a	P-GTCAAGATCGGAAGAGCACACGTCT
Barcode 6b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGACT
Barcode 7a	P-ACTGAGATCGGAAGAGCACACGTCT
Barcode 7b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTT
Barcode 8a	P-CTAGAGATCGGAAGAGCACACGTCT
Barcode 8b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGT
Barcode 9a	P-TACGAGATCGGAAGAGCACACGTCT
Barcode 9b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTAT
Barcode 10a	P-GATCAGATCGGAAGAGCACACGTCT
Barcode 10b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCT
Barcode 11a	P-ATGCAGATCGGAAGAGCACACGTCT
Barcode 11b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATT
Barcode 12a	P-TGACAGATCGGAAGAGCACACGTCT
Barcode 12b	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCAT

**Table 2**

**Oligonucleotide sequences for library amplification primers. All sequences are given 5' to 3'; \* indicates phosphorothioate bond, which protects from nucleolytic degradation. Sequences are taken from [23]**

Illumina primers	Sequence
Multi/Std primer 1.0	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T
Multiplexing primer 2.0	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index primer 1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTT*C
Index primer 2	CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTT*C
Index primer 3	CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTGGAGTT*C
Index primer 4	CAAGCAGAAGACGGCATAACGAGATTGGTCAAGTGACTGGAGTT*C
Index primer 5	CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTT*C
Index primer 6	CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTT*C
Index primer 7	CAAGCAGAAGACGGCATAACGAGATGATCTGGTGACTGGAGTT*C
Index primer 8	CAAGCAGAAGACGGCATAACGAGATCAAGTGTGACTGGAGTT*C
Index primer 9	CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTT*C
Index primer 10	CAAGCAGAAGACGGCATAACGAGATAAGCTAGTGACTGGAGTT*C
Index primer 11	CAAGCAGAAGACGGCATAACGAGATGTAGCCGTGACTGGAGTT*C
Index primer 12	CAAGCAGAAGACGGCATAACGAGATTACAAGGTGACTGGAGTT*C

#### **2.4 Validation of Adapter Ligation**

1. Phusion DNA Polymerase (New England Biolabs M0530L).
2. Multi/Std primer 1.0 (25  $\mu$ M) (*see* Table 2 for sequence).
3. Multiplexing primer 2.0 (0.5  $\mu$ M) (*see* Table 2).
4. Index primers (25  $\mu$ M) (*see* Table 2).

#### **2.5 Pooling and BrdU Immunoprecipitation**

1. 10 $\times$  PBS: 1.37 M NaCl, 27 mM KCl, 100 mM Na<sub>2</sub>HPO<sub>4</sub>, 18 mM KH<sub>2</sub>PO<sub>4</sub>; store at room temperature.
2. 2 $\times$ IP buffer: 2 $\times$ PBS +0.10%(v/v) Triton-X100; prepare fresh.
3. Wash buffer: PBS +0.05%(v/v) Triton-X100; prepare fresh.
4. TE: 10 mM Tris-HCl pH 7.6, 1 mM EDTA, store at room temperature.
5. Elution buffer: TE +0.5%(w/v) SDS; store at room temperature.
6. Anti-BrdU antibody (Invitrogen 033900).
7. DynaBeads Protein G (Invitrogen 10004D) (*see* Note 5).
8. Deionized, nuclease-free H<sub>2</sub>O.
9. Microcentrifuge tube rotator.



10. Magnetic rack for standard microcentrifuge tubes.
11. Qiagen QIAquick PCR purification kit (includes buffers PB, PE, and EB) (*see Note 1*).

### **2.6 Library Amplification**

1. Phusion DNA Polymerase (New England Biolabs M0530L).
2. Multi/Std primer 1.0 (25  $\mu$ M) (*see Table 2* for sequence).
3. Multiplexing primer 2.0 (0.5  $\mu$ M) (*see Table 2*).
4. Index primers (25  $\mu$ M) (*see Table 2*).
5. AMPure beads (Beckman Coulter A63880).
6. Magnetic rack for standard microcentrifuge tubes.
7. Nanodrop spectrophotometer (Cole-Palmer) or alternative instrument for measurement of DNA concentration in a low volume.

### **2.7 Quality Control and Quantification**

1. BioAnalyzer (Agilent Technologies) (*see Note 6*).

### **2.8 DNA Sequencing**

1. Illumina sequencing instrument (*see Note 7*).

### **2.9 Data Analysis**

1. Software for analysis of sequencing data (ea-utils, Bowtie2, SAMtools, BEDTools2, MACS, R-studio, Diffbind; alternatives are available).
2. High-performance computing platform.

---

## **3 Methods**

### **3.1 Growth, Synchronization, and Harvesting of Cells**

1. Grow 10–20 mL cells per sample to OD  $\sim$  0.5 in YEPD at 23  $^{\circ}$ C (*see Note 8*).
2. Transfer culture to a 15 or 50 mL sterile polypropylene tube(s) and pellet cells by centrifugation in a swinging-bucket rotor at  $\sim$ 1500  $\times g$ , for 3 min; discard the supernatant.
3. Resuspend cells in equal volume of fresh YEPD containing  $\alpha$ -factor ( $\sim$ 5 nM for *bar1 $\Delta$*  strains and  $\sim$ 5  $\mu$ M for *BAR1* strains), and transfer culture to a conical flask and incubate for 4 h at 23  $^{\circ}$ C (or as otherwise appropriate) on a platform shaker or shaking water bath ( $\sim$ 180 rpm) (*see Note 9*).
4. For synchronous release, pellet cells by centrifugation in 15 or 50 mL sterile polypropylene tubes in a swinging-bucket rotor at  $\sim$ 1500  $\times g$ , for 3 min. Discard the supernatant.
5. Resuspend cells in an equal volume of fresh YEPD containing 400  $\mu$ g/mL BrdU +200  $\mu$ g/mL Pronase (for *bar1 $\Delta$*  strains, to degrade  $\alpha$ -factor) +200 mM hydroxyurea (optional). Sonicate the resuspended cells while still in the centrifuge tube with a

probe sonicator to disaggregate cells, which aggregate in response to  $\alpha$ -factor (*see Note 10*). Transfer to a conical flask for incubation at 23 °C with shaking.

6. Harvest cells at the appropriate time(s) after release (*see Note 11*). Transfer cells to a 15 or 50 mL screw-cap tube containing an appropriate volume of 10% sodium azide to achieve 0.1% final after addition of culture, and mix by inversion. Pellet cells by centrifugation as above and discard the supernatant.
7. Resuspend cells in 10 mL ice-cold TBS. Pellet cells by centrifugation as above and discard the supernatant.
8. Resuspend cells in 1 mL ice-cold TBS and transfer to 2 mL screw-cap (FastPrep) tube. Pellet cells in microcentrifuge by spinning at full speed for ~3 s. Discard the supernatant. Proceed to DNA isolation or freeze samples at -20 °C.

### **3.2 Cell Lysis, Genomic DNA Isolation, and Chromatin Fragmentation**

1. *Cell lysis*: If cells were frozen, thaw on ice.
2. Resuspend cells in 500  $\mu$ L genomic lysis buffer. To each tube, add an equal volume of glass beads (a 0.5 mL microcentrifuge tube may be used to measure and dispense the beads).
3. Disrupt cells by bead-beating with FastPrep (45 s, speed 5.5). Other instruments may be used for the mechanical disruption of the yeast cell wall. Efficiency of cell breakage may be determined by light microscopy.
4. Pellet cells for 1 min at full speed in a microcentrifuge to collapse foam; add 25  $\mu$ L 5 M NaCl and close the tube tightly. Repeat **step 3** with FastPrep. Place the tube on ice for ~1 min to chill and reduce any pressure build-up in the tube for the next step.
5. Make sure the tube cap is tight, wipe off any ice with a kimwipe, invert the tube and flick to cause beads and liquid to drop into cap. Poke two holes near the bottom of the tube using a red-hot 26G hypodermic needle and insert the tube into a 5 mL polypropylene tube (snap-cap tube without cap). Spin 2 min at 2000 rpm (~1000  $\times g$ ) in table-top or clinical, swinging-bucket centrifuge.
6. Transfer sample including pellet into a 2 mL phase-lock tube.
7. *Genomic DNA isolation*: Add equal volume (~500  $\mu$ L) of PCI and extract by shaking vigorously by hand for ~15 s, spin 5 min at full speed in microcentrifuge
8. Repeat **step 7** in the same phase-lock tube and decant the aqueous phase (after centrifugation) into a microcentrifuge tube. Add 1 mL 100%EtOH, mix and spin 10 min at full speed in microcentrifuge.

9. Discard the supernatant, rinse pellet with 1 mL 70%EtOH, and discard (do not loosen pellet; if it comes loose, centrifuge again), and dry pellet briefly.
10. Dissolve pellet in 100  $\mu$ L TE. Heat to 45–50  $^{\circ}$ C to facilitate DNA solubilization, and once pellet is dissolved, add RNaseA to 0.2 mg/mL, and incubate for 30 min at 37  $^{\circ}$ C.
11. Add 10  $\mu$ L 10 $\times$  SDS-PK and incubate for 30 min at 50  $^{\circ}$ C.
12. Purify over QIAquick PCR purification column according to the Qiagen protocol, except elute in 120  $\mu$ L pre-warmed (55  $^{\circ}$ C) elution buffer, EB.
13. Measure DNA concentration with Nanodrop or other instrument requiring minimal sample volume. The yield from 20 mL culture ( $OD_{600} \sim 1$ ) should be at least 2  $\mu$ g.
14. *DNA fragmentation*: Transfer sample to Covaris microTUBE and fragment DNA using the Covaris S2 sonicator (duty cycle = 10%, intensity = 4, 200 cycles per burst for 120 s). These settings yield DNA sheared to an average of 300–400 bp (*see Note 2*).
15. Transfer the sheared DNA to a new 1.7 mL microcentrifuge tube.

### 3.3 DNA Barcoding/ Adapter Ligation

Samples are now end-labeled with unique barcoded adapters that are compatible with the subsequent library-amplification and sequencing primers for the sequencing platform to be used (i.e., Illumina). These barcodes will uniquely identify the source sample for every DNA fragment, allowing the samples to be pooled for all subsequent operations.

1. *End repair and A-tailing*: We recommend using KAPA Hyper prep kit (Kapa Biosystems, KR0961-v3.15) according to the Kapa protocol, except that we carry out half reactions. In a 0.2 mL PCR tube, on ice, combine the following:

0.5 $\mu$ g fragmented, genomic DNA (use H <sub>2</sub> O to adjust volume)	25 $\mu$ L
End-repair and A-tailing buffer	3.5 $\mu$ L
End-repair and A-tailing enzyme mix	1.5 $\mu$ L
Total	30 $\mu$ L

2. Mix by pipetting and immediately place in thermocycler and incubate for 30 min at 20  $^{\circ}$ C, followed by 30 min at 65  $^{\circ}$ C, and hold at 4  $^{\circ}$ C.
3. *Adaptor Ligation*: To the End-repair and A-tailing reaction (on ice or in 4  $^{\circ}$ C thermocycler block) add each of the following reagents from the KAPA kit:

Water	2.5 $\mu$ L
Ligation buffer	15 $\mu$ L
Adapter (15 $\mu$ M stock)	2.5 $\mu$ L
DNA ligase	5 $\mu$ L

- Mix by pipetting and incubate for 15 min at 20 °C in thermocycler.
- Purify each sample using the QIAquick PCR purification kit according to the Qiagen protocol, eluting in 50  $\mu$ L pre-warmed (55 °C) EB.
- Measure DNA concentration with Nanodrop.

### 3.4 Validation of Adapter Ligation

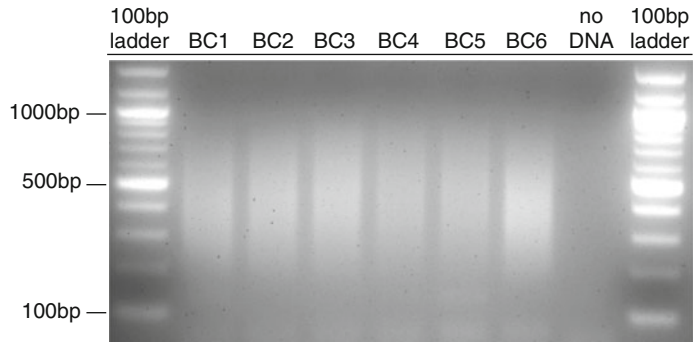
Before proceeding to pooling and IP, the critical step of adapter ligation is confirmed for individual samples by PCR with library amplification primers. Whereas it is sufficient to ensure that individual samples amplify with qualitatively similar efficiencies as described here, quantitative PCR may be used to achieve more exact measurements and potentially to adjust amounts for pooling in the next section.

- Pilot amplification with adapter-compatible primers.  
In separate 0.2 mL PCR tubes on ice, place 1  $\mu$ L of each barcoded sample; prepare an additional tube containing 1  $\mu$ L of H<sub>2</sub>O for “no DNA” control.
- In a microcentrifuge tube on ice, prepare the following reaction mix multiplied by the number of samples to be examined plus a few extra for the “no DNA” control and to account for pipetting losses.

H <sub>2</sub> O	16.75 $\mu$ L
5 $\times$ Phusion HF buffer	5 $\mu$ L
Multi/Std primer 1.0 (25 $\mu$ M)	0.5 $\mu$ L
Multiplexing primer 2.0 (0.5 $\mu$ M)	0.5 $\mu$ L
Index primer (25 $\mu$ M)	0.5 $\mu$ L
10 mM dNTP mix	0.5 $\mu$ L
Phusion DNA polymerase	0.25 $\mu$ L
Total	24 $\mu$ L

PCR mix (per sample):

- Dispense 24  $\mu$ L of the PCR mix to each of the PCR sample tubes.
- Place samples in thermocycler and carry out the following PCR protocol:



**Fig. 2** Validation of adapter ligation. PCR analysis using Illumina amplification primers as described in Subheading 3.4 was carried out on six different adapter ligation reactions, each using a different barcoded adapter (BC1-BC6); an additional PCR reaction was performed without added template (no DNA). After addition of gel loading dye, half of each PCR reaction was subjected to electrophoresis on a 2% agarose gel.

Step 1: 98 °C 1 min.

Step 2: 98 °C 10s.

Step 3: 65 °C 30s.

Step 4: 72 °C 45 s.

Step 5: Repeat **steps 2–4**, 14–17 more times (use the minimum number of cycles necessary to amplify enough DNA for quality control and sequencing below).

Step 6: 72 °C 5 min.

Step 7: 4 °C hold.

5. Analyze PCR products on agarose gel (*see* Fig. 2).

### 3.5 Pooling and BrdU Immunoprecipitation

Samples are pooled for IP at a DNA concentration of 1 µg/mL. 100 ng of each barcoded sample is recommended, but should be scalable. IP volume should be scaled according to number of samples pooled, and amount used per sample. It is not necessary to pool equal amounts of each sample, as parallel analysis of an aliquot of the pooled Input will allow for correction of DNA amounts through normalization of the data.

1. In a 1.7 mL microcentrifuge tube (or larger tube for greater than 15 samples),

Combine (per sample): 100 ng DNA (adjust volume with H <sub>2</sub> O)	45 µL
10× PBS	5 µL
Total (per sample)	50 µL

Mix thoroughly.

2. Transfer 2% of the total volume of the pooled sample to a new microcentrifuge tube and set-aside until Subheading 3.6; this is the “Input” DNA.
3. Heat the pooled sample from **step 1** for 5 min at 95 °C and immediately place in an ice water bath for 2 min.
4. Add an equal volume of ice-cold 2×IP buffer.
5. Add an appropriate amount of anti-BrdU antibody (e.g., 1:200) and incubate on rotating wheel for 45 min at 4 °C.
6. Add 30 µL of Protein G-Dynabeads (resuspend the slurry immediately before use) and incubate on rotating wheel for 30 min at 4 °C. Place the tube on a magnetic rack to collect beads, and discard buffer by decanting and/or aspirating, being careful not to aspirate beads.
7. Remove the tube from magnet and wash beads by adding 1 mL cold wash buffer by inverting tubes until beads are fully suspended. Place the tube on a magnetic rack to collect beads and remove wash buffer as above. Repeat this step two more times with wash buffer. (This and the following steps may be performed at room temperature with cold buffer).
8. Wash beads once as in **step 7** with 1 mL ice-cold TE.
9. Elute sample by resuspending beads in 100 µL elution buffer and incubating 10 min at 65 °C in heat block or water bath.
10. Place the tube on a magnetic rack to collect beads and transfer the eluate to a new microcentrifuge tube.
11. Purify with QIAquick PCR purification kit according to Qia-gen protocol, except elute from column with 30 µL pre-warmed H<sub>2</sub>O (65 °C).

### 3.6 Library Amplification

High-fidelity Phusion polymerase is used for amplification; Kapa polymerase is also recommended. The IP and Input DNA samples are amplified in separate reactions with a different index primer to allow combined sequencing on Illumina.

1. Library amplification with Illumina Indexed Primers.

In a 0.2 mL PCR tube on ice, combine the following and mix:

DNA (volume adjusted with H <sub>2</sub> O)	34.5 µL
5× Phusion HF buffer	10 µL
Multi/Std primer 1.0 (25 µM)	1 µL
Multiplexing primer 2.0 (0.5 µM)	1 µL
Index primer* (25 µM)	1 µL

\*a different Index primer is used for each reaction

10 mM dNTP mix	1.5 $\mu$ L
Phusion DNA polymerase	1 $\mu$ L
Total	50 $\mu$ L

- Place samples in thermocycler and carry out the following PCR protocol:

Step 1: 98 °C 1 min.

Step 2: 98 °C 10 s.

Step 3: 65 °C 30 s.

Step 4: 72 °C 45 s.

Step 5: Repeat **steps 2–4**, 14–17 more times (use the minimum number of cycles necessary to amplify enough DNA for quality control and sequencing below).

Step 6: 72 °C 5 min.

Step 7: 4 °C hold.

- DNA Purification with AMPure beads.

AMPure beads are effective at removal of primers and larger primer dimers, which are common by-products of PCR that should be removed prior to DNA sequencing.

Transfer the library amplification reaction to a 1.7 mL microcentrifuge tube and purify DNA according to the AMPure protocol using an equal volume of AMPure bead slurry, and final elution with 20  $\mu$ L of pre-warmed TE (65 °C). Use the magnetic rack as described for the Dynabeads in Subheading 3.5.

- Measure DNA concentration with Nanodrop.

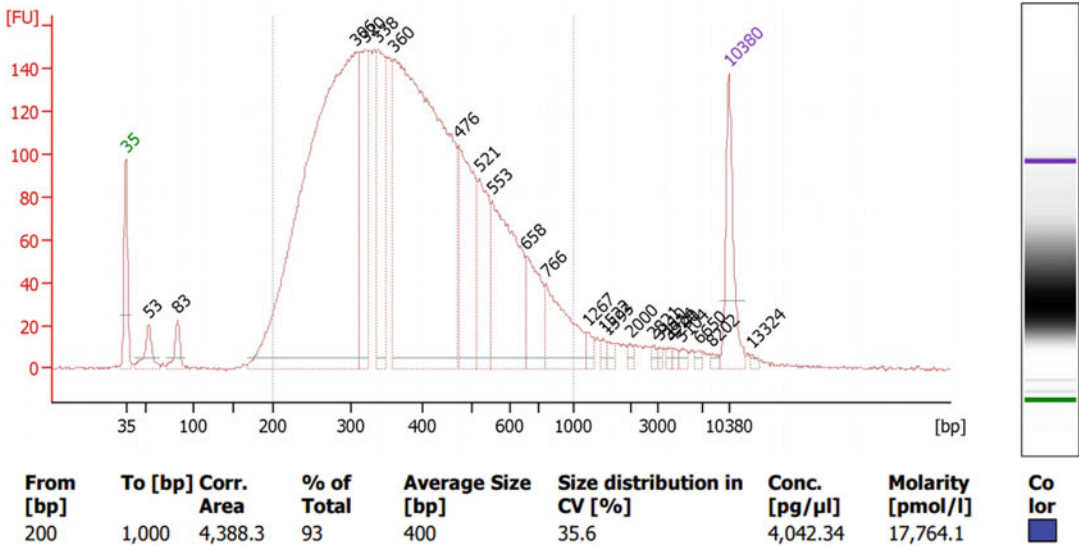
### 3.7 Quality Control and Quantification

The Agilent Technologies Bioanalyzer permits quantification and determination of DNA fragment size distribution of the amplified samples. This analysis confirms the library quality based on the size distribution, expected to show an average fragment size of ~400 bp, as the linker and primer additions add ~130 bp (Fig. 3). Common artifacts such as primer dimers that will consume sequence reads will be revealed, allowing their possible elimination through further purification using AMPure beads. Quantification at this step allows adjustment of sample concentrations for the desired sequencing platform.

- Analyze up to 10 ng DNA on Bioanalyzer (*see Note 6*).

### 3.8 DNA Sequencing

- Sequence samples using an Illumina instrument (*see Note 7*). 10 ng (~2 ng/ $\mu$ L) of DNA is sufficient.



**Fig. 3** Quality control and quantification. 1–10 ng of the amplified and purified library from the completion of Subheading 3.6 was run on a Bioanalyzer chip (High Sensitivity DNA Assay). The tracing shows Intensity in fluorescence units (FU) on the y-axis and DNA length in base-pairs (bp) on the x-axis of the gel image to the right. The lower (35 bp) and upper molecular weight markers (10,380 bp) are indicated in *green* and *purple* respectively. Results are summarized below the graph indicating an average fragment size (400 bp) and concentration (~4 ng/μL).

### 3.9 Data Analysis

Sequenced reads are processed using a series of software programs that are freely available online. The following analysis pipeline allows for the extraction of the data, mapping the reads to the reference genome, quantitative normalization, and data display.

1. *ea-utils*: Sequenced reads are first processed using the fastq-multx tool [15, 16]. This de-multiplexes the reads based on the specified barcode sequences. The first five bases are trimmed from the 5' end of each read to remove the barcodes for genome alignment.
2. *bowtie2*: The *S. cerevisiae* genome sequence is accessed from Saccharomyces Genome Database ([www.yeastgenome.org](http://www.yeastgenome.org)) and indexed using bowtie2, *bowtie2-build* [17]. The raw reads are then aligned to the indexed genome using bowtie2.
3. *Samtools*: Filter reads that map to more than one region in the genome using *samtools view*. Sort by coordinates of the reference genome using *samtools sort*. Remove PCR duplicates using *samtools rmdup*.
4. *BEDtools*: Change file format by converting files from .bam to .bed files. Bin the aligned reads into 50 bp nonoverlapping bins using *BEDtools* [18, 19].
5. *QBU normalization*: Binned read counts for IP are divided by the corresponding binned read counts for Input, for the



corresponding samples (i.e., IP and Input with the same barcodes). The resulting data is median-smoothed over a 1000 bp window for both IP and Input. Replicate data sets may be averaged for plotting and further analysis. The values from the IP/Input normalization are unitless but are directly comparable between samples in the same pool. Optional: We find it useful to scale the data from each pool to a maximum value of “1” for one of the samples (e.g., the wild-type sample). This is achieved as follows: Any replicates are first averaged. The maximum bin value in the wild-type data set is determined, and all bins for all samples in that pool are divided by that maximum bin value.

*Additional analyses:* MACS may be used for peak-calling, using filtered, aligned reads [20]; Diffbind (using DESeq) may be used for analysis for differential data signal (e.g., IP signal) [21, 22].

---

## 4 Notes

1. Alternative kits or methods for DNA purification may be substituted.
2. Alternative sonicators including probe sonicators may be used, but will require prior determination of conditions for optimal DNA shearing. This protocol targets ~300 bp average length. DNA may also be fragmented by enzymatic digestion (e.g., KAPA Hyper Plus kit).
3. Alternative library preparation kits or individual reagents may be used for Illumina sequencing; alternative sequencing platforms will require alternative library preparation reagents or kits.
4. Oligonucleotide pairs must be annealed for ligation. Oligonucleotide pairs are diluted (from 100  $\mu\text{M}$  stocks in TE) to 15  $\mu\text{M}$  using annealing buffer (10 mM Tris-HCl pH 7.6, 50 mM NaCl, 1 mM EDTA). Place in heat block 5 min at 95  $^{\circ}\text{C}$ . Remove the heat block (with samples) from the heating unit and allow to cool to room temperature. A thermocycler may be used with a cooling rate of  $\sim 2$   $^{\circ}\text{C}$  per minute; store at  $-20$   $^{\circ}\text{C}$ .
5. Protein A- and anti-IgG-coupled beads may also be used as appropriate for the anti-BrdU antibody used; sepharose beads may be used instead of magnetic beads.
6. The Bioanalyzer allows electrophoretic analysis of nanogram quantities of DNA. If a Bioanalyzer is not available, larger amounts on the order of 100 ng may be analyzed by agarose gel electrophoresis using a sensitive dye such as SYBR Green (Molecular Probes) and an appropriate scanner. Real-time PCR (using the library amplification primers) may also be used for

quantification, and PCR products may be analyzed by gel electrophoresis.

7. Other sequencing platforms may be used but will require modification(s) of the method, particularly the library construction steps.
8. Culturing conditions, including media, temperature, synchronization method, etc., may be modified according to experimental or strain-specific requirements. DNA content analysis by flow cytometry and/or morphological analysis of budding is/are recommended to confirm expected cell cycle synchronization and progression.
9. Monitor cells microscopically during this incubation to confirm synchronization and determine the appropriate time for release into S phase (remove ~0.5 mL of culture and sonicate gently to disaggregate cells before placing ~10  $\mu$ L on slide with cover slip). For  $\alpha$ -factor-arrested cells, unbudded cells will accumulate (>95% for *bar1* $\Delta$  strains after 3–4 h) and change morphologically from roughly spherical- or oval-shaped cells to elongated, pear-shaped cells called “schmoos.” Cells are held arrested ~4 h to allow daughter cells to grow sufficiently to release synchronously with the larger mothers.
10. Sonicator settings will depend on the instrument and the volume of the culture and the centrifuge tube. For example, for a 15 mL culture volume in a 50 mL tube, the following settings are used for a Branson 250 Sonicator: power: 2.5, duty cycle: 30%, six pulses.
11. For hydroxyurea-treated cells, 60–90 min is most commonly used. For time-course experiments, a large culture may be continuously labeled with BrdU and sampled at regular intervals, such as 10 min; alternatively, cells from a large unlabeled culture (i.e., lacking BrdU) may be pulse-labeled for defined intervals (e.g., 10–15 min) by incubating aliquots removed from the culture with BrdU.

---

## Acknowledgments

We thank current and previous members of the Aparicio lab for comments on the protocol. Research in the Aparicio lab has been funded by NIH R01 GMS05494.

## References

1. Dolbeare F (1996) Bromodeoxyuridine: a diagnostic tool in biology and medicine, Part III. Proliferation in normal, injured and diseased tissue, growth factors, differentiation, DNA replication sites and in situ hybridization. *Histochem J* 28(8):531–575
2. Lengronne A, Pasero P, Bensimon A, Schwob E (2001) Monitoring S phase progression

- globally and locally using BrdU incorporation in TK(+) yeast strains. *Nucleic Acids Res* 29 (7):1433–1442
3. Vernis L, Piskur J, Diffley JF (2003) Reconstitution of an efficient thymidine salvage pathway in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31(19):e120
  4. Viggiani CJ, Aparicio OM (2006) New vectors for simplified construction of BrdU-incorporating strains of *Saccharomyces cerevisiae*. *Yeast* 23(14–15):1045–1051
  5. Vickers MF, Mani RS, Sundaram M, Hogue DL, Young JD, Baldwin SA, Cass CE (1999) Functional production and reconstitution of the human equilibrative nucleoside transporter (hENT1) in *Saccharomyces cerevisiae*. Interaction of inhibitors of nucleoside transport with recombinant hENT1 and a glycosylation-defective derivative (hENT1/N48Q). *Biochem J* 339(Pt 1):21–32
  6. Hodson JA, Bailis JM, Forsburg SL (2003) Efficient labeling of fission yeast *Schizosaccharomyces pombe* with thymidine and BUdR. *Nucleic Acids Res* 31(21):e134
  7. Sivakumar S, Porter-Goff M, Patel PK, Benoit K, Rhind N (2004) In vivo labeling of fission yeast DNA with thymidine and thymidine analogs. *Methods* 33(3):213–219. doi:10.1016/j.ymeth.2003.11.016
  8. Green MD, Sabatinos SA, Forsburg SL (2015) Microscopy techniques to examine DNA replication in fission yeast. *Methods Mol Biol* 1300: 13–41. doi:10.1007/978-1-4939-2596-4\_2
  9. Bianco JN, Poli J, Saksouk J, Bacal J, Silva MJ, Yoshida K, Lin YL, Tourriere H, Lengronne A, Pasero P (2012) Analysis of DNA replication profiles in budding yeast and mammalian cells using DNA combing. *Methods* 57(2):149–157. doi:10.1016/j.ymeth.2012.04.007
  10. Peace JM, Ter-Zakarian A, Aparicio OM (2014) Rif1 regulates initiation timing of late replication origins throughout the *S. cerevisiae* genome. *PLoS One* 9(5):e98501. doi:10.1371/journal.pone.0098501
  11. Ryba T, Battaglia D, Pope BD, Hiratani I, Gilbert DM (2011) Genome-scale analysis of replication timing: from bench to bioinformatics. *Nat Protoc* 6(6):870–895. doi:nprot.2011.328 [pii] 1038/nprot.2011.328
  12. Viggiani CJ, Knott SR, Aparicio OM (2010) Genome-wide analysis of DNA synthesis by BrdU immunoprecipitation on tiling microarrays (BrdU-IP-chip) in *Saccharomyces cerevisiae*. *Cold Spring Harb Protoc* 2010(2.) pdb prot5385. doi:2010/2/pdb.prot5385 [pii] 1101/pdb.prot5385
  13. Rotem A, Ram O, Shores N, Sperling RA, Goren A, Weitz DA, Bernstein BE (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 33 (11):1165–1172. doi:10.1038/nbt.3383
  14. Peace JM, Villwock SK, Zeytounian JL, Gan Y, Aparicio OM (2016) Quantitative BrdU immunoprecipitation method demonstrates that Fkh1 and Fkh2 are rate-limiting activators of replication origins that reprogram replication timing in G1 phase. *Genome Res* 26 (3):365–375. doi:10.1101/gr.196857.115
  15. Aronesty E (2011) Command-line tools for processing biological sequencing data. <https://github.com/ExpressionAnalysis/ea-utils>.
  16. Aronesty E (2013) Comparison of sequence utility programs. *Open Bioinform J* 7:1–8. doi:10.2174/1875036201307010001
  17. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* 9 (4):357–359. doi:10.1038/nmeth.1923
  18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16):2078–2079. doi:10.1093/bioinformatics/btp352
  19. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842. doi:10.1093/bioinformatics/btq033
  20. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9 (9):R137. doi:10.1186/gb-2008-9-9-r137
  21. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106. Doi:gb-2010-11-10-r106 [pii] 1186/gb-2010-11-10-r106
  22. Stark R, Brown R (2013) DiffBind: differential binding analysis of ChIP-Seq peak data. <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>
  23. Dunham JP, Friesen ML (2013) A cost-effective method for high-throughput construction of illumina sequencing libraries. *Cold Spring Harb Protoc* 2013(9):820–834. doi:10.1101/pdb.prot074187

# Chapter 17

## Strand-Specific Analysis of DNA Synthesis and Proteins Association with DNA Replication Forks in Budding Yeast

Chuanhe Yu, Haiyun Gan, and Zhiguo Zhang

### Abstract

DNA replication initiates at DNA replication origins after unwinding of double-strand DNA(dsDNA) by replicative helicase to generate single-stranded DNA (ssDNA) templates for the continuous synthesis of leading-strand and the discontinuous synthesis of lagging-strand. Therefore, methods capable of detecting strand-specific information will likely yield insight into the association of proteins at leading and lagging strand of DNA replication forks and the regulation of leading and lagging strand synthesis during DNA replication. The enrichment and Sequencing of Protein-Associated Nascent DNA (eSPAN), which measure the relative amounts of proteins at nascent leading and lagging strands of DNA replication forks, is a step-wise procedure involving the chromatin immunoprecipitation (ChIP) of a protein of interest followed by the enrichment of protein-associated nascent DNA through BrdU immunoprecipitation. The isolated ssDNA is then subjected to strand-specific sequencing. This method can detect whether a protein is enriched at leading or lagging strand of DNA replication forks. In addition to eSPAN, two other strand-specific methods, (ChIP-ssSeq), which detects potential protein-ssDNA binding and BrdU-IP-ssSeq, which can measure synthesis of both leading and lagging strand, were developed along the way. These methods can provide strand-specific and complementary information about the association of the target protein with DNA replication forks as well as synthesis of leading and lagging strands genome wide. Below, we describe the detailed eSPAN, ChIP-ssSeq, and BrdU-IP-ssSeq protocols.

**Key words** Chromatin immunoprecipitation (ChIP), BrdU immunoprecipitation, Single-strand DNA library preparation, Next-generation sequencing, eSPAN, ChIP-ssSeq and BrdU-IP-ssSeq

---

## 1 Introduction

Eukaryotic DNA replication initiates at multiple DNA replication origins in a tempo-spatial order [1]. Once started, DNA synthesis proceeds bi-directionally, with continuous synthesis of leading strand DNA and discontinuous synthesis of lagging strand or Okazaki fragment [2]. Multiple proteins, including DNA polymerases, replicative DNA helicase MCM complexes, DNA ligase, proliferating cell nuclear antigen (PCNA), and single-strand DNA binding protein (SSB), are involved in DNA synthesis [3–9]. Chromatin immunoprecipitation (ChIP) coupled with PCR analysis and/or

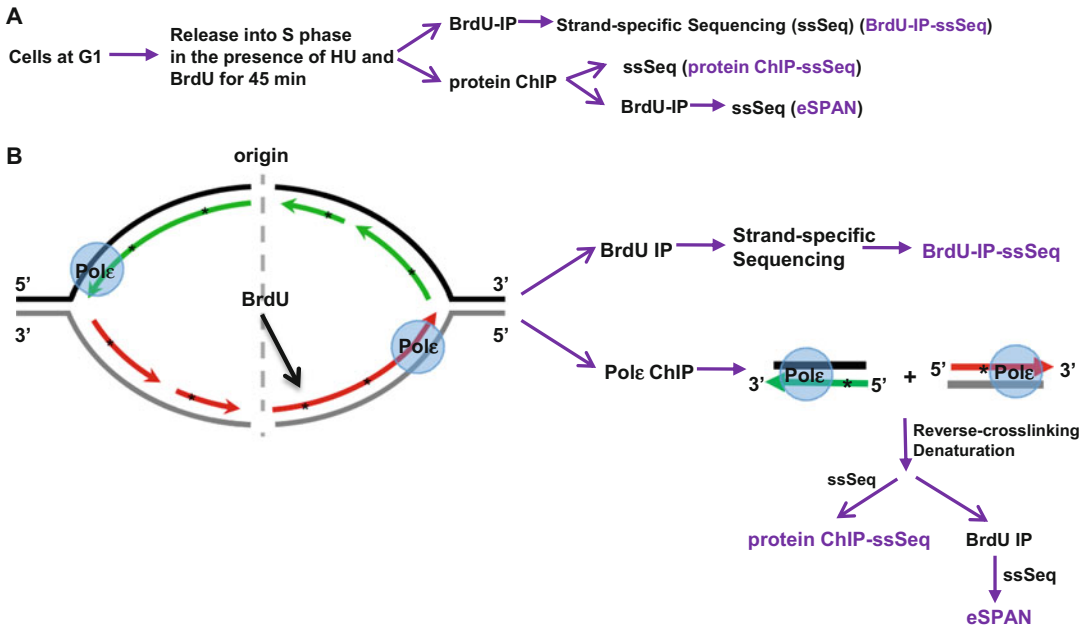
microarray has been used to map DNA replication origins and to monitor the replication proteins traveling along with the replication fork [10, 11]. All these methods cannot differentiate leading strand from lagging strand. We reason that monitoring of DNA synthesis and the association of a protein of interest with DNA replication forks in a strand-specific manner will yield novel insight into the DNA replication process. Therefore, we developed the eSPAN method (enrichment and Sequencing of Protein-Associated Nascent DNA) [12] to measure the relative amount of protein at leading and lagging strands DNA replication forks (Fig. 1). Using this method, we show that leading strand DNA Polymerase  $\epsilon$  and lagging strand polymerase  $\delta$  [13] associate preferentially with leading and lagging strands of DNA replication forks, respectively (Fig. 2) [14]. We further showed that PCNA is unloaded from lagging strands when replication forks stall. These results demonstrate the efficiency and power of eSPAN. During the development of the eSPAN method, we also performed nascent DNA strand-specific sequencing (BrdU-IP-ssSeq) and targeted protein ChIP strand-specific sequencing (ChIP-ssSeq) (Fig. 1). BrdU-IP-ssSeq can monitor the leading and lagging strands synthesis independently; whereas ChIP-ssSeq can monitor replication protein travel on the template strand and detect potential protein-ssDNA binding. These combined methods provide unprecedented information on the synthesis of leading and lagging strands of DNA replication forks and the association of proteins with leading and lagging strands of DNA replication forks. Here, we describe the protocols for ChIP-ssSeq, BrdU-IP-ssSeq, and eSPAN in budding yeast. Targeted protein ChIP, labeling DNA with BrdU, and BrdU immunoprecipitation protocols were adapted from [15, 16] and the ssDNA library preparation protocol was adapted from [17, 18].

---

## 2 Materials

### 2.1 *Yeast Cell Culture and Sample Collection*

1. YPD medium: Bacto peptone 20 g, yeast extract 10 g, glucose 20 g and transfer to a 1 l cylinder. Add 950 ml water. Stir until fully suspended. Make up to 1 l with water. Autoclave.
2. Yeast strain (W303A strain).
3. Alpha factor (WHWLQLKPGQPMY) stock: 5 mg/ml. Dissolve 50 mg in 10 ml 100% ethanol. Store at  $-20^{\circ}\text{C}$ .
4. Tris-buffered saline (TBS;  $10\times$ ): 1.5 M NaCl, 0.1 M Tris-HCl, pH 7.6. 24.2 g Tris base and 80 g NaCl. Transfer to a 1 l cylinder. Add 900 ml and adjust pH to 7.6 with HCl. Stir until fully suspended. Make up to 1 l with water. Autoclave.
5. PMSF ( $100\times$ ): Dissolve 17.4 mg PMSF in 1 ml isopropanol.

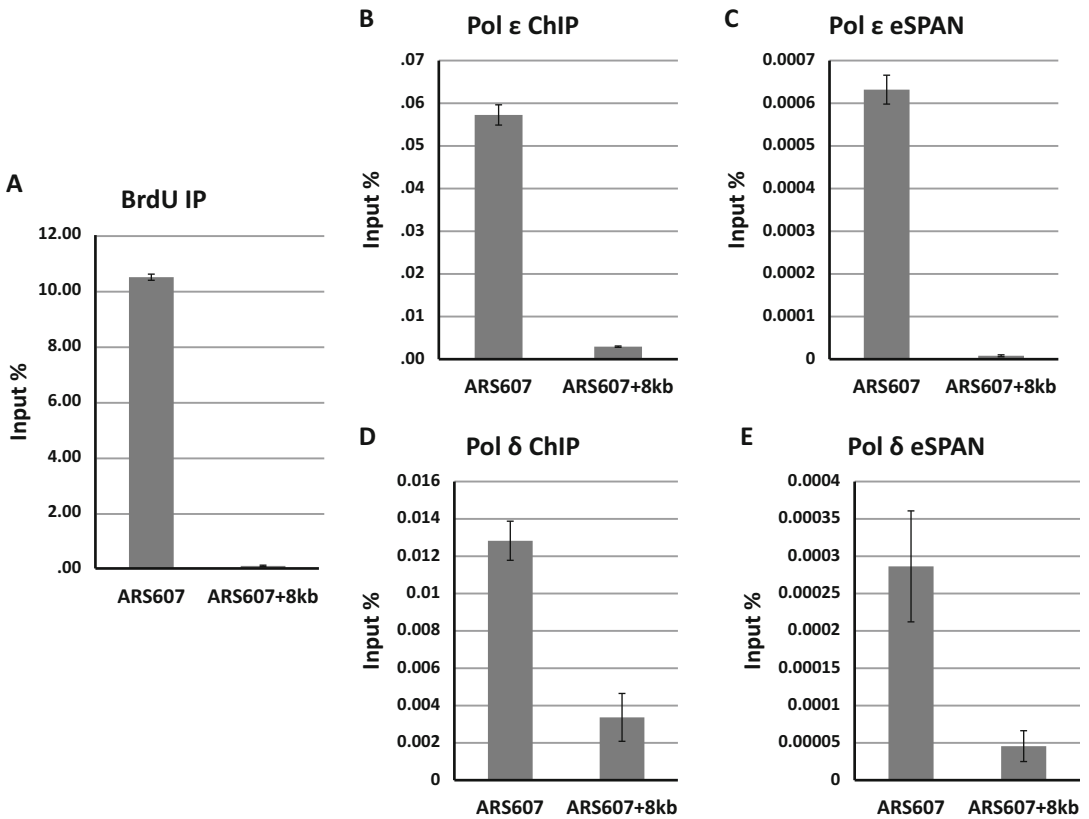


**Fig. 1** An outline of experimental procedures of BrdU IP-ssSeq, ChIP-ssSeq, and eSPAN. (a) A flowchart and (b) a schematic of the experimental procedures. Yeast cells arrested in G1 were released into fresh medium containing HU and BrdU at 30 °C for 45 minutes. Early S phase cells were collected for BrdU-IP-ssSeq, ChIP-ssSeq, and eSPAN. After cells were physically broken and chromatin was sheared by sonication, a fraction of the chromatin (2–5%) was saved as input as well as for BrdU-immunoprecipitation (BrdU-IP). The rest of the chromatin (95–98%) was used for ChIP of the protein of interest. After the purification of ChIP DNA, ChIP DNA was separated into two parts, one (13%) was used directly for ssDNA library preparation (ChIP-ssSeq), the rest (87%) was used for the enrichment and Sequencing of Protein-Associated Nascent DNA (eSPAN) procedure. For each experiment, four DNA samples including input DNA, BrdU IP DNA, ChIP-ssSeq DNA, and eSPAN DNA were used to prepare libraries following the single-stranded DNA library preparation protocol. The sequencing reads of Watson strand (*red*) and Crick strand (*green*) were independently mapped to the yeast reference genome

6. TBS (1×) + PMSF buffer (*see Note 1*): dilute 10× TBS by adding nine times water in volume, then add 1/100 volume PMSF(100×) solution.
7. BrdU (e.g., Sigma B5002).

## 2.2 Chromatin Immunoprecipitation

1. PBS (10×): 1.37 M NaCl, 27 mM KCl, 80 mM Na<sub>2</sub>HPO<sub>4</sub>, and 20 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.4. 25.6 g of Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 80 g of NaCl, 2 g of KCl and 2 g of KH<sub>2</sub>PO<sub>4</sub>. Make up to 1 l with water. Autoclave.
2. NaOH (10 N): Add 40 g NaOH to a suitable container. Add distilled water to make solution up to 100 ml.
3. 37% PFA (freshly made): Weigh 4.7 g of paraformaldehyde and transfer to a 50 ml falcon tube. Add 25 μl of NaOH (10 N) and 7.5 ml of 1× PBS buffer. Put the tube into a cup containing 2/



**Fig. 2** Newly synthesized DNA and DNA Polymerases  $\epsilon$  and  $\delta$  are enriched at replicating DNA using real-time PCR analysis. (a) BrdU Immunoprecipitation (IP) assay showed that BrdU was incorporated at the early replication origin *ARS607*, but not 8-kb away from *ARS607* (*ARS607* + 8 kb). (b) and (c) DNA Pol  $\epsilon$  and  $\delta$  associate with replicating DNA. (d) and (e) QPCR analysis of Pol  $\epsilon$  and Pol  $\delta$  eSPAN DNA. ChIP assays were performed using antibodies against FLAG-tagged Pol2 or Pol32

3 of boiling water and agitate the tube once every 1 min until PFA is completely dissolved. Bring to 12.5 ml with  $1 \times$  PBS and filter through a  $0.45 \mu\text{M}$  filter. Use on the same day.

- 2.5 M Glycine: dissolve 93.8 g glycine in 500 ml water. Store at room temperature after autoclaving.
- Glass beads, 425–600  $\mu\text{m}$  in diameter (e.g., Sigma G8772), acid washed and heat-sterilized.
- Hypodermic needles (16 gauge  $\times$  1 in).
- 1 M HEPES/KOH, pH 7.5: Add 700 ml water to 238.3 g HEPES in a 1 l cylinder, stir and adjust pH to 7.5 with KOH. Make up to 1 l with  $\text{H}_2\text{O}$ , filter and store at  $-20^\circ\text{C}$ .
- ChIP lysis buffer: 0.1 M HEPES/KOH, pH 7.5; 0.1 M NaCl; 0.01 M EDTA; 1% TX-100; 1% Na-deoxycholate (deoxycholic acid, sodium salt). Add 50 ml 1 M HEPES/KOH, pH 7.5; 10 ml 5 M NaCl; 1 ml 0.5 M EDTA; 5 ml TX-100 and 0.5 g

Na-deoxycholate to a 1 l cylinder. Stir and make to 500 ml with water. Store at 4 °C.

9. Protein G-Sepharose beads (GE).
10. Protease inhibitor and antibiotic mix: PMSF (100×): add 17.4 mg to 1 ml isopropanol and mix until fully dissolved; AEBSF (4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride; 200×): add 50 mg into 1 ml water and mix until fully dissolved; Benzamidine (1000×): add 157 mg benzamidine into 1 ml water and mix until fully dissolved; Bacitracin (100×): 100 mg bacitracin in 1 ml water and mix until fully dissolved. Store all at -20 °C.
11. 1 M Tris, pH 8.0: Add 121.1 g Tris base to 0.8 l water, adjust pH to 8.0 with HCl. Bring up to 1 l.
12. Tris/LiCl buffer: 0.5 M LiCl, 0.5% NP-40, 0.1% sodium deoxycholate, 0.1 M Tris-HCl (pH 8.0). For 500 ml, add 50 ml of 1 M Tris-HCl pH 8.0, 50 ml of 5 M LiCl (for 5 M LiCl, dissolve 21.2 g LiCl in 90 ml distilled water. Adjust the volume to 100 ml. Store at 4 °C, 2.5 ml of 100% Nonidet P-40, and 0.5 g of sodium deoxycholate. Adjust the volume to 500 ml with distilled water.
13. ChIP lysis buffer + 0.5 M NaCl: add 72 µl 5 M NaCl to 1 ml ChIP lysis buffer.
14. Tris/EDTA washing buffer: 50 mM Tris, pH 8.0; 10 mM EDTA. Mix 25 ml 1 M Tris, pH 8.0 and 10 ml 0.5 M EDTA. Bring to 500 ml with H<sub>2</sub>O.
15. 2× TE buffer: 20 mM Tris, pH 8.0, 2 mM EDTA. 10 ml of 1 M Tris (pH 8.0) and 2 ml EDTA (0.5 M) are made to 100 ml with distilled water.
16. Anti-flag antibody or specific antibody for your protein of interest.
17. 10% Chelex-100 (BioRad): Add 1 g Chelex-100 to 10 ml water. It can be stored at room temperature for several months.
18. Covaris E210 sonicator, or equivalent.

### **2.3 BrdU Immunoprecipitation**

1. BrdU IP buffer: dilute 10× PBS by adding nine volumes of water, and then add Triton X-100 to 0.0625% (v/v).
2. 1× TE buffer: dilute 2× TE buffer by adding an equal volume of distilled water.
3. TE + 1% SDS buffer: Mix 5 ml 10% SDS, 25 ml 2× TE buffer, and 20 ml distilled water.
4. BrdU antibodies (BD bioscience Cat. No. 555627).
5. E.Coli tRNA (e.g., Sigma Cat. No. R1753).
6. MiniElute PCR Purification Kit (QIAGEN Cat. No. 28006).
7. SYBR Green PCR Master MIX (ThermoScientific 4,309,155).



### 3 Methods

#### 3.1 Yeast Cell Synchronization and Sample Collection (See Note 2)

1. Grow 10 ml of yeast cells on a shaking platform at 30 °C overnight.
2. Dilute into 100 ml fresh YPD medium to OD<sub>600</sub> ~ 0.2 in the morning. When it reaches OD<sub>600</sub> ~ 0.4–0.5, add 100 µl  $\alpha$ -factor (5 mg/ml) and continue to culture at 25 °C (yeast cells are released more synchronously at 25 °C than at 30 °C). After 1.5 h, add another 100 µl  $\alpha$ -factor (5 mg/ml) and culture yeast cells for another 1.5 h. During that time, prepare the 37.5% paraformaldehyde solution and warm up 100 ml YPD + 1.52 g hydroxyurea + 40 mg BrdU medium to 30 °C.
3. 15 min before the 3 h arrest, check the cells with a light microscope to make sure that at least 90% yeast cells are in G1 phase (pear shaped in morphology).
4. Pellet the cells by centrifugation at 1600  $\times g$  for 10 min at 4 °C. Wash twice with 20 ml cold H<sub>2</sub>O. Resuspend cells in pre-warmed 100 ml YPD containing 200 mM HU and 400 mg/l BrdU.
5. Continue culturing cells on a shaker for 45 min at 30 °C, add 3 ml fresh 37% paraformaldehyde, and incubate with shaking for another 20 min at 25 °C.
6. Quench formaldehyde with 6 ml 2.5 M glycine for 5 min at 25 °C on a shaker.
7. Pellet the cells at 1600  $\times g$  for 10 min. Then wash the cells twice with 20 ml cold TBS buffer containing 1 mM PMSF. Resuspend the cell pellet in 1 ml TBS + PMSF buffer and transfer into two 2 ml tubes (50 ml cell/tube) with screw cap. Pellet the cells, remove the supernatant, and store at –80 °C.

#### 3.2 ChIP and DNA Extraction

1. Dissolve cell pellets in 0.1 ml ChIP lysis buffer with protease inhibitors and antibiotic mix and add ~100 µl glass beads. Lyse by bead beating for 30 s on, 1 min off, in a 4 °C cold room. Repeat bead beating for a total of four times. Punch a hole in the bottom of the tube with a 16-gauge hot needle. Nest the tube into a 1.5 ml empty Eppendorf tube and centrifuge at 1600  $\times g$  for 1 min. All of the liquid and cell debris—without beads—should now be at the bottom of the new tube.
2. Aspirate the supernatant (chromatin is in the cell debris fraction). Resuspend the cell pellets in 0.25 ml ChIP lysis buffer with proteinase inhibitor by pipetting up and down several times.
3. To shear the chromatin, sonicate the resuspended cell lysate with a Covaris E210 (10 s on, 30 s off) for 25 times in the tubes

provided by Covaris (*see Note 3*). Transfer the lysate to a new Eppendorf tube and spin at  $10,000 \times g$  for 5 min at 4 °C.

4. Collect the supernatant and transfer to a new Eppendorf tube. Centrifuge at  $10,000 \times g$  for another 10 min at 4 °C. Save 50  $\mu$ l of the clarified lysate as whole cell DNA input and for the BrdU IP (Fig. 1).
5. For the rest of clarified lysate (about 200  $\mu$ l), add a suitable amount of antibody (*see Note 4*) and incubate on a rotating platform (20–30 rotations per min) at 4 °C overnight.
6. The next day, clarify by centrifuging at  $10,000 \times g$  for 10 min at 4 °C. Wash protein G beads twice with ChIP lysis buffer. Transfer the supernatant into an Eppendorf tube containing 25  $\mu$ l washed protein G beads.
7. Incubate at 4 °C for 2 h on a rotating platform (20–30 rotations per min).
8. Spin at  $1000 \times g$  for 1 min and remove the supernatant. Wash beads with the following buffers in order: (a) 1 ml cold lysis buffer; (b) 1 ml cold lysis buffer with 0.5 M NaCl; (c) 1 ml cold Tris/LiCl buffer; (d) 1 ml room temperature  $1 \times$  Tris/EDTA washing buffer. For each wash, rotate 5 min on a rotating platform, spin the beads at  $1000 \times g$  for 1 min, and remove all liquid using a 16G needle.
9. Add 100  $\mu$ l 10% Chelex to the washed beads. Vortex to mix well.  
For the 50  $\mu$ l clarified lysate input, also add 0.1 ml 10% Chelex.
10. Boil for 10 min in a 100 °C heat block, and then cool at room temperature.
11. Add 5  $\mu$ l 20 mg/ml proteinase K to the reaction tube, mix well, and incubate at 55 °C for 30 min while shaking at 600 rpm.
12. Boil beads for another 10 min in a 100 °C heat block.
13. Spin at a top speed, save the supernatant.
14. Add 100  $\mu$ l  $2 \times$  TE to the pellet, vortex, spin down, save the supernatant.
15. Mix the supernatants taken above (about 180  $\mu$ l). For the ChIP sample, save 30  $\mu$ l for library preparation (ChIP-ssSeq, Fig. 1), the rest (about 150  $\mu$ l) will be subjected to BrdU IP (eSPAN, Fig. 1).
16. For the input samples, also save 30  $\mu$ l for input library preparation (input), the rest 150  $\mu$ l will be used for BrdU IP (BrdU-IP-ssSeq) described below

### 3.3 BrdU IP

1. Samples (including the input and ChIP) (about 150  $\mu$ l) are boiled 3 min in a 100 °C heat block, and immediately cooled on ice for 3 min. Add BrdU IP mix (15  $\mu$ l  $10 \times$  PBS, 1.35 ml

BrdU IP buffer, 0.5  $\mu$ l anti-BrdU Ab and 1  $\mu$ l E.coli tRNA (10 mg/ml)). Rotate for 2 h at 4 °C.

2. Add 20  $\mu$ l protein G sepharose bead slurry (washed twice with cold BrdU IP buffer). Incubate at 4 °C for 1 h on a rotor.
3. Spin at  $800 \times g$  for 1 min and remove the supernatant. Wash the beads three times using 1 ml cold BrdU IP buffer, 3 min incubation each time, then wash beads once with 1 ml TE buffer.
4. Add 100  $\mu$ l elution buffer (1 $\times$  TE + 1%SDS), vortex to mix well. Incubate at 65 °C for 15 min. After a short spin, the supernatant is collected for DNA purification.

### **3.4 ChIP DNA Purification**

1. Purify the recovered DNA (four DNA samples including DNA from the input sample, protein ChIP, BrdU-IP, and eSPAN for each experiment) using the Mini-elute PCR purification kit according to the manufacturer's protocol.
2. After purification, DNA is eluted with 17  $\mu$ l EB buffer. At this point, samples can be stored at 4 °C for several days or at -20 °C for months.

### **3.5 Quantitative PCR Analysis of ChIP, BrdU IP, and eSPAN at DNA Replication Origins**

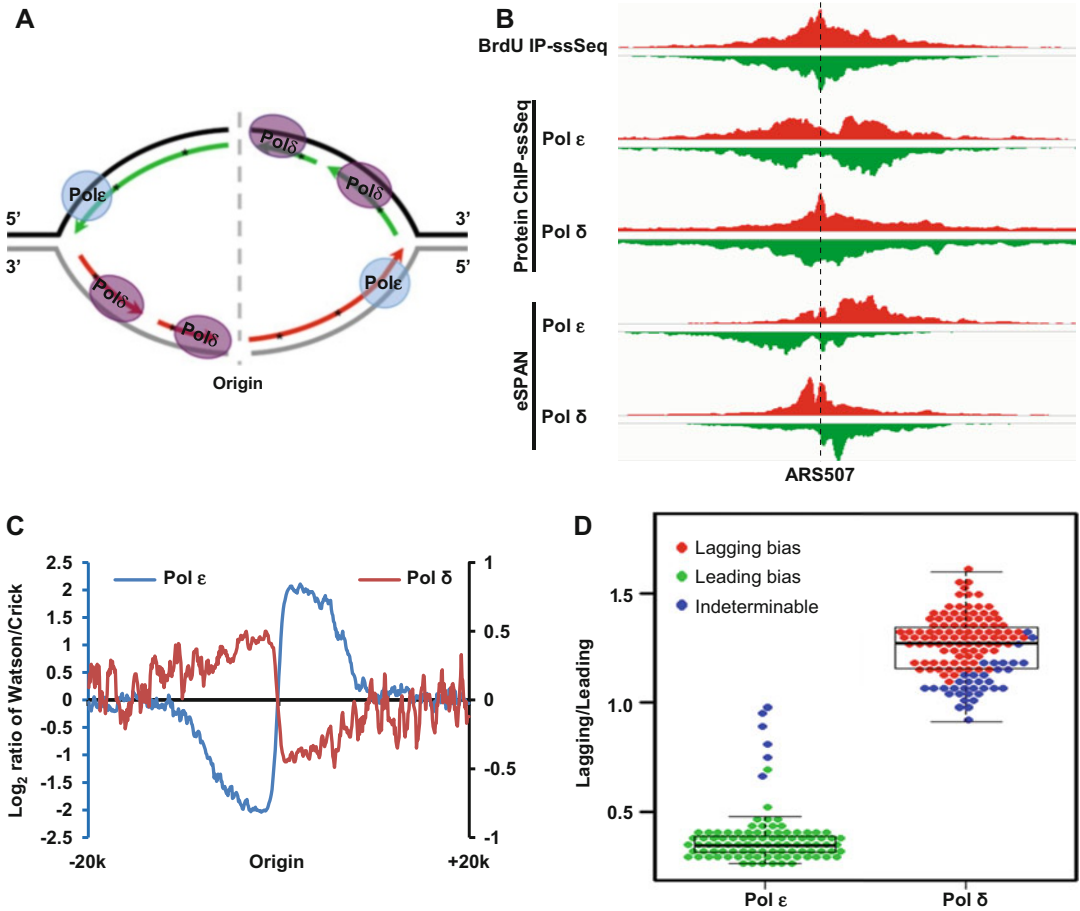
1. Dilute 2  $\mu$ l purified DNA eightfold by adding 14  $\mu$ l of nuclease-free water. Use the diluted DNA for the analysis using real-time PCR. The remaining 15  $\mu$ l of DNA from each sample is used for ssDNA library preparation.
2. Make a 12.5  $\mu$ l reaction with 2.5  $\mu$ l of diluted DNA, 6.25  $\mu$ l SYBR Green PCR Master Mix (ThermoScientific 4,309,155), 0.5  $\mu$ l primer mix (forward and reverse primer at 1.5  $\mu$ M, ARS607 and ARS607 + 8 kb (*see Note 5*)) and 3.25  $\mu$ l water.
3. Perform QPCR: 95 °C for 3 min; 40 cycles of 95 °C for 10 s, 61 °C for 30 s, 72 °C for 35 s; and a final melt curve stage, consisting of 95 °C for 15 s, cooling to 55 °C and 80 repeats of heating for 5 s, starting at 55 °C and with 0.5 °C increments (Fig. 2).

### **3.6 ssDNA Library Preparation and Next-Generation Sequencing**

1. Single-stranded DNA libraries are prepared according to [17]. Omit the endonuclease VIII and UDG treatments.
2. Perform parallel paired-end sequencing with the Illumina Hi-seq 2000 or 2500 platform (*see Note 6*).

### **3.7 Data Analysis**

1. Align the sequencing reads to the yeast genome (sacCer3) using Bowtie2 software (example shown in Fig. 3) [19].
2. Separate the Watson and Crick strand reads by Perl codes, then calculate the genome-wide read coverage of Watson and Crick strands using BEDTools [20].



**Fig. 3** DNA polymerase Pol  $\epsilon$  (Pol  $\epsilon$ ) and DNA polymerase Pol  $\delta$  (Pol  $\delta$ ) showing leading and lagging strands bias at HU-stalled DNA replication forks, respectively. **(a)** A cartoon showing Pol  $\epsilon$  and Pol  $\delta$  bind to the leading and lagging strands, respectively. **(b)** A snapshot of BrdU IP-ssSeq, protein ChIP-ssSeq, and eSPAN peaks at *ARS507* for Pol  $\epsilon$  and Pol  $\delta$ . Red and green colors represent the Watson and Crick strands, respectively. **(c)** The average bias pattern of Pol  $\epsilon$  and Pol  $\delta$  eSPAN peaks at HU-stalled DNA replication forks. A 200-bp sliding window was used to scan from 20Kb upstream to 20 Kb downstream 20 kb of fired replication origins. **(d)** Dot-and-box plot showing the bias pattern of Pol  $\epsilon$  and Pol  $\delta$  eSPAN peaks at the 134 individual early origins

3. Use both Watson and Crick strand reads from BrdU-IP-ssSeq to call replication peaks using MACS software [21].

There are two methods to determine whether a protein of interest is enriched at the leading and lagging strands of DNA replication forks.

4. Analysis of the average bias pattern of eSPAN peaks at all fired origins (Fig. 3c)

(a) Extract the regions with  $\pm 20$  kb surrounding the fired replication origins using Perl codes.

- (b) Use a 200 bp sliding window to scan each replication origin and calculate the  $\log_2$  ratio of Watson/Crick strand reads using Perl codes.
  - (c) Calculate the average  $\log_2$  ratio of Watson/Crick strand reads of all fired replication origins using Perl codes.
  - (d) email us for the detailed analysis protocol
5. Analysis of the bias pattern of eSPAN peaks at individual origins (Fig. 3d)
- (a) Based on the location of the replication origin and Watson and Crick strands, split each BrdU peak region into the following four quadrants: Watson strand on the left (WL), Watson strand on the right (WR), Crick strand on the left (CL), and Crick strand on the right (CR).
  - (b) Count the number of sequence reads in each four quadrants using Perl codes.
  - (c) Calculate the  $p$ -value to determine whether sequence reads of leading-strand (WL + CR) were significantly different from lagging-strand reads (WR + CL) at each replication fork using the binomial distribution with  $10^{-5}$  as the cut-off  $p$ -value.
  - (d) Calculate the  $\log_2$  ratio at each replication origin using the following formula:  $\log_2 (WL + CR)/(WR + CL)$ , which represents the  $\log_2$  ratio of lagging/leading, to determine whether the fork exhibited a lagging or leading bias pattern.
  - (e) Separate the origins into one of three categories based on  $p$ -values and the  $\log_2$  ratio of leading/lagging sequence reads.

Leading strand bias:  $p < 10^{-5}$  and  $\log_2 (WL + CR)/(WR + CL) < 0$ .

Lagging strand bias:  $p < 10^{-5}$  and  $\log_2 (WL + CR)/(WR + CL) > 0$ .

Indeterminable bias:  $p > 10^{-5}$ .

---

## 4 Notes

1. PMSF is poisonous and not stable in water. TBS+ PMSF buffer should be made just prior to use.
2. Although unsynchronized cells also work for eSPAN experiments, synchronized early S phase cells work much better. In

the presence of hydroxyurea, 2–3 kb regions at early replication origins replicate, which help generate clean eSPAN results.

3. The sonication condition should be tested experimentally before ChIP experiments. The average sheared DNA fragments should be 300 bp–500 bp in length.
4. The amount of each antibody used depends on the target protein and antibody specificity. It is better to determine the amount of each antibody before going through the eSPAN protocol.
5. In synchronized early S phase cells in the presence of hydroxyurea, early replication origin *ARS607* is replicated, whereas the distal site (*ARS607* + 8 k), which is 8 kb distal from the *ARS607*, is not. Therefore, BrdU and targeted protein ChIP and BrdU IP are expected to enrich at *ARS607* compared to the *ARS607* + 8 k (Fig. 2).
6. Follow the instructions for multiplex sequencing provided by Illumina. Replace the first read primer by the custom primer CL72 (ACACTCTTTCCCTACACGACGCTCTTCC) because one ligation adaptor is deleted of 4 nt in the library preparation step [17, 18]. We sequence a maximum of 32 samples per lane.

---

## Acknowledgment

We thank Dr. Albert Serra Cardona for editing this protocol. This work was supported by NIH grants GM118015 to Z.Z and C.Y. is supported by the Edward C. Kendall Fellowship.

## References

1. Stillman B (2005) Origin recognition and the chromosome cycle. *FEBS Lett* 579 (4):877–884. Doi:S0014-5793(04)01538-8 [pii]10.1016/j.Febslet.2004.12.011
2. Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Sugino A (1968) Mechanism of DNA chain growth. I Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc Natl Acad Sci U S A* 59 (2):598–605
3. O'Donnell M, Langston L, Stillman B (2013) Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb Perspect Biol* 5(7.) doi:10.1101/cshperspect.a010108a010108
4. Bell SP, Dutta A (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* 71:333–374. doi:10.1146/annurev.biochem.71.110601.135425110601.135425 [pii]
5. Waga S, Stillman B (1998) The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* 67:721–751. doi:10.1146/annurev.biochem.67.1.721
6. Bell SD, Botchan MR (2013) The Minichromosome maintenance replicative helicase. *Cold Spring Harb Perspect Biol* 5(11):a012807. doi:cshperspect.a012807v1 [pii]10.1101/cshperspect.a012807cshperspect.a012807 [pii]
7. Bell SP, Kaguni JM (2013) Helicase loading at chromosomal origins of replication. *Cold Spring Harb Perspect Biol* 5(6.) doi:cshperspect.a010124 [pii]10.1101/cshperspect.a010124

8. MacAlpine DM, Almouzni G (2013) Chromatin and DNA replication. *Cold Spring Harb Perspect Biol* 5(8):a010207. doi:cshperspect.a010207 [pii]10.1101/cshperspect.a010207
9. Arias EE, Walter JC (2007) Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* 21(5):497–518. doi:21/5/497 [pii]10.1101/gad.1508907
10. Aparicio OM, Stout AM, Bell SP (1999) Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proc Natl Acad Sci U S A* 96(16):9130–9135
11. Aparicio OM, Weinstein DM, Bell SP (1997) Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* 91(1):59–69
12. Yu C, Gan H, Han J, Zhou ZX, Jia S, Chabes A, Farrugia G, Ordog T, Zhang Z (2014) Strand-specific analysis shows protein binding at replication forks and PCNA unloading from lagging strands when forks stall. *Mol Cell* 56(4):551–563. doi:S1097-2765(14)00749-7 [pii]10.1016/j.molcel.2014.09.017
13. Pursell ZF, Isoz I, Lundstrom EB, Johansson E, Kunkel TA (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317(5834):127–130. doi:DOI 10.1126/science.1144067
14. Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA (2008) Division of labor at the eukaryotic replication fork. *Mol Cell* 30(2):137–144. doi:10.1016/j.molcel.2008.02.022S1097-2765(08)00168-8 [pii]
15. Viggiani CJ, Aparicio OM (2006) New vectors for simplified construction of BrdU-incorporating strains of *Saccharomyces cerevisiae*. *Yeast* 23(14–15):1045–1051. doi:10.1002/yea.1406
16. Viggiani CJ, Knott SR, Aparicio OM (2010) Genome-wide analysis of DNA synthesis by BrdU immunoprecipitation on tiling microarrays (BrdU-IP-chip) in *Saccharomyces cerevisiae*. *Cold Spring Harb Protoc* 2010(2.) pdb prot5385. doi:2010/2/pdb.prot5385 [pii] 10.1101/pdb.prot5385
17. Gansauge MT, Meyer M (2013) Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat Protoc* 8(4):737–748. doi:nprot.2013.038 [pii]10.1038/nprot.2013.038
18. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prüfer K, de Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N, Tandon A, Siebauer M, Green RE, Bryc K, Briggs AW, Stenzel U, Dabney J, Shendure J, Kitzman J, Hammer MF, Shunkov MV, Derevianko AP, Patterson N, Andres AM, Eichler EE, Slatkin M, Reich D, Kelso J, Paabo S (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338(6104):222–226. doi:10.1126/science.1224344science.1224344 [pii]
19. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* 9(4):357–359. doi:nmeth.1923 [pii]10.1038/nmeth.1923
20. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842. doi:10.1093/bioinformatics/btq033btq033 [pii]
21. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9):R137. doi:10.1186/gb-2008-9-9-r137gb-2008-9-9-r137 [pii]

## Analysis of Replicative Polymerase Usage by Ribonucleotide Incorporation

Andrea Keszthelyi, Izumi Miyabe, Katie Ptasińska, Yasukazu Daigaku, Karel Naiman, and Antony M. Carr

### Abstract

Mapping the usage of replicative DNA polymerases has previously proved to be technically challenging. By exploiting mutant polymerases that incorporate ribonucleotides into the DNA with a significantly higher proficiency than their wild-type counterparts, we and others have developed methods that can identify what proportion of each DNA strand (i.e., the Watson and Crick strands) is replicated by a specific DNA polymerase. The incorporation of excess ribonucleotides by a mutated polymerase effectively marks, in each individual cells, the DNA strand that is replicated by that specific mutated polymerase. Changes to DNA polymerase usage can be examined at specific loci by Southern blot analysis while a global analysis of polymerase usage can be achieved by applying next-generation sequencing. This genome-wide data also provides a direct measure of replication origin efficiency and can be used to indirectly calculate replication timing.

**Key words** Pu-seq, Fork restart, rNMP incorporation, Origins, HTP sequencing

---

### 1 Introduction

The division of labor for eukaryotic replicative polymerases was initially mapped in *S. cerevisiae* using mutation spectra analysis [1]. In this work, the mutational bias of a specific error-prone mutant polymerase allows the assignment of specific base changes to the mis-incorporation of dNMPs on either the Watson or the Crick strand. An alternative approach was subsequently developed that used Southern blot analysis to map ribonucleotide incorporation at specific loci in strains that contained a specific polymerase mutation that increases the mis-incorporation of rNMPs [2]. Collectively, these approaches demonstrated that Pole synthesized the leading strand, Pol $\delta$  synthesized the lagging strand and that this division of labor between Pole and Pol $\delta$  was conserved between *S. cerevisiae* and *S. pombe*. The analysis of rNMP incorporation by Southern blot analysis has also been used to demonstrate that Pol $\delta$



synthesizes both DNA strands after replication is restarted by homologous recombination [3]. A limitation of such studies is that they are focused on a defined locus.

More recently, the basic idea of exploiting polymerase mutants with a higher promiscuity toward ribonucleotide incorporation has provided a global picture of replication polymerase usage in both *S. cerevisiae* and *S. pombe* by coupling the approach with next-generation sequencing [4–8]. The subsequent genome-wide polymerase usage data also allowed the identification of origins of replication throughout the genome and provided a direct measure of their efficiency (the percentage of the population in which a given origin is fired) [5]. This is of particular interest because previous methods that measured replication dynamics genome-wide were restricted to directly measuring replication timing and only indirectly provide a measure of replication origin efficiency [9].

In this chapter, we describe two approaches to studying polymerase usage. Both are based on increased ribonucleotide incorporation by mutant DNA polymerases. The first provides information on a specific locus and exploits Southern blotting. The second—which we have termed Polymerase usage sequencing (Pu-seq)—involves next-generation sequencing and provides a map of polymerase usage throughout the whole genome. The first method is cost efficient, but is restricted to a few defined loci. The second method requires additional financial input but extends the analysis to the whole genome and provides appreciably better quantitative information.

Both the methods require strains with specific backgrounds. In order to ensure specificity of ribonucleotide incorporation to an individual polymerase, the steric gate of the given replicative enzyme is mutated [10]. This results in promiscuity toward the incorporation of ribonucleotides. However, ribonucleotides in the genome are normally rapidly removed by ribonucleotide excision repair (RER) [11]. In order to maintain the mis-incorporated rNMP residues in the DNA, it is thus necessary to disable RER. We have achieved this by deleting the catalytic subunit of RNase H2 (*rnh201*) (*see* Subheading 3.1). Once the ribonucleotides are stably incorporated into the DNA by the mutant polymerase, DNA can be prepared and the site of incorporation turned into a strand break in the DNA backbone. We achieve this by using alkali treatment (rNMPs are alkali labile), which also makes the DNA single stranded (*see* Subheading 3.2). Fragmented ssDNA can subsequently be subjected to either Southern blotting with strand-specific probes to visualize the extent of the fragmentation (*see* Subheading 3.3) or it can be the starting point to create a library for next-generation sequencing (*see* Subheading 3.4). During the latter process, the 5' end of each fragment generated by alkali treatment allows for the precise identification of the site of ribonucleotide incorporation by the mutant polymerase.

While this protocol describes the analysis of polymerase usage by ribonucleotide mapping, ribonucleotides incorporated at physiological levels into the DNA can also be mapped, i.e., without the presence of specific polymerase mutations. In such experiments the generation of a library for next-generation sequencing should be carried out as described below.

### **1.1 Alkaline Cleavage of DNA Containing Ribonucleotides**

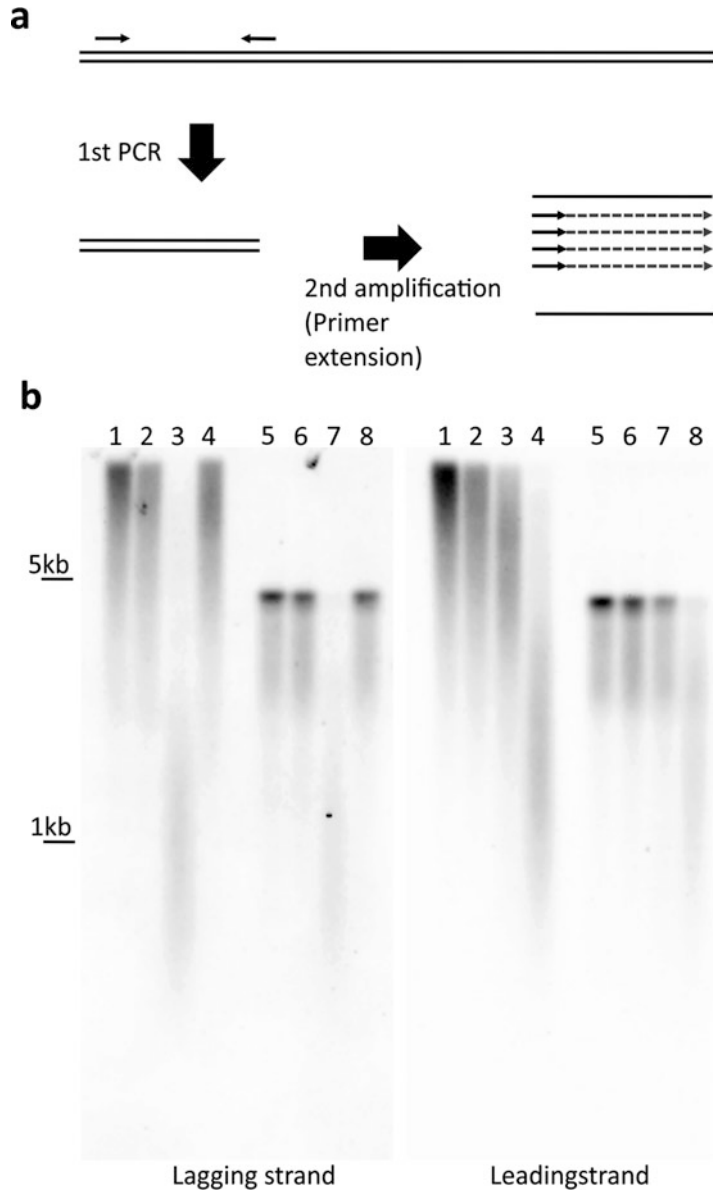
The specific fragmentation of the DNA at the site of the ribonucleotides can be achieved by alkaline treatment or by the use of specific enzymes such as RNase H2. During alkaline treatment the phosphate backbone is cleaved as a result of a hydroxyl radical attack on the 2'-OH group. The products of this reaction are a cyclic 2'-3'-phosphate and a 5'-OH end. In contrast, cleavage by RNase H2 results in 5'-phosphate and 3'-OH ends that are non-ligatable and which thus require further processing for library production. Alkaline treatment also denatures the DNA and thus the analysis of alkali-treated DNA by either Southern blot or by library production for next-generation sequencing must account for this.

### **1.2 Required Genetic Background**

Because ribonucleotides are efficiently removed by ribonucleotide excision repair (RER), downregulation of this system is required so that a sufficient signal-to-noise ratio can be achieved either for Southern blot analysis or for next-generation sequencing. In our methodology we delete the *rnh201* gene that encodes the catalytic subunit of RNase H2. Theoretically, other means of ablating RER would be adequate. The next-generation sequencing method, but not the Southern blotting method, is sufficiently sensitive that it can be used to analyze DNA purified from otherwise unaltered strains to study physiological ribonucleotide incorporation and/or repair. To follow polymerase usage specifically between the Watson and Crick strands of the DNA, two specific polymerase mutations, in Pol $\epsilon$  and Pol $\delta$  respectively, are required that incorporate ribonucleotide with higher proficiency. In this protocol, we describe a data analysis pipeline for analyzing sequence data derived from paired strains containing a mutation in each of these two replicative polymerases (*see* Subheading 3.4.6). However, the method can be extended to the third major replicative polymerase (Pol $\alpha$ ) and potentially other polymerases involved in replication or repair.

### **1.3 Mapping Polymerase Usage Locally or Genome Wide**

Alkali-treated ssDNA from each of the two polymerase mutants can be run on a denaturing gel and be analyzed using strand-specific probes by Southern blot (*see* Subheading 3.3. and Fig. 1). This method has been used to establish the polymerase usage during canonical replication [2] and examine replication polymerase usage during site-specific replication fork arrest and restart events [3]. In each case, pairs of strand-specific probes were designed to separately

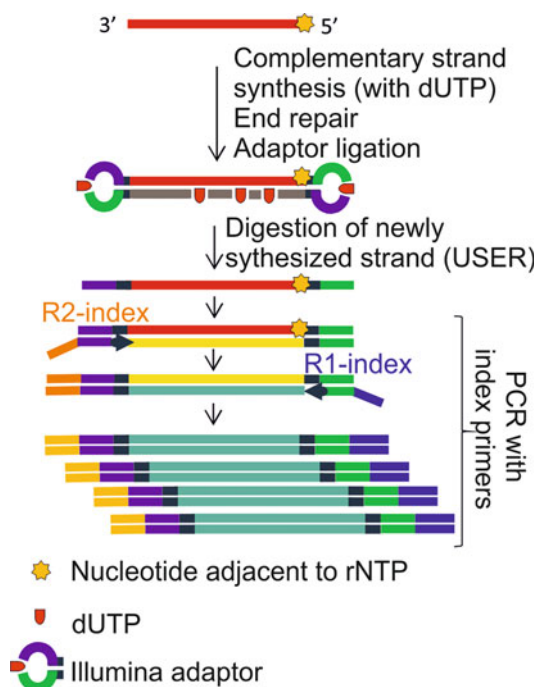


**Fig. 1** Detection of alkali-sensitive sites in genomic DNA by Southern blot. **(a)** Schematic representation of strand-specific probe labeling. *Blue dotted lines* show labeled DNA molecules. **(b)** Typical result. *Left panel:* Probed with the lagging strand. *Right panel:* Probed with the leading strand. Genomic DNA from *S. pombe* wild-type (*lanes 1 and 5*), *rnh201Δ* (*lanes 2 and 6*), *rnh201Δcdc6 (polδ)-L591G* (*lanes 3 and 7*), *rnh201Δcdc20(pole)-M630F* (*lanes 4 and 8*) are subjected to alkaline gel electrophoresis without enzyme digestion (*lanes 1–4*) or after digestion with *XmaI* (*lanes 5–8*)

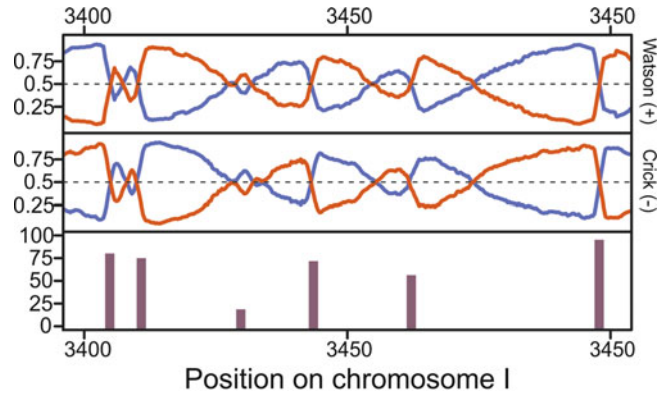
reveal either the Watson or the Crick strand of a specific locus following restriction enzyme digestion.

A far more comprehensive analysis has been achieved by exploiting next-generation sequencing (*see* Subheading 3.4). In this protocol, the position of the nucleotide on the 5' end of each ssDNA fragment generated by the alkali treatment is mapped. This nucleotide was originally adjacent to the ribonucleotide before the alkali cleavage. In order to attach specific adaptors (*see* Subheading 3.4.4, **step 2**) the alkali-treated ssDNA is first converted to dsDNA using random primers (*see* Subheading 3.4.3). To ensure we can assign strand specificity dUTP is used instead of dTTP in the dNTP mix. This enables the selective digestion by the USER enzyme mix of the newly synthesized complementary strand after adaptor ligation (*see* Subheading 3.4.4, **step 4**). Index primers are next used to amplify and create the library. If multiple samples are processed sample-specific index primers allow multiplex runs. For a schematic view on the complementary strand synthesis and library preparation, please refer to Fig. 2.

The library is sequenced on an Illumina sequencing platform to result in at least ten million reads per sample (*see* Subheading 3.4.5). Reads are then aligned to the reference genome and the count numbers of each position (i.e., the relative amount of incorporated ribonucleotides) are calculated (*see* Subheading 3.4.6, **steps 1–6**).



**Fig. 2** Schematic representation of complementary strand synthesis and library preparation. *See* details in the text (Subheadings 1 and 3.4)



**Fig. 3** Typical representation of the Pu-seq specific data analysis. Pol $\epsilon$  (red) and Pol $\delta$  (blue) usage on the Watson (top panel) and on the Crick strand (middle panel) and origin efficiencies (lower panel)

These calculations can be carried out on data from any genetic background and bespoke scripts can be then used for detailed analysis according to the specific aim of the study. Here, we present a data analysis tool (*see* Subheading 3.4.6, **step 7**) that is specific for Pu-seq and generates datasets on the relative contribution of the two main replicative polymerases Pole and Pol $\delta$ . Origins are represented in this dataset as sharp reciprocal changes in the usage of Pole and Pol $\delta$  and the efficiencies of the origins are calculated by determining the height of these changes. The pipeline presented here identifies these positions as origins and calculates their efficiency. A typical result from this analysis can be seen in Fig. 3.

#### 1.4 Data Analysis— Generating Count Numbers

Next-generation sequencing should be carried out as paired-end sequencing where library fragments are sequenced from both directions to allow high accuracy of alignment. Paired-end sequencing generates two FASTQ files—each containing the 3' or the 5' end sequence of the fragments. These reads are aligned to the reference genome using the bowtie2 program that is freely available (*see* Subheading 3.4.6, **steps 1–4**). This process generates a single SAM file from the two FASTQ files for each strain. The SAM file contains information on the position where each fragment was aligned, in which direction they were orientated and statistics on the quality of each alignment (*see* <https://samtools.github.io/hts-specs/SAMv1.pdf>). This information is subsequently converted into count numbers by a bespoke perl program (pe-sam-to-bin-count.pl) which we have deposited on the github website (<https://github.com/yasukasu/sam-to-bincount>). Briefly, the program finds the fragments where both ends were aligned concordantly and defines the position of the base next to the 5' end of the original fragment (corresponding to the ribonucleotide). These positions are then divided into two groups, according to whether they were

aligned to the Watson or to the Crick strand. Last, the reads are sorted by chromosome and position, the genome is divided into bins with a specified size (300 bp by default) and the numbers of reads are counted for each bin. In this way two CSV files are generated—one for the Watson and one for the Crick strand—where each row corresponds to a bin and each column contains the name of the chromosome, the middle position of the bin and the counts of reads within that bin. These CSV files can be generated from any paired-end sequencing data and can be a starting point for various bespoke analyses. In the next section, we describe a Pu-seq specific pipeline that starts from the CSV files generated from sequences from Pole and Pol $\delta$  mutant strains that outputs polymerase usage data for each strand in addition to a list of the positions and the efficiency of origins.

### 1.5 Data Analysis— Pu-Seq-Specific Pipeline

We provide a program written in the R language (<https://github.com/yasukasu/Pu-seq.R>) where the CSV files described above (that contain the counts of sequenced fragments from Pole and Pol $\delta$  mutants—both in an *rnh201A* background) are used as input. The program generates WIG and BEDGRAPH files containing data on polymerase usage related to chromosome position and the efficiency of origins throughout the genome. These files can be used to visualize the data in a genome browser (e.g., IGV).

To calculate the relative contribution of each polymerase on each strand at each bin the program inputs four datasets; the counts (C) on the forward (+) and reverse (−) strands from both the Pol $\delta$  ( $\delta$ ) and from the Pole ( $\epsilon$ ) strains ( $C_{\delta}^{+}$ ,  $C_{\delta}^{-}$ ,  $C_{\epsilon}^{+}$ , and  $C_{\epsilon}^{-}$ ). These counts are then normalized to enable a direct comparison between them: at each position (1) the count is divided by the sum of all the counts from that dataset (e.g., the normalized counts (N) at position  $i$  on the forward strand (+) from Pol $\delta$  ( $N_{\delta i}^{+}$ ) equals the counts at this position divided by the sum of all the counts from the Pol $\delta$  dataset for the forward strand:  $N_{\delta i}^{+} = C_{\delta i}^{+} / \sum C_{\delta}^{+}$ ). The ratio (R) of the usage of each polymerase is calculated using the assumption that each strand is synthesized by either Pol $\delta$  or Pole (i.e., ignoring the relatively small contribution of Pol $\alpha$ ). The ratio of Pol $\delta$  on the forward strand at each position is then calculated as  $R_{\delta i}^{+} = N_{\delta i}^{+} / (N_{\delta i}^{+} + N_{\epsilon i}^{+})$ . The relative contribution of Pol $\delta$  on the reverse strand is similarly calculated ( $R_{\delta i}^{-} = N_{\delta i}^{-} / (N_{\delta i}^{-} + N_{\epsilon i}^{-})$ ). The corresponding equations are used to generate the datasets for Pole usage on the forward and reverse strands: ( $R_{\epsilon i}^{+} = N_{\epsilon i}^{+} / (N_{\epsilon i}^{+} + N_{\delta i}^{+})$  and  $R_{\epsilon i}^{-} = N_{\epsilon i}^{-} / (N_{\epsilon i}^{-} + N_{\delta i}^{-})$ ). These ratios can be smoothed by a moving average specified by the user (*see* Subheading 3.4.6, **step 7**, Option `--ma1`).

Origins of DNA replication are represented by sharp transitions between the usages of the two polymerases. These changes appear as sudden increases in the Pol $\delta$  usage on the reverse strand ( $R_{\delta}^{-}$ ) and Pole usage on the forward strand ( $R_{\epsilon}^{+}$ ). To identify these steep

positive slopes the differentials of these two datasets (i.e., the differences of each neighboring data point) are calculated and smoothed by a moving average (*see* Subheading 3.4.6, **step 7**, Option `--ma2`). Positive peaks in a differential dataset represent sudden change of polymerase usage in relation to chromosome position. The positions of these peaks are then identified. To locate origins above the noise, only peaks with a maximum above the 30th percentile of all peaks are considered and peaks found within four bins are merged. The positions of these peaks are subsequently identified in the corresponding original polymerase usage dataset and the differences between the local maxima and the minima surrounding each peak (the heights of the positive changes) are calculated. This difference, multiplied by 100, represents the efficiency, expressed as a percentage, of the origin (i.e., in the population what percentage of the cells fired that particular origin). Lastly, only origins identified in both independent datasets (Pol $\delta$  usage on the reverse strand ( $R^-_{\delta}$ ) and Pole usage on the forward strand ( $R^+_{\epsilon}$ ) calculated from each of the matched mutant strains) within four bins are chosen and the average of their efficiencies from both datasets is calculated.

---

## 2 Materials

### 2.1 Cell Growth, DNA Extraction, and Alkali Treatment

1. YE medium: 0.5% Difco Yeast Extract 3% Glucose.
2. Liquid nitrogen.
3. NIB buffer: 50 mM MOPS (pH 7.2), 17% Glycerol, 150 mM potassium acetate, 2 mM MgCl<sub>2</sub> (*see* **Note 1**).
4. Lyticase (Sigma; L4025-1MU).
5. 1% (wt/vol) SDS.
6. Glass slide.
7. Qiagen buffer G2.
8. RNase A.
9. 30% (wt/vol) N-lauroyl sarcosine.
10. 20 mg/ml proteinase K.
11. Qiagen 100/G Genomic-tip.
12. TE.
13. 1 M NaOH.

### 2.2 Detection of Alkali-Sensitive Sites in Genomic DNA by Southern Blot

1. Ethanol.
2. 1 N KOH.
3. Agarose.

4. 10× Alkaline electrophoresis buffer: 500 mM NaOH, 10 mM EDTA. Prepare just before use. EDTA is necessary to prevent the precipitation of Mg(OH)<sub>2</sub> which may entrap DNA.
5. 1× Alkaline electrophoresis buffer: 50 mM NaOH, 1 mM EDTA. Prepare just before use.
6. 6× Alkaline loading buffer: 300 mM NaOH, 6 mM EDTA, 15% (w/v) Ficoll400, 0.25% (w/v) Xylene Cyanol FF, 0.15% (w/v) Bromocresol Green.
7. Neutralization buffer: 1 M Tris-HCl (pH 7.2), 1.5 M NaCl.
8. 0.5 µg/ml EtBr.
9. Uncharged nylon membrane.
10. 10× Standard Taq Reaction Buffer containing 15 mM MgCl<sub>2</sub>.
11. dNTP (2 mM each) without dCTP.
12. α-<sup>32</sup>P-dCTP.
13. Taq polymerase.
14. G-50 gel filtration spin column.

### **2.3 Genome-Wide Identification of Incorporated Ribonucleotides**

1. 2% (wt/vol) TBE gel in 0.5× TBE buffer.
2. 0.5 µg/ml acridine orange solution (2000× dilution in water from 10 mg/ml stock).
3. Macherey-Nagel gel extraction kit.
4. Qubit ssDNA assay kit.
5. 3 mg/ml of 8 N random primer.
6. 10× NEB 2.1 buffer.
7. 2 mM each dNTP with dUTP instead of dTTP.
8. T4 polymerase.
9. 0.5 M EDTA.
10. AMPure XP beads.
11. 80% (vol/vol) ethanol.
12. Magnetic rack.
13. High-sensitivity DNA chip.
14. NEBNext Ultra DNA library prep kit for Illumina.
15. NEBNext multiplex oligos for Illumina.
16. NEBNext USER enzyme.

---

## **3 Methods**

### **3.1 Strain Generation**

1. To follow ribonucleotide incorporation by mutant polymerases deletion of *rnh201*, or another means of RNase H2 down-regulation, is required.



2. In our hands the following *S. pombe* DNA polymerase mutants showed the optimal balance between a sufficient level of ribonucleotide incorporation and retaining viability:

- Polε: *cdc20-M630F*.
- Polδ: *cdc6-L591G*.
- Polα: *Poll-L850 M*.

For *S. cerevisiae* others have used the following equivalent mutations [6, 7]:

- Polε: *pol2-M644G*.
- Polδ: *pol3-L612G, pol3-L612 M*.
- Polα: *pol1-L868 M, pol1-Y869A*.

### 3.2 Cell Growth, DNA Extraction, and Alkali Treatment

1. Primary culture:  
Inoculate one colony of each strain into 10 ml of YE medium in a test tube and incubate it overnight at 30 °C with constant rotation.
2. Inoculate  $8 \times 10^7$  cells from the primary culture into 1 L of YE at 30 °C with 180 rpm shaking. Allow the culture to grow to  $3\text{--}5 \times 10^6$  cells/ml. Collect the cells by centrifugation ( $6000 \times g$  at 4 °C, 5 min) (*see Note 2*).
3. Remove the supernatant and resuspend the cell pellet in 40 ml H<sub>2</sub>O. Transfer the cells to Falcon tube and centrifuge again ( $4000 \times g$  at 4 °C, 5 min). After discarding the supernatant immediately freeze the cells in liquid nitrogen. Cell pellets can be kept at –80 °C for extended periods of time.
4. Resuspend the cell pellet in 2 ml of NIB buffer containing 5 mg/ml lyticase and incubate for 20 min at 37 °C. To check cell lysis, mix 1 µl of cells with 9 µl of 1% (wt/vol) SDS on a glass slide and observe the cells under the microscope. Lysed cells should appear hollow. If insufficient cells have lysed continue incubation for a further 10 min and check again.
5. When lysis is judged sufficient (>70% of cells lyse), add 20 ml of ice-cold water and mix. After centrifugation (at  $4000 \times g$  for 10 min) remove the supernatant and gently resuspend the pellet in 2 ml Qiagen buffer G2.
6. RNase treat the mix by adding 100 µl of 10 mg/ml RNase A, followed by incubation for 30 min at 37 °C.
7. Add 100 µl of 30% (wt/vol) *N*-lauroyl sarcosine and 100 µl of freshly prepared 20 mg/ml proteinase K. After 30 min incubation at 55 °C centrifuge the mixture at  $4000 \times g$  for 15 min at 4 °C and transfer the supernatant into an empty falcon tube. To increase DNA yield, repeat the proteinase treatment on the remaining pellet; resuspend it in 1 ml Qiagen buffer G2, 50 µl of 30% (wt/vol) *N*-lauroyl sarcosine, and 50 µl 20 mg/

ml proteinase K. Incubate for 30 min at 55 °C. After centrifugation, transfer the supernatant to the same tube as the supernatant from the first proteinase K treatment.

8. Using the combined supernatant, follow the Genomic-tip protocol for 100/G Genomic-tip described in the Qiagen Genomic DNA Handbook. After precipitation collect the DNA by centrifugation and resuspend it in 200 µl of TE.
9. Measure the DNA concentration with Nanodrop. The expected yield is 250–500 ng/µl. DNA can be stored at –20 °C for at least 1 month.

### 3.3 Detection of Alkali-Sensitive Sites in Genomic DNA by Southern Blot

#### 3.3.1 Alkaline Agarose Gel Electrophoresis (See Note 1)

1. Digest 5 µg of genomic DNA with appropriate restriction enzyme(s) (*see Note 3*).
2. Precipitate DNA with ethanol and resuspend in 28 µl of distilled water.
3. Add 12 µl of 1 N KOH and incubate for 2 h at 55 °C.
4. In the meantime, add the appropriate amount of agarose to a measured quantity of distilled water (in most cases, a 1% gel is sufficient). Heat the slurry in a microwave oven until the agarose dissolves.
5. After cooling the agarose solution to 55 °C add 0.1 volume of 10× alkaline electrophoresis buffer and pour immediately into the gel tray (*see Note 4*). Leave the gel until it is completely set.
6. Mount the gel in the electrophoresis tank. Add freshly prepared 1× alkaline electrophoresis buffer until the gel is completely covered.
7. Add 8 µl of 6× alkaline loading buffer to the DNA samples after **step 3** is completed.
8. Load the samples into the wells and electrophorese at <5 V/cm until the xylene cyanol has migrated approximately half of the length of the gel (*see Note 5*).
9. Soak the gel in neutralization buffer and incubate for 30 min at room temperature with gentle shaking.
10. Soak the gel in 0.5 µg/ml EtBr solution and incubate overnight. Some other staining dyes, such as SYBR Gold, should also work. This step can be skipped if it is judged that the visualization of the total DNA used for the subsequent Southern transfer is unnecessary.
11. Transfer DNA to an uncharged nylon membrane according to any standard Southern blot protocol.

#### 3.3.2 Southern Hybridization with Strand-Specific Probes

1. Amplify the DNA fragment (100–500 bp) of interest using a standard PCR protocol.

2. Purify the PCR product with Qiagen PCR purification or Gel extraction kit.
3. Prepare the reaction mix for primer extension (*see* **Note 6**).
  - 1  $\mu\text{l}$  Template (1  $\mu\text{g}/\mu\text{l}$  purified PCR product).
  - 1  $\mu\text{l}$  primer (10  $\mu\text{M}$ , *see* **Note 7**).
  - 2.5  $\mu\text{l}$  10 $\times$  Standard Taq Reaction Buffer containing 15 mM  $\text{MgCl}_2$ .
  - 1  $\mu\text{l}$  dNTP (2 mM each) without dCTP.
  - 2  $\mu\text{l}$   $\alpha$ - $^{32}\text{P}$ -dCTP.
  - 0.1  $\mu\text{l}$  Taq polymerase.
  - Distilled water up to 25  $\mu\text{l}$ .
4. Perform the reaction with the following conditions:
  - (a) 94  $^{\circ}\text{C}$  for 5 min.
  - (b) 94  $^{\circ}\text{C}$  for 1 min.
  - (c) 55  $^{\circ}\text{C}$  for 30 s.
  - (d) 72  $^{\circ}\text{C}$  for 1 min.
  - (e) Repeat (c) and (d) 25 times.
  - (f) 72  $^{\circ}\text{C}$  for 5 min.
  - (g) Hold 4  $^{\circ}\text{C}$ .
5. Purify the reaction using a G-50 gel filtration spin column.
6. Hybridization can be performed using any standard protocol.
7. Strip the membrane and re-probe with the other strand-specific probe (Fig. 1).

### **3.4 Genome-Wide Identification of Incorporated Ribonucleotides**

#### *3.4.1 Alkali Treatment of Genomic DNA*

#### *3.4.2 Size Selection of the Alkali-Treated Genomic DNA*

1. Dilute 20  $\mu\text{g}$  of the DNA in 70  $\mu\text{l}$  of  $\text{H}_2\text{O}$  and add 30  $\mu\text{l}$  of 1 M NaOH (final concentration of 0.3 M). Incubate for 2 h at 55  $^{\circ}\text{C}$ .

The protocol can be executed without this step. However, we found that size selection results in more efficient and reproducible complementary strand synthesis (*see* Subheading 3.4.3). If the method is used for the identification of ribonucleotide incorporation in DNA from wild-type cells (i.e., no downregulation of RNase H2, and/or no DNA polymerase mutation), it is unlikely that the concentration of ribonucleotides in the DNA is sufficient to result in small fragments after alkali treatment. In this case, it is thus not possible to size select DNA fragments. Therefore, this step should not be included in the protocol.

1. Run 10  $\mu\text{g}$  of DNA (50  $\mu\text{l}$ ) on a 2% (wt/vol) TBE gel in  $0.5\times$  TBE buffer for 2 h at 100 V (*see Note 8*).
2. Stain the gel in a 0.5  $\mu\text{g}/\text{ml}$  acridine orange solution for 2 h at room temperature with gentle shaking followed by destaining in water overnight at room temperature with gentle shaking. Changing the water a couple of times helps to decrease the presence of background staining.
3. Visualize the gel under long-wave UV illumination. As acridine orange stains ssDNA red and dsDNA green a red smear should be observed.
4. Excise a slice of gel that contains ssDNA fragments between 300 and 500 bp and isolate the ssDNA from the gel slice with the Macherey-Nagel gel extraction kit as instructed by the manufacturer (*see Note 9*).

**3.4.3 Complementary Strand Synthesis (See Fig. 2)**

In our hand if the protocol is executed without size selection of the ssDNA (*see Subheading 3.4.2*) klenow fragment from BioPrime DNA labeling system works with higher efficiency than T4 polymerase (*see Note 10*). If the size selection step was included follow the protocol below.

1. Measure the ssDNA concentration with a fluorimeter using the Qubit ssDNA assay kit and dilute 100 ng into a final volume of 30  $\mu\text{l}$  in  $\text{H}_2\text{O}$ .
2. Add 5  $\mu\text{l}$  of 3 mg/ml of 8 N random primer and 5  $\mu\text{l}$  of  $10\times$  NEB 2.1 buffer to the ssDNA and boil for 5 min at 95  $^\circ\text{C}$ . Immediately place it on ice for 5 min.
3. Add 5  $\mu\text{l}$  of 2 mM each dNTP (use dUTP instead of dTTP), 4  $\mu\text{l}$  of  $\text{H}_2\text{O}$  and 1  $\mu\text{l}$  of T4 polymerase.
4. Incubate the mixture at 37  $^\circ\text{C}$  for 20 min. Immediately add 5  $\mu\text{l}$  of 0.5 M EDTA (pH 8) to stop the reaction.

**3.4.4 Library Preparation (See Fig. 2)**

1. Purification of dsDNA.
  - (a) Resuspend the AMPure XP beads by vortexing and add 99  $\mu\text{l}$  ( $1.8\times$ ) of beads to the reaction from the previous step.
  - (b) Mix well by pipetting up and down and incubate for 5 min at room temperature. Briefly spin the tube to collect any sample from the wall of the tube.
  - (c) Place the tube on a magnetic rack and wait until the beads separate from the supernatant ( $\sim 5$  min). Carefully remove and discard the supernatant.
  - (d) Add 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol) ethanol to the beads (leave the tube in the magnetic rack) and

incubate at room temperature for 30 s, then carefully remove and discard the supernatant.

- (e) Repeat the wash step with 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol).
  - (f) Keep the tubes in the magnetic rack and air-dry the beads with the lid open for 10 min.
  - (g) Elute the DNA by adding 60  $\mu\text{l}$  of nuclease-free water to the beads. Mix well by pipetting up and down. Briefly spin the tube to collect any sample from the wall of the tube. Place it in the magnetic rack and wait until the solution clears.
  - (h) Remove 58  $\mu\text{l}$  of the supernatant and transfer it to a clean nuclease-free PCR tube.
  - (i) Run 1  $\mu\text{l}$  on the Bioanalyzer for a quality check using a high-sensitivity DNA chip (*see Note 11*).
2. End repair and adaptor ligation.
- (a) Transfer 55.5  $\mu\text{l}$  of the purified dsDNA to a PCR tube and add 6.5  $\mu\text{l}$  of NEBNext end repair reaction buffer (10 $\times$ ) and 3  $\mu\text{l}$  of NEBNext end prep enzyme mix.
  - (b) Incubate the sample in a thermal cycler for 30 min at 20  $^{\circ}\text{C}$  and then for 30 min at 65  $^{\circ}\text{C}$ .
  - (c) Add the followings to the sample:
    - 15  $\mu\text{l}$  Blunt/TA ligase master mix.
    - 1  $\mu\text{l}$  10 $\times$  diluted NEBNext adapter (dilute the adapter freshly in nuclease-free water).
    - 2.5  $\mu\text{l}$  nuclease-free water.
  - (d) Incubate the mixture for 15 min at 20  $^{\circ}\text{C}$  in a thermal cycler.
3. Size selection of the adaptor ligated fragment.

To ensure a consistent size of insert in the library and to discard adaptor dimers and other unwanted residuals, a size selection step is used at this point. For Illumina sequencing the optimal fragment size is around 400 bp. However, to ensure a sufficient yield, we have found it works better if we recover fragments of a slightly higher size range (250–600 bp).

- (a) Add 16.5  $\mu\text{l}$  of nuclease-free water to the reaction to adjust the final volume to 100  $\mu\text{l}$ .
- (b) Resuspend the AMPure XP beads by vortexing and add 35  $\mu\text{l}$  (0.35 $\times$ ) to the mixture.
- (c) Mix well by pipetting up and down and incubate it for 5 min at room temperature.

- (d) Briefly spin the tube to collect any sample from the wall of the tube.
  - (e) Place the tube on a magnetic rack and wait until the beads separate from the supernatant (~5 min) and then carefully remove and transfer the supernatant to a new tube (*see Note 12*).
  - (f) Add 35  $\mu\text{l}$  ( $0.26\times$ ) of resuspended AMPure XP beads to the supernatant, mix well, and incubate for 5 min at room temperature.
  - (g) Briefly spin the tube to collect any sample from the wall of the tube.
  - (h) Place the tube on a magnetic rack and wait until the beads separate from the supernatant (~5 min) and then carefully remove and discard the supernatant.
  - (i) Add 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol) ethanol to the beads (leave the tube in the magnetic rack) and incubate it at room temperature for 30 s, then carefully remove and discard the supernatant.
  - (j) Repeat the washing step twice with 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol) for a total three washes.
  - (k) Keep the tubes in the magnetic rack and air-dry the beads for 10 min with the lid open.
  - (l) Elute the DNA by adding 25  $\mu\text{l}$  of  $0.1\times$  TE, pH 8.0 to the beads. Mix well by pipetting up and down.
  - (m) Briefly spin the tube and then place it in the magnetic rack and wait until the solution is clear (~5 min), transfer 23  $\mu\text{l}$  to a new tube. As residual beads may affect the following polymerase reaction be careful not to transfer any beads.
4. USER excision and PCR library enrichment.
- (a) Take the Q5 Hot Start Hifi Master mix out of storage and let it warm up to room temperature.
  - (b) Transfer 20  $\mu\text{l}$  size-selected DNA to a PCR tube, and add the followings:
    - 3  $\mu\text{l}$  NEBNext USER enzyme.
    - 25  $\mu\text{l}$  Q5 Hot Start Hifi Master mix.
    - 1  $\mu\text{l}$  Universal PCR primer (25  $\mu\text{M}$ ).
    - 1  $\mu\text{l}$  Sample-specific index primer (*see Note 13*) (25  $\mu\text{M}$ ).
  - (c) Perform a PCR with the following conditions:
    - 37 °C for 15 min (USER digestion).
    - 98 °C for 30 s (Initial denaturation).
    - 98 °C for 10 s (Denaturation).

- 65 °C for 75 s (Annealing/extension).
- Repeat the above two steps nine times (Cycle numbers might need optimization *see* **Note 14**).
- 65 °C for 5 min (Final extension).
- Hold 4 °C.

#### 5. Purification of the library

To ensure that there is no residual primer or adaptor dimer in the final library, we routinely perform the AMPure purification twice.

- (a) Resuspend the AMPure XP beads by vortexing and add 50  $\mu\text{l}$  (1 $\times$ ) to the library.
- (b) Briefly spin the tube to collect any sample from the wall of the tube.
- (c) Place the tube on a magnetic rack and wait until the beads separate from the supernatant (~5 min) and then carefully remove and discard the supernatant.
- (d) Add 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol) ethanol to the beads (leave the tube in the magnetic rack) and incubate it at room temperature for 30 s, then carefully remove and discard the supernatant.
- (e) Repeat the washing step once with 200  $\mu\text{l}$  of freshly prepared 80% (vol/vol) for a total of two washes.
- (f) Keep the tubes in the magnetic rack and air-dry the beads for 10 min with the lid open.
- (g) Elute the DNA by adding 51  $\mu\text{l}$  of nuclease-free water to the beads. Mix well by pipetting up and down.
- (h) Briefly spin the tube and then place it in the magnetic rack and wait until the solution is clear (~5 min), transfer 50  $\mu\text{l}$  to a new tube.
- (i) Repeat the clean-up procedure from **steps (a) to (f)**.
- (j) Elute the DNA by adding 23  $\mu\text{l}$  of nuclease-free water to the beads. Mix well by pipetting up and down.
- (k) Briefly spin the tube and then place it in the magnetic rack and wait until the solution is clear (~5 min), transfer 20  $\mu\text{l}$  to a new tube. As residual beads may affect the following sequencing reactions, be careful not to transfer any beads.
- (l) Dilute 1  $\mu\text{l}$  of the purified library in 4  $\mu\text{l}$  of nuclease-free water and run 1  $\mu\text{l}$  from this dilution in an Agilent DNA high-sensitivity chip to assess library quality (*see* **Note 15**).

#### 3.4.5 Library Sequencing

For Pu-seq analysis, sequence the library on an Illumina high-throughput sequencing system. We aim for a minimum of ~10 million reads from each strain. We use the NextSeq system with 150 cycle paired-end sequencing. For details, *see* documentation on the Illumina webpage (<https://support.illumina.com/>).

## 3.4.6 Data Analysis

Below are commands in Linux syntax which can be used to output the Pu-seq information from paired end reads stored in FASTQ formats.

1. Download and install Bowtie2, follow the instructions on the Bowtie2 homepage (<http://bowtie-bio.sourceforge.net/bowtie2/>).
2. Create an index from a reference genome FASTA file with the following command (*see Note 15*):

```
filepath/bowtie2-build filepath/reference_FASTA_file index_name
```

3. Align the paired end reads (R1 and R2 index files) from the library created from the polymerase  $\delta$  mutant background, generating a single SAM file.

```
filepath/bowtie2 -x filepath/index_name -trim5 1 -trim3 30 -1 pol-d_R1.fastq -2 pol-d-R2.fastq -S pol-d.sam
```

This command uses parameters that give good genome-wide alignment of reads to the *S. pombe* reference genome. They can be modified for better coverage of, e.g., repetitive regions. For more options *see* Bowtie2 online manual (<http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml>).

4. Align the paired end reads from the library created from the polymerase  $\epsilon$  mutant background, generating a single SAM file.

```
filepath/bowtie2 -x filepath/index_name -trim5 1 -trim3 30 -1 pol-e_R1.fastq -2 pol-e-R2.fastq -S pol-e.sam
```

5. Download the custom-made Perl program `pe-sam-to-bincount.pl` from <https://github.com/yasukasu/sam-to-bincount> and use it to extract concordantly aligned reads from the SAM file generated from the  $\text{pol}\delta$  dataset, sort them according to forward and reverse strand and calculate the number of reads in each bin.

```
perl filepath/pe-sam-to-bincount.pl -i pol-d.sam -trim5 1 -strand -end 1 -n 1 -w 300 -ref filepath/reference_FASTA_file
```

This outputs two sets of files to the working directory; one for the reverse and one for the forward strand. The CSV files contain the numbers of reads in each bin (size of which is set by `--w`). The counts are calculated by the position adjacent to the 5' end of the R1 mate, set by the `--end` and `-n` parameters. These parameters can be modified to extract a different subset of reads. Details of all of the options can be found by typing in the help command –



```
perl filepath/pe-sam-to-bincount.pl -h
```

In addition to the CSV files, for each strand an ALN and a SORT.ALN file will be created. These files are intermediate files for the final CSV file and contain a list of the position on each read (in the SORT.ALN file these positions are sorted by chromosome and position). The ALN and SORT.ALN files are not necessary for further Pu-Seq analysis.

6. Extract concordantly aligned reads from the SAM file generated from the pole dataset and sort them.

```
perl filepath/pe-sam-to-bincount.pl -i pol-e.sam -trim5 1 -strand -end 1 -n 1 -w 300 -ref filepath/reference_FASTA_file
```

7. Download the custom-made R program Pu-seq.R from <https://github.com/yasukasu/Pu-seq.R> and use it to transform the raw count data from the four CSV files generated in **steps 4** and **5** into polymerase usage data and origin datasets (*see Note 16*).

```
R < filepath/Pu-seq.R -vanilla -slave -args -prefix strain_name -df pol-d.e1.fw300.count.csv -dr pol-d.e1.r-w300.count.csv -ef pol-e.e1.f-w300.count.csv -er pole.e1.r-w300.count.csv -ma1 3 ma2 3 -ori
```

Execute this command in the working directory in which the CSV files from **steps 4** and **5** are located. It is very important that each of the files is entered into the command at the appropriate option (for example, the CSV files with counts from the polymerase  $\delta$  mutant background on the forward and reverse strands should be entered as `--df` and `--dr`, respectively; similarly, for reads from the polymerase  $\epsilon$  mutant background).

A more detailed description of the options and how they can be modified can be found on <https://github.com/yasukasu/Pu-seq.R>.

This command will generate four WIG files, describing the usage of the two polymerase on the forward and reverse strands. A BEDGRAPH file, with information about the origins of replication and their firing efficiencies, will also be created.

8. Visualize the WIG files and BEDGRAPH file using a genome browser, e.g., IGV (<https://www.broadinstitute.org/igv/>).

We recommend overlaying the usage of polymerases  $\delta$  and  $\epsilon$  on each strand, creating two panels, one for the forward strand and another for the reverse (*see Fig. 3*).

---

## 4 Notes

1. We have occasionally observed that prolonged incubation with lyticase in NIB buffer caused degradation of DNA, which affected reproducibility of our Southern analysis. This problem can be resolved by replacing 2 mM MgCl<sub>2</sub> with 40 mM EDTA in NIB buffer and using 1 mg/ml Zymolyase 100T as opposed to lyticase to remove the cell wall.
2. It is important not to let the cultures overgrow and to grow different strains to the same cell concentration: we have found that different growth states apparently result in different ribonucleotide incorporation levels.
3. A restriction digestion is not always necessary. However, it helps to detect alkali sensitivity at a specific genomic locus. Using the *S. pombe* strains described above, the alkaline degradation of DNA is obvious within a restriction fragment as small as 2 kb.
4. Because hydrolysis of the polysaccharides occurs in alkaline condition at high temperature, NaOH should be added after cooling the agarose solution to <55 °C.
5. The appropriate voltage depends on the equipment to be used. The gel and buffer temperature should be kept lower than 40–45 °C to prevent alkaline hydrolysis of the gel. Gel tanks with a large buffer volume, or a buffer cooling system, will allow the gel to be run at high voltages. Therefore, the voltage should be adjusted adequately to the equipment.
6. Nonradioactive nucleotide analogues such as fluorescein-12-dUTP should work as well as radioisotopes. In this case, unmodified nucleotide should be added at the appropriate ratio. For example, fluorescein-12-dUTP and unmodified dTTP are frequently used in the ratio of 0.35:0.65.
7. The individual primers used to amplify the template should work well for the labeling reaction, assuming that the original PCR is specific. Designing internal primers may help to avoid the labeling of nonspecific PCR products.
8. To avoid overloading the gel, divide the sample into two and run 25–25 µl (5–5 µg) per lane.
9. In our hands freezing the gel slice on –20 °C for at least 20 min before extraction increases the final yield of isolated ssDNA.
10. If the size selection step is not performed during the protocol, we recommend the use of Klenow fragment obtained from the BioPrime DNA labeling system instead of T4 polymerase. In this case, starting from 100 ng ssDNA, follow the instruction manual provided by the supplier but use 2 mM each dNTP mix

with dUTP substituted for dTTP instead of the labeled dNTP mix provided in the kit and reduce the length of incubation on 37 °C to 40 min. In our hands, the Klenow fragments obtained from other manufacturers did not perform as well as the Klenow fragment from the BioPrime DNA labeling system.

11. Our typical yield is at least 300–600 pg/μl dsDNA from size-selected ssDNA and ~100 pg/μl dsDNA from non-size-selected ssDNA. The size distribution is usually between 300 and 500 bp from size-selected ssDNA and somewhat smaller (200–400 bp) from a non-size-selected template.
12. Do not discard the supernatant. At this step due to the PEG concentration added with the AMPure XP only the large (>600 bp) unwanted fragments are precipitated and bound to the beads.
13. Multiplex sequencing is possible by using sample-specific index primers. In this case several samples can be run on the same sequencing lane. On a Next-seq platform we typically mix 28–30 samples which results in ~10 million paired-end reads from each sample.
14. A clear peak of 300–600 pg/μl DNA between 200 and 500 bp should be visible in the bioanalyzer profile. An additional peak at higher molecular weight could be the result of over-amplification and thus the number of PCR cycles should be reduced. On the other hand, if an insufficient quantity of DNA is present, additional PCR cycles could be applied. Additional peaks at 50–150 bp indicate the presence of an adapter dimer. In this case, it is advised to clean up the library again with 1× AMPure beads.
15. Do not copy paste commands from the Windows to Linux, as this often creates syntax errors.
16. R software environment can be downloaded from <https://www.r-project.org/>.

---

## Acknowledgment

We thank Conrad Nieduszynski, Carolin Muller, and members of the Carr and Murray labs for discussion.

## References

1. Pursell ZF et al (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317(5834):127–130
2. Miyabe I, Kunkel TA, Carr AM (2011) The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet* 7(12): e1002407
3. Miyabe I et al (2015) Polymerase delta replicates both strands after homologous recombination-dependent fork restart. *Nat Struct Mol Biol* 22(11):932–938

4. Keszthelyi A et al (2015) Mapping ribonucleotides in genomic DNA and exploring replication dynamics by polymerase usage sequencing (Pu-seq). *Nat Protoc* 10(11):1786–1801
5. Daigaku Y et al (2015) A global profile of replicative polymerase usage. *Nat Struct Mol Biol* 22(3):192–198
6. Clausen AR et al (2015) Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol* 22(3):185–191
7. Reijns MA et al (2015) Lagging-strand replication shapes the mutational landscape of the genome. *Nature* 518(7540):502–506
8. Ding J et al (2015) Genome-wide mapping of embedded ribonucleotides and other non-canonical nucleotides using emRiboSeq and EndoSeq. *Nat Protoc* 10(9):1433–1444
9. Retkute R, Nieduszynski CA, de Moura A (2012) Mathematical modeling of genome replication. *Phys Rev E Stat Nonlinear Soft Matter Phys* 86(3 Pt 1):031916
10. Brown JA, Suo Z (2011) Unlocking the sugar “steric gate” of DNA polymerases. *Biochemistry* 50(7):1135–1142
11. Sparks JL et al (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* 47(6):980–986

# Chapter 19

## Dynamic Architecture of Eukaryotic DNA Replication Forks In Vivo, Visualized by Electron Microscopy

Ralph Zellweger and Massimo Lopes

### Abstract

The DNA replication process can be heavily perturbed by several different conditions of genotoxic stress, particularly relevant for cancer onset and therapy. The combination of psoralen crosslinking and electron microscopy has proven instrumental to reveal the fine architecture of in vivo DNA replication intermediates and to uncover their remodeling upon specific conditions of genotoxic stress. The replication structures are stabilized in vivo (by psoralen crosslinking) prior to extraction and enrichment procedures, allowing their visualization at the transmission electron microscope. This chapter outlines the procedures required to visualize and interpret in vivo replication intermediates of eukaryotic genomic DNA, and includes an improved method for enrichment of replication intermediates, compared to previously used BND-cellulose columns.

**Key words** Electron microscopy, DNA replication, Psoralen crosslinking, In vivo replication intermediates, Replication fork reversal, ssDNA, Nucleosome position

---

### 1 Introduction

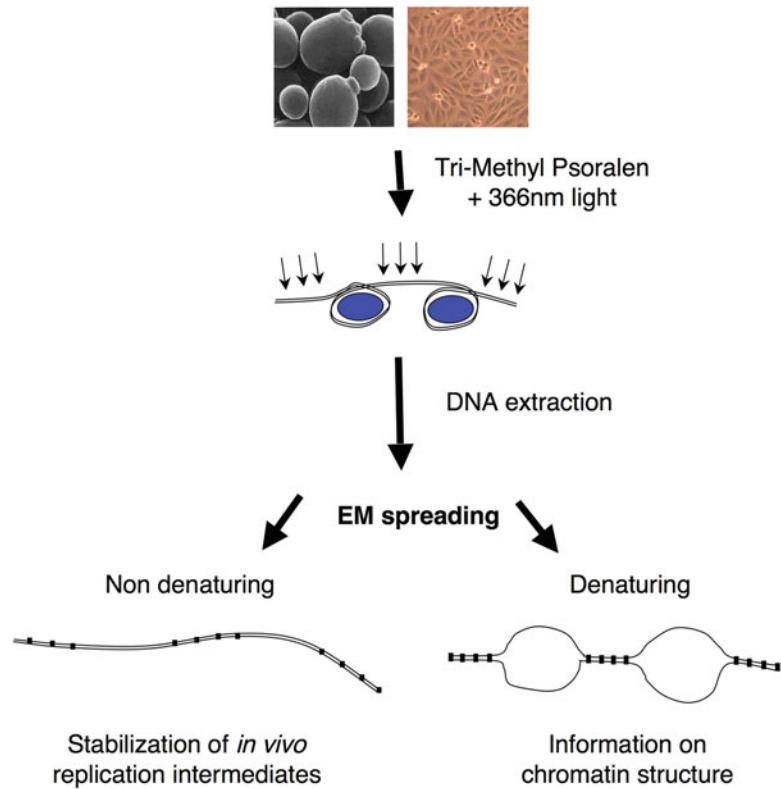
Visualizing DNA replication intermediates (RIs) by structural approaches has proven invaluable to complement standard cell and molecular biology studies on the DNA replication process. Owing to the high magnification that can be achieved, transmission electron microscopy (EM) has helped uncovering the fine architecture of DNA replication forks. Since its original establishment by Dr. José Sogo at the ETH Zurich, this EM approach improved our understanding of the physiological process of genomic or episomal DNA duplication, in different model systems ranging from bacteriophages to mammalian cells [1–4]. The same approach was later extensively used to investigate specific conditions of replication stress in yeast cells, revealing profound changes of RI architecture upon fork stalling and other types of genotoxic stress [5–9]. More recently, several studies have made use of this technique in higher eukaryotic systems, uncovering the remodeling of replication forks

as a global response to cancer-relevant replication stress and identifying cellular factors orchestrating such a response [10–19]. Overall, this approach has played a pivotal role in providing structural insight in the DNA replication process and has showed the potential to confirm, disprove, or refine long-standing models and dogmas in this field.

This chapter will focus on the experimental procedures required to extract genomic RIs from yeast, mammalian cells, and *Xenopus* egg extracts and analyze their fine architecture in vivo using the Transmission Electron Microscope (TEM). We refined several experimental steps of previously published protocols [20, 21]. In particular, we optimized a new protocol for RI enrichment (*see* Subheading 3.7), which no longer relies on BND cellulose and which showed improved enrichment of RIs over linear non-replicating DNA fragments, increasing the speed of the analysis. The technique described here is in principle reproducible in any lab, provided that a TEM and a proper high vacuum evaporator of carbon and platinum/carbon are available.

A crucial prerequisite for these investigations is in vivo psoralen crosslinking of DNA, achieved by repetitive exposure of living cells to tri-methyl-psoralen (TMP), followed by irradiation pulses with 365–366 nm monochromatic light. In vivo psoralen crosslinking is critical for two reasons (Fig. 1): (a) the formation of inter-strand crosslinks induced by this treatment prevents branch migration of cellular DNA and generally stabilizes in vivo RIs during DNA extraction and de-proteinization; (b) the inaccessibility of nucleosomal DNA to the crosslinks provides an opportunity (by proper modification of the technique, *see* Subheading 3.9.2) to obtain important information on nucleosome positioning in vivo and nucleosome dynamics at DNA replication forks [2, 5, 22–24]. Under denaturing conditions, DNA strands are separated wherever a nucleosome was positioned in vivo and are kept together by the interstrand crosslinks induced in the linker DNA (Fig. 1), giving to double-stranded DNA (dsDNA) the appearance of a string of single-stranded DNA (ssDNA) bubbles (Figs. 1 and 8). In fact, this technique has not only been coupled to EM, but also to more standard molecular biology approaches to study in vivo chromatin structure [25].

After in vivo-psoralen crosslinking, genomic DNA is extracted from the cells by standard procedures, minimizing mechanical forces to avoid the shearing of chromosomal DNA. DNA replication intermediates are then enriched by binding, washing, and elution in a QIAGEN-genomic tip 20 column, taking advantage of the high affinity of this resin to ssDNA, which is invariably present at DNA replication forks. With optimal stringency, this enrichment procedure allows an approximately 20-fold enrichment of RIs, although the majority of the recovered molecules are still linear duplex DNA (*see* Subheading 3.11, step 3). The non-



**Fig. 1** Graphic representation of the psoralen crosslinking procedure and advantages for studies on DNA replication intermediates and nucleosome dynamics (see Subheading 1 for details). For the sake of clarity, linker- and nucleosomal-DNA are not drawn to scale in the graphical representation

saturation conditions described here for RI enrichment exclude that the recovered RI pool is biased toward molecules containing longer ssDNA stretches.

Once the DNA sample is enriched for RIs, it is usually concentrated in size-exclusion columns and used for protein-free DNA spreadings in the presence of the cationic detergent benzyltrimethylammoniumchloride (BAC), using water as hypophase, with a modified version of the method originally described by Vollenweider et al. [26]. The low molecular weight of this spreading agent (compared to protein-based methods) generally allows a better visualization of details (i.e., secondary structures) along the DNA molecules and an easier identification of ssDNA regions.

The monolayer nucleic acid film is then absorbed to freshly prepared carbon-coated grids and stained with uranyl acetate to improve the contrast of DNA molecules (in particular of ssDNA, [26, 27]). The grids are then subjected to flat angle-rotary shadowing with Platinum, which allows for the visualization of individual DNA molecules over the background granularity of the carbon

support. The reproducibility of the shadowing angle is a crucial point for the success of this technique. Once a high quality sample is obtained (in terms of molecule concentration, unfolding, and contrast over the background), a high number of RIs are photographed and carefully analyzed for important structural features (presence of ssDNA, secondary structures, nucleosome dynamics, etc.). Contour length measurements on the digital files complement the visual investigation and lead to accurate statistical analysis of the RI population analyzed.

---

## 2 Materials

### 2.1 *In Vivo* Psoralen Crosslinking (*S. cerevisiae*)

1. Ice-cold deionized water (*S. cerevisiae*).
2. Standard Petri dish, diameter 8.5 cm (*S. cerevisiae*).
3. Ice-cold 1 × PBS (mammalian cells).
4. 96-Well plates (Xenopus egg extracts).
5. Ice-cold EB buffer (Xenopus egg extracts).
6. Tissue culture dishes, 60 × 15 mm (mammalian cells).
7. 4,5',8-Trimethylpsoralen (TMP) stock solution: dissolve 200 µg/ml TMP (Sigma, cat. T6137) in Ethanol 100%. Stir extensively until the compound is dissolved completely. The solution can be stored at 4 °C for at least 1 year. Stir briefly at RT before each usage. Due to its DNA-modifying potential, TMP-containing solutions should be handled with gloves, lab-coat, and protection glasses.
8. Biolink Crosslinker 365 nm, cat. B89110 (*see Note 1*).
9. Monochromatic 365 nm lamps for Biolinker (replacement UV bulbs, 8 W, cat. B89270).
10. UVP UVX Radiometer, with 365 nm sensor.
11. Freezing pack, flat, 1.5–2 cm thick.
12. Flat metal support (1 cm), precooled at –20 °C. The surface should be large enough to accommodate several Petri dishes (to crosslink more samples at the same time), but small enough to be easily inserted and removed from the available Crosslinker. This support is placed on the top of the freezing pack, to prevent heating of the samples during the incubation/irradiation cycles required for psoralen crosslinking.

### 2.2 *In Vivo* Psoralen Crosslinking (Mammalian Cells)

1. Ice-cold deionized water (*S. cerevisiae*).
2. Standard Petri dish, diameter 8.5 cm (*S. cerevisiae*).
3. Ice-cold 1 × PBS (mammalian cells).
4. 96-Well plates (Xenopus egg extracts).



5. Ice-cold EB buffer (*Xenopus* egg extracts).
6. Tissue culture dishes, 60 × 15 mm (mammalian cells).
7. 4,5',8-Trimethylpsoralen (TMP) stock solution: dissolve 200 µg/ml TMP (Sigma, cat. T6137) in Ethanol 100%. Stir extensively until the compound is dissolved completely. The solution can be stored at 4 °C for at least 1 year. Stir briefly at RT before each usage. Due to its DNA-modifying potential, TMP-containing solutions should be handled with gloves, lab-coat, and protection glasses.
8. Biolink Crosslinker 365 nm, cat. B89110 (*see Note 1*).
9. Monochromatic 365 nm lamps for Biolinker (replacement UV bulbs, 8 W, cat. B89270).
10. UVP UVX Radiometer, with 365 nm sensor.
11. Freezing pack, flat, 1.5–2 cm thick.
12. Flat metal support (1 cm), precooled at –20 °C. The surface should be large enough to accommodate several Petri dishes (to crosslink more samples at the same time), but small enough to be easily inserted and removed from the available Crosslinker. This support is placed on the top of the freezing pack, to prevent heating of the samples during the incubation/irradiation cycles required for psoralen crosslinking.

### **2.3 Psoralen Crosslinking of Sperm DNA (*Xenopus* Egg Extracts)**

1. Ice-cold deionized water (*S. cerevisiae*).
2. Standard Petri dish, diameter 8.5 cm (*S. cerevisiae*).
3. Ice-cold 1 × PBS (mammalian cells).
4. 96-Well plates (*Xenopus* egg extracts).
5. Ice-cold EB buffer (*Xenopus* egg extracts).
6. Tissue culture dishes, 60 × 15 mm (mammalian cells).
7. 4,5',8-Trimethylpsoralen (TMP) stock solution: dissolve 200 µg/ml TMP (Sigma, cat. T6137) in Ethanol 100%. Stir extensively until the compound is dissolved completely. The solution can be stored at 4 °C for at least 1 year. Stir briefly at RT before each usage. Due to its DNA-modifying potential, TMP-containing solutions should be handled with gloves, lab-coat, and protection glasses.
8. Biolink Crosslinker 365 nm, cat. B89110 (*see Note 1*).
9. Monochromatic 365 nm lamps for Biolinker (replacement UV bulbs, 8 W, cat. B89270).
10. UVP UVX Radiometer, with 365 nm sensor.
11. Freezing pack, flat, 1.5–2 cm thick.
12. Flat metal support (1 cm), precooled at –20 °C. The surface should be large enough to accommodate several Petri dishes (to crosslink more samples at the same time), but small enough

to be easily inserted and removed from the available Cross-linker. This support is placed on the top of the freezing pack, to prevent heating of the samples during the incubation/irradiation cycles required for psoralen crosslinking.

**2.4 Genomic DNA  
Extraction  
(*S. cerevisiae*) (See  
Note 2)**

1. Spheroplasting buffer: 1 M sorbitol, 100 mM EDTA pH 8.0, 0.1%  $\beta$ -Mercaptoethanol, 100 U/ml Lyticase. This buffer is freshly prepared from the following stocks: 2 M Sorbitol; 0.5 M EDTA pH 8.0; 14.3 M  $\beta$ -Mercaptoethanol (pure liquid); Lyticase (Sigma, cat. L4025) 1000 U/ml in water (store 1 ml aliquots at  $-20^{\circ}\text{C}$ ).
2. RNaseA stock solution (10 mg/ml in Tris-HCl pH 7.4) prepared according to [28], freeze aliquots at  $-20^{\circ}\text{C}$ .
3. Proteinase K stock solution (20 mg/ml in water), freeze aliquots at  $-20^{\circ}\text{C}$ .
4. Chloroform/isoamylalcohol 24:1. This solution can be stored at RT in a hood for unlimited time.
5. Kimble glass tubes (KIMBLE HS No. 45500-30).
6. Solution I: 2% w/v CTAB (cetyltrimethylammoniumbromide), 1.4 M NaCl, 100 mM Tris-HCl pH 7.5, 25 mM EDTA pH 8.0 (*see Note 3*). Filter Solution I and Solution II to avoid the formation of aggregates during DNA preps. These solutions can be usually stored for 2–3 months at RT. They should be refiltered or freshly prepared if a precipitate is detectable.
7. Solution II: 1% CTAB, 50 mM Tris-HCl pH 7.5, 10 mM EDTA (*see Note 3*).
8. Solution III: 1.4 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA.
9. Isopropanol.
10. 70% EtOH, stored at RT.
11.  $1\times$  TE buffer.

**2.5 Genomic DNA  
Extraction  
(Mammalian Cells)**

1. Lysis buffer (QIAGEN buffer C1: 1.28 M Sucrose; 40 mM Tris-HCl pH 7.5; 20 mM  $\text{MgCl}_2$ ; 4% Triton X-100). Dissolve 483.14 g sucrose, 4.06 g  $\text{MgCl}_2\cdot 6\text{H}_2\text{O}$ , and 4.84 g Tris base in 680 ml ddH<sub>2</sub>O. Add 42 g Triton X-100 (100%). Adjust the pH to 7.5 with HCl. Adjust the volume to 1 L with ddH<sub>2</sub>O. Keep at  $4^{\circ}\text{C}$ .
2. Digestion buffer (QIAGEN buffer G2: 800 mM Guanidine-HCl; 30 mM Tris-HCl pH 8.0; 30 mM EDTA pH 8.0; 5% Tween-20; 0.5% Triton X-100). Dissolve 76.42 g guanidine HCl, 11.17 g  $\text{Na}_2\text{EDTA}\cdot 2\text{H}_2\text{O}$ , and 3.63 g Tris base in 600 ml of ddH<sub>2</sub>O. Add 250 ml 20% Tween-20 solution and 50 ml 10% Triton X-100 solution. Adjust the pH to 8.0 with NaOH. Adjust the volume to 1 L with ddH<sub>2</sub>O. Store at RT.

3. Ice-cold ddH<sub>2</sub>O.
4. Ice-cold 1× PBS.
5. Proteinase K stock solution 20 mg/ml in ddH<sub>2</sub>O (Roche, cat. 03115852001). Store at −20 °C.
6. Chlorophorm/isoamylalcohol 24:1.
7. Isopropanol.
8. Ethanol 70%.
9. 30 ml glass centrifugation tubes, KIMBLE HS No. 45500-30.
10. Eppendorf centrifuge 5810R, rotor A-4-81.
11. Sorvall Evolution RC, rotor HB-6 swinging (8000 rpm = 10,459 × *g*).

**2.6 Sperm DNA  
Extraction (*Xenopus*  
Egg Extracts)**

1. EB EDTA buffer: 50 mM Hepes-KOH pH −7.5, 100 mM KCl, 2.5 mM MgCl<sub>2</sub>, 1 mM EDTA.
2. Cy3™-dCTP (GE HealthCare cat. PA35021). Dilute 1:50 in EB-EDTA buffer and store in aliquots at −20 °C.
3. EB EDTA sucrose buffer: EB EDTA +30% sucrose.
4. Proteinase K stock solution 20 mg/ml in ddH<sub>2</sub>O (Roche cat. 03115852001). Store at −20 °C.
5. Chlorophorm/isoamylalcohol 24:1.
6. Ethanol 70%.
7. 1× TE buffer.

**2.7 DNA Digestion  
and Enrichment  
of Replication  
Intermediates**

1. RNase A (Ribonuclease A Type I-AS, Sigma-Aldrich, R5503) 10 mg/ml. Store in aliquots at −20 °C.
2. RNase III (ShortCut RNase III, NEB M0245) 10 mg/ml. Store aliquots at −20 °C.
3. QBT equilibration buffer. Dissolve 43.83 g NaCl, 10.46 g MOPS (free acid) in 800 ml distilled water. Adjust the pH to 7.0 with NaOH. Add 150 ml pure isopropanol and 15 ml 10% Triton X-100 solution (v/v). Adjust the volume to 1 L with distilled water.
4. Stock solution 10 mM Tris-HCl pH 8.0, 300 mM NaCl. The solution can be prepared in large volume (50 ml), stored at RT in plastic tubes, and reused for different experiments. 5 M NaCl and 1 M Tris-HCl pH 8.0 stocks are required to prepare this and the following solutions.
5. Stock solution 10 mM Tris-HCl pH 8.0, 900 mM NaCl can be prepared in large volume (50 ml), stored at RT in plastic tubes, and reused for different experiments.
6. Stock solution 10 mM Tris-HCl pH 8.0, 1 M NaCl can be prepared in large volume (50 ml), stored at RT in plastic tubes, and reused for different experiments.

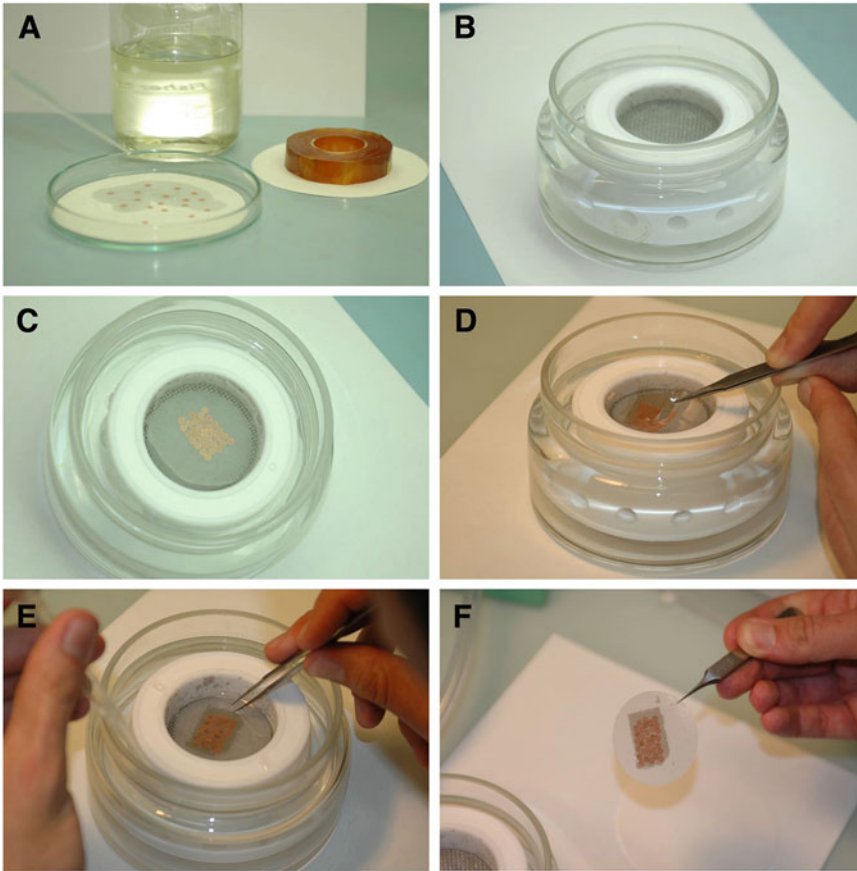
7. Stock solution 10 mM Tris-HCl pH 8.0, 1 M NaCl, 1.8% caffeine (w/v) can be prepared in large volume (50 ml), but requires extensive incubation at 50 °C and stirring or vortexing to help caffeine dissolution. The stock needs to be filtered (syringe filter, 0.2 µM), can be stored at RT in plastic tubes, and can be reused for different experiments during 4–6 months. A new stock should be prepared when a precipitate is detectable.
8. QIAGEN Genomic-tip 20/G (QIAGEN, Cat. No. 10223).
9. Millipore size-exclusion columns: Amicon ultra 100K membrane (Millipore, cat. UFC510096).
10. 1 × TE buffer.

### **2.8 Preparation of Carbon-Coated Grids**

1. High Vacuum Evaporator MED 020 (BalTec), with two EK 030 electron guns (C and Pt/C), quartz crystal, thin film monitor QSG 100, and control unit EVM 030 (*see Note 4*).
2. Mica Sheets (Plano, mica high grade quality V1, 25 × 76 mm, cat. 56).
3. Wolfram cathodes for electron guns (BALTIC, BP2317).
4. Carbon rods (BALTIC, 3 × 50 mm, cat. BP2217 P = 6).
5. Scotch solution: 20–30 cm of Scotch tape in 100 ml of chloroform in a glass bottle (Fig. 2a). After quick stirring of the bottle, the chloroform turns yellowish, dissolving the tape adhesive (the cellophane support does not dissolve). This solution can be stored at RT for at least 1 year.
6. Supporting Teflon-wire mesh stand (Fig. 2b; *see Note 5*).
7. Filter paper circles (Macherey-Nagel, 4310045, diameter: 45 mm).
8. Filter paper circles (Macherey-Nagel, 431009, diameter: 90 mm).
9. EM-grade water (*see Note 6*).
10. Copper 3.05 mm grids, 400 mesh (Plano, cat. G2400C; *see Note 7*).

### **2.9 DNA Spreading by the “BAC Method”**

1. P2 pipette.
2. EM-grade water (*see Note 6*).
3. EtBr stock (10 mg/ml, Sigma, cat. E1510). It can be stored at 4 °C for at least 1 year.
4. EtBr working solution (33.3 µg/ml). It is freshly prepared before each set of spreadings, adding 1 µl of EtBr stock (10 mg/ml) to 300 µl of EM-grade water.
5. Formamide (pure, Sigma-Aldrich, cat. 47680).
6. Glyoxal: 40% solution in water (Merck, cat. 4910).



**Fig. 2** Series of photographs showing crucial steps in the preparation of carbon-coated grids (*see* Subheading 3.8 for details)

7. BAC stock: Benzyltrimethylammonium chloride (Bayer-Leverkusen, n-alkyl mixture: C<sub>12</sub>H<sub>25</sub>, 60%; C<sub>14</sub>H<sub>29</sub>, 40%; *see* **Note 8**) 0.2% w/v in Formamide. This solution can be stored at RT for at least 1 year.
8. BAC working solution (BAC 1:20): just before each set of spreadings, the BAC stock is diluted 1:20 in 1× TE (20 μl of the diluted stock are prepared in a microfuge tube and are normally sufficient for all spreadings performed during the day).
9. Ethanol 100%, molecular biology grade.
10. Uranyl acetate stock: uranyl acetate (UrAc, Fluka, cat. 73943) 5 mM in HCl 5 mM. This stock solution can be stored at 4 °C for at least 1 year. Uranyl acetate has a typical radioactivity of 0.37–0.51 μCi/g. This mild radioactivity level is not sufficient to be harmful while the material remains external to the body. The use of standard protective clothing (gloves, glasses, lab coat) is therefore sufficient to work safely.

11. Uranyl Acetate working solution (0.5 mM UrAc, 0.5 mM HCl, 90% Ethanol): just before each set of spreadings, the UrAc stock is diluted 1:10 in 100% Ethanol.
12. Round filter paper (Macherey-Nagel, 431009, diameter: 90 mm).
13. Tissue culture dishes, 60 × 15 mm (*see Note 9*).
14. Mica sheets (Plano, mica high grade quality V1, 25 × 76 mm, cat. 56).
15. Graphite powder (BalTec, cat. LZ 02096 VN).
16. Fine tweezers with bent points (*see Note 10*) or Dumont pinzette (Plano T539).

### **2.10 Platinum-Carbon Rotary Shadowing**

1. High Vacuum Evaporator MED 020 (BalTec; *see Note 4*), with two EK 030 electron guns (C and Pt/C), quartz crystal, thin film monitor QSG 100, control unit EVM 030, rocking rotary stage, and specimen table (54 mm) for 20 grids (3.05 mm, Fig. 4a; *see Note 11*).
2. Carbon/Platinum rods: Carbon rods 2 × 20 mm (BALTIC BP2260); platinum insets (BALTIC BP2261).
3. Micrometric control of rocking rotary stage angle: the standard knob of the MED020 is substituted by a Precision Rotation Platform PR01 with Adapter Plate PR01A, Thorlabs, Newton, NJ, USA (*see Note 12*; Fig. 4c and d, black arrow).

### **2.11 Visualization at the Transmission Electron Microscope, Contour Length Measures, and Statistics**

1. Transmission Electron Microscope, connected to a computer-driven CCD-camera. Minimal resolution required is 1k × 1k if the camera is mounted on the bottom port, while it should preferably reach 2k × 3.5k if mounted on the side 35 mm port.
2. Software for Camera Control and storage of acquired images in original and \*.tiff format (i.e., Digital Micrograph).
3. Software package to produce a composite image (montage) from adjacent images.
4. ImageJ or other software allowing contour length measurements on \*.tiff files.
5. Standard statistic applications (i.e., Microsoft Office Excel).
6. Large hard disk capacity or external hard disk for storage of a high number of large image files (*see Note 13*).

---

## **3 Methods**

### **3.1 In Vivo Psoralen Crosslinking (*S. cerevisiae*)**

1. Samples are typically collected at different time points during synchronization experiments (*see Note 14*). Every sample corresponds to  $4 \times 10^9 - 1 \times 10^{10}$  cells (400 ml of culture

1– $2.5 \times 10^7$  cells/ml). Spin cells down at  $3200 \times g$  for 10 min. Resuspend the cells in ice-cold water, transfer the suspension to 50 ml tubes. Spin cells down at  $3200 \times g$  for 5 min at 4 °C (*see Note 15*).

2. Resuspend the pellet in 20 ml of ice-cold water. Transfer cells into an 8.5 cm diameter Petri dish. Several samples can be simultaneously crosslinked with the following procedure.
3. Install the five monochromatic 365 nm lamps in the crosslinker. Make sure that they are all properly inserted and that they all light up when starting the crosslinker.
4. Add 1 ml TMP stock solution (10 µg/ml final concentration) to the cell suspension in the Petri dish. Mix well with a pipette and incubate for 5 min in the dark on the precooled metal support (*see Note 16*). Place the precooled metal support with the Petri dishes on the freezing pack. Irradiate the sample for 3 min.
5. Repeat **step 4** once (*see Note 17*).
6. Transfer the cell suspension to a 50 ml tube and keep it on ice. Wash the dish twice with 1 ml water to remove all the cells and pool the washes in the same 50 ml tube. Spin down the cells at  $3200 \times g$  for 10 min. Use the cell pellet for DNA extraction (*see Subheading 3.4*).

### **3.2 In Vivo Psoralen Crosslinking (Mammalian Cells)**

1. Samples can be collected from asynchronously growing cells or at different time points in synchronization experiments (*see Note 14*). Every sample corresponds to  $2.5\text{--}5.0 \times 10^6$  cells ( $150 \times 20$  mm tissue culture dish, 50–80% confluency for U2OS cells). After standard trypsinization (or collection, for cells in suspension), transfer the cells to 15 ml Falcon tubes and spin them down at  $600 \times g$  for 5 min. Wash the cell pellet once (by resuspension/centrifugation) with 5 ml ice-cold  $1 \times$  PBS (*see Note 15*).
2. Resuspend the pellet in 10 ml ice-cold  $1 \times$  PBS. Transfer cells into a tissue culture dish  $60 \times 15$  mm. Several samples can be simultaneously crosslinked with the following procedure.
3. Insert the five monochromatic 365 nm lamps in the crosslinker. Make sure that they are all properly inserted and that they all light up when starting the crosslinker.
4. Add 0.5 ml of TMP stock solution (10 µg/ml final concentration) to the cell suspension in the Petri dish. Mix well with a pipette and incubate for 5 min in the dark on the precooled metal surface (*see Note 16*). Place the metal support with the Petri dishes on the top of the freezing pack. Irradiate the sample for 3 min.
5. Repeat **step 4** once (*see Note 17*).

6. Transfer the suspension back to a 50 ml tube and keep it on ice. Wash the dish twice with 1 ml 1× PBS to remove all cells. Spin down the cells at  $600 \times g$  for 5 min. Resuspend the cell pellet in 2 ml 1× PBS and proceed with DNA extraction (*see* Subheading 3.5).

### 3.3 Psoralen Crosslinking of Replicating Sperm DNA (*Xenopus* Egg Extracts)

1. Samples are typically collected 30–60 min after sperm nuclei addition in the *Xenopus* extracts. Replicating interphase extracts are prepared as described in [29]. Every sample corresponds to 200–300  $\mu$ l of extract to which 4000 nuclei/ $\mu$ l are added.
2. Split the sample into 100  $\mu$ l aliquots in microfuge tubes and incubate at 23 °C.
3. (Optional) To control the timing of replication in the extract, add 5  $\mu$ l Cy-3 dCTP to 100  $\mu$ l of the sample (extract + sperm) and check visible incorporation by standard immunofluorescence microscopy (*see* **Note 18**).
4. When 80–90% of the nuclei start incorporating Cy-3 dCTP (ideally 45–50 min after sperm addition), arrest the DNA replication by diluting each 100  $\mu$ l aliquot with 200  $\mu$ l of cold EB-EDTA buffer and incubate on ice.
5. Pool the aliquots corresponding to each sample.
6. Underlay the samples with 2 volumes of cold EB-EDTA Sucrose Buffer.
7. Spin at  $8600 \times g$  at 4 °C for 5 min.
8. Remove the supernatant carefully so as not to disturb the pellet.
9. Resuspend the pellet in 100  $\mu$ l of ice-cold EB-EDTA using a P200 pipette with cut tips.
10. Transfer the samples to a precooled round-bottom microtiter plate on a precooled metal support.
11. Insert the five monochromatic 365 nm lamp in the crosslinker. Make sure that they are all properly inserted and that they all light up when starting the crosslinker.
12. Add 5  $\mu$ l of TMP stock solution (10  $\mu$ g/ml final concentration) to each 100  $\mu$ l nuclei suspension and mix by pipetting with cut tips. Incubate for 5 min in the dark on a precooled metal support. Place the metal support with the microtiter plate on the top of the freezing pack and irradiate the sample for 3 min.
13. Repeat **step 12** two more times.
14. Recover the nuclei suspension from the wells into microfuge tubes. Centrifuge at  $5000 \times g$  for 5 min and resuspend nuclei in 300  $\mu$ l of EB-EDTA buffer using a precut 1 ml tip. Proceed with DNA extraction (*see* Subheading 3.6).



**3.4 Genomic DNA  
Extraction by CTAB  
Method (*S. cerevisiae*)  
(See Note 2)**

1. Resuspend the cells in 5 ml of Spheroplasting buffer. Incubate at 30 °C for 45 min. Invert the tube several times during incubation. Spin down the spheroplasts at 6000 × *g* for 10 min.
2. Resuspend the spheroplasts in 2 ml of distilled water and rapidly add 2.5 ml of Solution I, 200 µl of RNaseA (10 mg/ml). Incubate for 30 min at 50 °C, then add 200 µl of Proteinase K (20 mg/ml). Gently mix the sample.
3. Incubate for 1.5–2 h at 50 °C. Invert the tube several times during incubation. Note that the solution has to become clear (no clumps!); if necessary, use precut 1 ml tips to break clumps. If clumps are still present after 2 h, add 100 µl of Proteinase K (20 mg/ml) and incubate O.N. at 30 °C.
4. Centrifuge for 10 min at 3200 × *g*; keep the pellet (cellular debris) for further extraction (*see* Subheading 3.4, **steps 9–11**). Transfer the supernatant carefully to a 15 ml tube containing 2.5 ml of chloroform/isoamylalcohol 24:1 at RT.
5. Mix by inverting the tubes several times and spin at RT at 3200 × *g* for 10 min. Note that a white protein layer is formed between the two phases. Carefully remove the clear upper phase and transfer it to a 30 ml Kimble glass tube.
6. Gently add 10 ml (2 volumes) of Solution II, cover with parafilm and invert several times. The solution should slowly become turbid. If necessary, incubate for 10–15 min at RT until some turbidity is detectable.
7. Centrifuge for 10 min at 12,000 × *g* in a proper swinging bucket rotor and discard the supernatant.
8. Add 2 ml of Solution III and incubate the solution briefly at 37 °C until the pellet is completely dissolved.
9. Resuspend the pellet from **step 4** in 2 ml of Solution III, mix vigorously, and incubate at 50 °C for 1 h (use precut 200 µl tips to help complete resuspension). Check that the solution is finally homogeneous.
10. Transfer the solution carefully to a 15 ml tube containing 1 ml of chloroform/isoamylalcohol 24:1 at RT.
11. Mix by inverting the tubes several times and spin at RT at 3200 × *g* for 10 min. Note that a white protein layer is formed between the two phases. Carefully remove the clear upper phase (ca. 2 ml) and pool it in the Kimble glass tube from **step 8** (4 ml final volume).
12. Precipitate the DNA by adding 4 ml of isopropanol (RT). Gently mix the sample to ensure a proper DNA precipitation. Spin down the DNA at 12,000 × *g* for 10 min in a swinging bucket rotor.
13. Discard the supernatant and wash the pellet with 1 ml of Ethanol 70% (RT).

14. Remove as much ethanol as possible and briefly spin the pellet to accumulate residual ethanol at the bottom of the tube. Remove it using a 200  $\mu$ l pipette. Further dry the pellet by briefly incubating the open Kimble glass tube in a 37 °C water bath.
15. Add 200  $\mu$ l of 1 $\times$  TE to the dried pellet. Cover the Kimble glass tube with parafilm and incubate for 30 min at 37 °C (or O/N at RT) to ensure proper resuspension of the DNA pellet. Spin briefly and transfer the DNA solution to a microfuge tube using a precut 200  $\mu$ l tip, to avoid shearing of genomic DNA.
16. Check the quality and the concentration of DNA preps, as well as possible RNA contamination (*see Note 21*) by agarose-gel electrophoresis. The standard yield, starting with  $4 \times 10^9$ – $1 \times 10^{10}$  cells, is therefore 10–60  $\mu$ g of genomic DNA from each sample.

### **3.5 Genomic DNA Extraction (Mammalian Cells)**

1. Add to the cell suspension in 1 $\times$  PBS (Subheading 3.2, step 6) 2 ml of ice-cold lysis buffer and 6 ml of ice-cold ddH<sub>2</sub>O.
2. Mix by inverting the tube several times and incubate on ice for 10 min.
3. Spin the lysed cells at 4 °C for 15 min at 1300  $\times g$ . Discard the supernatant.
4. Add 1 ml of ice-cold lysis buffer and 3 ml of ice-cold ddH<sub>2</sub>O. Resuspend the pelleted nuclei completely by vortexing.
5. Spin the lysed cells at 4 °C for 15 min at 1300  $\times g$ . Discard the supernatant completely.
6. Add 100  $\mu$ l of ice-cold PBS and resuspend the nuclei with a cut 200  $\mu$ l tip. Make sure that no clumps of cells are left.
7. Add 5 ml of digestion buffer. It is crucial NOT to VORTEX or rotate the tube to avoid clumping of the cells in the center of the tube. Add 200  $\mu$ l of proteinase K stock and incubate at 50 °C until the solution is clear (1–2 h).
8. Let the sample cool down to RT, then transfer it into a 50 ml tube containing 5 ml of chloroform/isoamylalcohol. Close the tube properly and invert it vigorously but carefully 30 times (solution turns milky). Immediately pour sample into a glass centrifugation tube.
9. Centrifuge at 4 °C for 20 min at 10,500  $\times g$  (phase separation).
10. Carefully transfer upper phase into a new glass tube using a cut 1 ml tip. Add the same volume of isopropanol and mix well to precipitate the DNA.
11. Centrifuge at 4 °C for 10 min at 10500  $\times g$ .
12. Wash DNA with 70% ethanol. Spin at 4 °C for 5 min at 10,500  $\times g$ .

13. Air dry the pellet and resuspend in  $1 \times$  TE buffer (200–400  $\mu$ l). Pay attention not to overdry DNA pellet to prevent irreversible “agglutination” of DNA molecules. The standard yield is 10–50  $\mu$ g of genomic DNA for  $2.5\text{--}5.0 \times 10^6$  cells.

### **3.6 DNA Extraction from Sperm Nuclei (*Xenopus* Egg Extracts)**

1. Add to the crosslinked nuclei (Subheading 3.3, step 14) 1.5  $\mu$ g/ml final concentration of proteinase K and incubate for 2 h at 50 °C.
2. Add equal volume of phenol:chloroform:isoamylalcohol—25:24:1 to the above solution and shake vigorously. Spin at  $18,500 \times g$  for 10 min.
3. Transfer the supernatant to a new microfuge with a cut 1 ml tip.
4. Add equal volume of 100% isopropanol to the supernatant to precipitate the DNA. Incubate at 4 °C for 10 min. Spin the sample at  $18,500 \times g$  for 10 min at 4 °C. Discard the supernatant.
5. Wash pellet with 500  $\mu$ l of 70% ethanol and spin down the sample for 5 min.
6. Discard the supernatant and briefly centrifuge at  $1700 \times g$ . Remove the residual ethanol carefully with a 20  $\mu$ l pipette.
7. Incubate the tube at 37 °C to evaporate the excess ethanol.
8. Resuspend the pellet in 100  $\mu$ l TE.
9. Assess quality and quantity of the DNA by UV spectrophotometry and by gel electrophoresis.

### **3.7 DNA Digestion and Enrichment of Replication Intermediates (See Note 19)**

1. 10–15  $\mu$ g (yeast, mammalian cells) or 5  $\mu$ g of DNA (*Xenopus*) is digested with 50–100 U of restriction enzyme in the proper buffer for 3–5 h at 37 °C. Standard restriction enzymes used are PvuI for *S. cerevisiae* genomic DNA, PvuII HF for mammalian genomic DNA, and NdeI for *Xenopus* sperm DNA (see Note 20). The volume of the restriction reaction is normally set to 250  $\mu$ l, but can be increased to account for more diluted DNA preps (the reaction mix is anyway diluted in step 5 in Subheading 3.7). If necessary, add small amounts of the proper RNase enzymes (see Note 21).
2. Equilibrate QIAGEN-tip 20 by applying 1 ml Buffer QBT. Allow the column to empty by gravity flow. Flow of buffer will begin automatically by reduction in surface tension due to the presence of the detergent in the equilibration buffer (Triton X-100).
3. Wash the QIAGEN-tip 20 column three times with 1 ml of 10 mM Tris–HCl pH 8, 1 M NaCl.
4. Equilibrate the column six times with 1 ml of 10 mM Tris–HCl pH 8, 300 mM NaCl.

5. After the required incubation time (3–5 h, *see step 1*), adjust the digestion mix to 300 mM NaCl final concentration, by adding 5 M NaCl stock (check the restriction buffer composition). Adjust the final volume to 600  $\mu$ l with 10 mM Tris–HCl pH 8, 300 mM NaCl.
6. Apply the pre-equilibrated digestion mix and allow it to enter the resin by gravity-flow.
7. Wash the QIAGEN-tip 20 with two times 1 ml 10 mM Tris–HCl pH 8, 850 mM NaCl. Collect the flow-through in 2 ml tube and save for an analytical gel.
8. Add to the column 600  $\mu$ l of 10 mM Tris–HCl pH 8, 1 M NaCl, 1.8% caffeine at 50 °C. Collect the flow-through, enriched in RIs in 1.5 ml tube.
9. Purify and concentrate the DNA (removing residual RNA, small linear fragments, and microscopic dirty particles) using an Amicon size-exclusion column.  
Load the 600  $\mu$ l of elution from **step 8** into Amicon column. Spin the column for 8 min at 9000 rcf.
10. Wash the membrane with 200  $\mu$ l of 1 $\times$  TE and spin the column for 5 min at 9000 rcf. Wash again with 200  $\mu$ l of 1 $\times$  TE and spin the column for 4 min at 8000 rcf until 15–30  $\mu$ l remains.
11. Invert the Amicon filter and short spin into a fresh Amicon tube.  
Note: Transfer the sample to a fresh 1.5 ml eppendorf tube to prevent evaporation of the sample solution.
12. Load a 1  $\mu$ l aliquot on an agarose gel to check DNA quality and concentration. If necessary, adjust the final volume (by adding 1 $\times$  TE or concentrating the sample in a standard vacuum evaporator) to reach the optimal DNA concentration of 10–50 ng/ $\mu$ l (*see Note 22*).

### **3.8 Preparation of Carbon-Coated Grids**

1. Cleave a 2 cm  $\times$  2 cm sheet of mica and place it (with the freshly cleaved surface facing up) on the support plate of the MED 020, at a distance of about 12 cm to the carbon evaporator gun. Place the quartz sensor as close as possible to the mica. Cover the mica with the tilting shutter, properly position the glass vacuum chamber wall, and start the turbomolecular pump of the MED020.
2. At a vacuum of about  $3 \times 10^{-5}$  mbar, preheat the filament of the carbon electron gun. Adjust voltage and current applied to the electron gun to reach a constant evaporation rate of 0.03–0.05 nm/s (detected on the QSG 100 thin film monitor). Open the tilting shutter and start measuring the carbon film thickness on the QSG 100 (*see Note 23*). When the thickness readout is 3.5–4.5 nm (50–70 Hz), close the shutter

and shut off the electron gun. Generally, 2–3 carbon films can be produced in series (*see* **Note 24**).

3. Remove the carbon-coated mica sheet from the MED 020. Carbon-coated mica sheets can be stored at this stage up to 4–5 weeks.
4. When ready to transfer the carbon film on the grids, place the carbon-coated mica sheet (carbon side up) in a Petri dish on wet filter paper and incubate it at 37–42 °C for 30 min–1 h.
5. Spread copper grids on filter paper in a glass Petri dish and make sure that the glossy side of all grids is facing up. 30–35 grids should be used for each 2 cm × 2 cm carbon-coated mica sheet. Place the Petri dish in a hood and, using a Pasteur pipette, rinse each grid with one drop of Scotch solution (Fig. 2a). Air-dry the grids. Repeat the procedure two to three times to make sure that all grids have been extensively rinsed in Scotch solution. The tape adhesive will keep the carbon film attached.
6. Fill the supporting wire mesh stand with EM-grade water (Fig. 2b). A round filter paper (diameter: 45 mm) is submerged in water and 30–35 grids (glossy side up) are placed on its surface in a close and ordered distribution (Fig. 2c).
7. The carbon-coated mica sheet is then removed from the wet filter paper, briefly dried on its lower side (no carbon) by a napkin, and slowly lowered into the water (carbon side up) at an angle of approximately 45° (Fig. 2d), until the carbon film is completely released and floating on the water surface (*see* **Note 25**). Discard the mica support.
8. The carbon film is finally placed on the grids by carefully lowering the water level in the supporting Teflon-wire mesh stand, using an aspirator connected to a vacuum pump. Use tweezers to correctly position the carbon film on the grids, while lowering the water level (Fig. 2e).
9. Once the water has been removed, take the filter paper with the carbon-coated grids (Fig. 2f), cut off the excess of wet filter paper around the grids, and let the carbon-coated grids dry for at least one night before using them for DNA spreading experiments. Although some variability has been observed, carbon-coated grids are usually reliable for DNA absorption for about 4–5 weeks after carbon-film production.

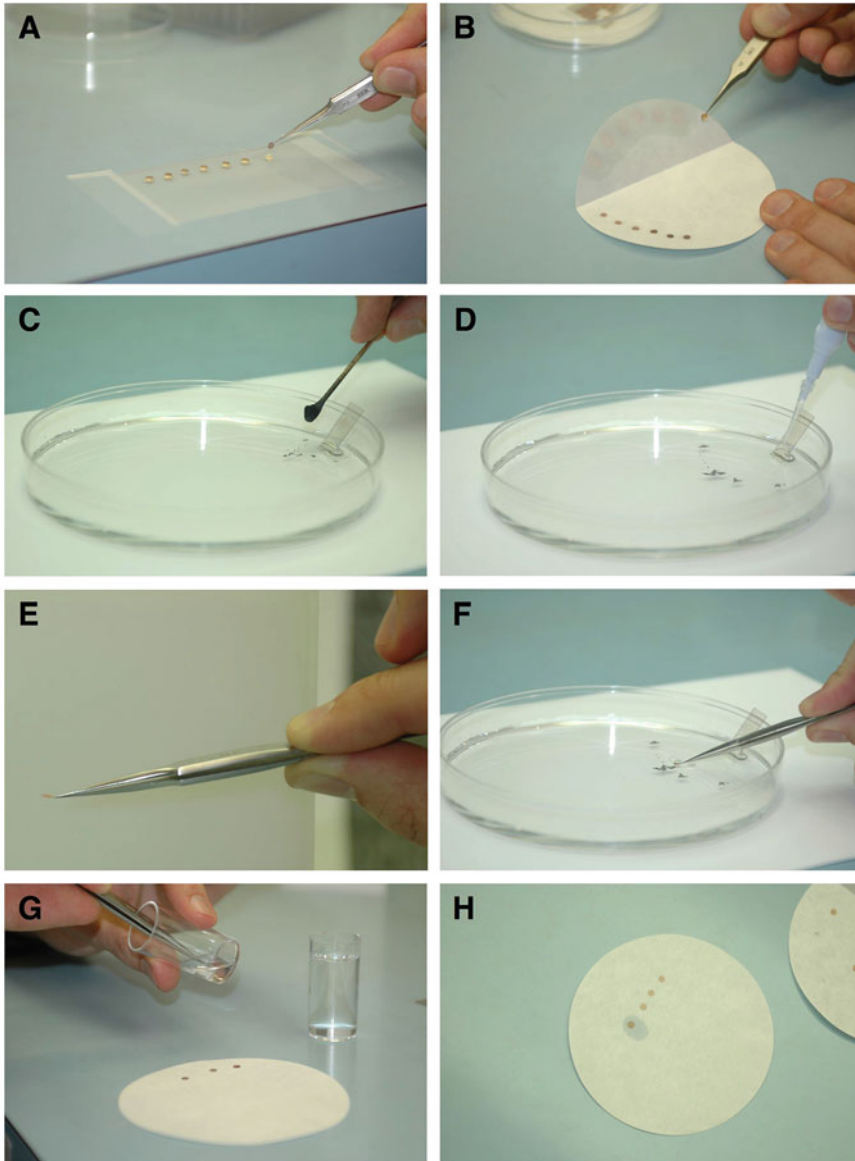
### **3.9 DNA Spreading for EM Visualization**

#### *3.9.1 “Native” DNA Spreading by the “BAC Method”*

1. For each DNA spreading, distribute up to eight droplets (10–15 µl) of EtBr working solution on a piece of parafilm. Carefully place a carbon-coated grid on the top of each drop, with the carbon-side facing the liquid (Fig. 3a; *see* **Note 26**). Incubate grids for 20–45 min. Prevent evaporation by covering the parafilm with the lid of the 15 cm Petri dish. Just before

spreading the sample on the hypophase (Subheading 3.9.1, **step 6**), take each grid, remove the excess of EtBr solution contacting the filter paper, and place the dried grids (carbon side down) on the top of filter paper (Fig. 3b).

2. Using a P2 pipette, mix at the very bottom of a 1.5 ml-microfuge tube: 1  $\mu$ l of Formamide, 0.4  $\mu$ l of BAC 1:20 (*see Note 27*).



**Fig. 3** Series of photographs showing crucial steps in the BAC-DNA spreading method (*see* Subheading 3.9 for details)

3. In a separate drop on the side of the same tube (*see Note 28*), mix: 1.5–3 ng of plasmid DNA of known size (3–10 kb; internal size maker) and 10–50 ng of sample DNA (genomic DNA enriched for Replication Intermediates, *see Subheading 3.7, step 12 and Note 29*). The total volume of the DNA drop should be 1.25  $\mu$ l; usually it consists of 0.25  $\mu$ l plasmid DNA (5 ng/ $\mu$ l) and 1.0  $\mu$ l sample DNA. Smaller sample volumes can be filled up with 1 $\times$  TE buffer (*see Note 29*). In case of very low DNA concentration in the sample, up to 4  $\mu$ l of the DNA sample can be added to the spreading. In this case an equal volume of Formamide should be added to the mix, while volumes of plasmid DNA and BAC remain unchanged
4. Pour approximately 20 ml (the minimum volume to cover the surface completely) of EM-grade water in the 15 cm Petri dish: this is called the hypophase. Cleave a mica sheet (about 1 cm  $\times$  2 cm) and place it in the water as a ramp, with the freshly cleaved surface facing up (Fig. 3c).
5. Spin the tube containing the sample for a few seconds in a microfuge and aspirate it completely in a 10  $\mu$ l tip. With a cotton swab sprinkle a few grains of graphite powder (no graphite flakes!) onto the water surface, in close proximity to the mica ramp (Fig. 3c).
6. Pipette the entire sample volume out of the tip and let the droplet touch the ramp few millimeters above the water surface. The BAC-containing drop will immediately slide down the ramp and spread over the water hypophase. The graphite powder will mark the border of the monomolecular detergent film containing the DNA molecules (Fig. 3d; *see Note 30*).
7. Using fine tweezers take one carbon-coated, EtBr-treated grid from the filter paper (carbon-side down, Fig. 3e) and pick up part of the DNA film touching the spreading surface in proximity to the graphite powder (*see Note 31*). Hold the tweezers to ensure full parallel contact of grid and surface (*see Note 10*). Enough pressure should be applied to ensure full contact between the carbon and the DNA-containing film (Fig. 3f). The grid is then removed from the surface and incubated for 15 s for staining in 1 ml of uranyl acetate working solution (in a flat-bottom 20 ml tube; Fig. 3g). After a brief wash (1–2 s) in 100% Ethanol, air-dry the grid (carbon-side up) on filter paper (Fig. 3h) and carefully wipe the tweezers with a napkin (*see Note 32*).
8. **Step 7** is repeated for the other grids, ensuring to collect DNA from different regions of the BAC film containing the DNA molecules (*see Note 33*).



### 3.9.2 Denaturing DNA Spreading by the “BAC Method”

To obtain information about in vivo-nucleosome positioning on the RIs (*see* Subheading 1), DNA samples can be denatured just before the BAC spreading. In this case, the presence of the denaturing agents (Formamide and Glyoxal) in the spreading mix is coupled with a short incubation at 42 °C, leading to DNA strand separation at each of the not-crosslinked regions (nucleosomal DNA; Figs. 1 and 8; *see* **Note 34**).

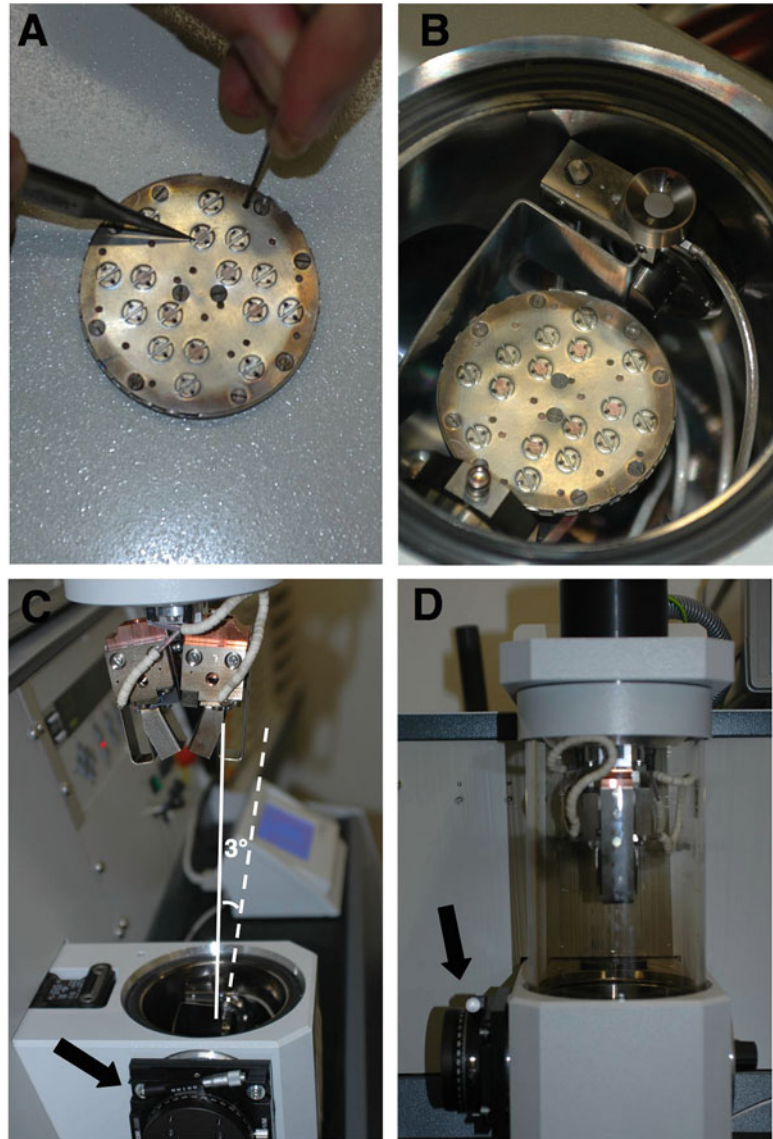
The spreading procedure is identical to the one described above (Subheading 3.9.1), with the exception of **steps 2** and **3**, performed as follows:

2. Using a P2 pipette, mix in a 1.5 ml-microfuge tube: 1.0 µl of Formamide, 0.2 µl of Glyoxal, and 1.0 µl of DNA sample (10–50 ng) from Subheading 3.7, **step 12**. Incubate for 10 min at 42 °C in a water bath and chill immediately after in ice-water.
3. Spin briefly the sample and add: 0.25 µl of plasmid DNA 5 ng/µl (internal size marker, *see* **Note 35**) and 0.4 µl of BAC 1:10. Immediately proceed with **steps 4–8** of the BAC spreading.

### 3.10 Platinum-Carbon Rotary Shadowing

1. Place the grids on the specimen table, taking care that they are properly fixed (by clips or magnetic stripes) and flat (Fig. 4a). Position the specimen table on the rocking rotary stage and the quartz sensor as close as possible to the specimen table (Fig. 4b).
2. Tilt the rocking rotary stage using the Precision Rotation Platform PR01, so that the angle between the specimen table and the Pt/C gun is exactly and reproducibly 3° (Fig. 4c). The fine micrometric scale on the Rotation Platform enables the reproducibility of the angle (Fig. 4c and d, black arrow). Cover the rotary stage with the tilting shutter, properly position the glass vacuum chamber wall, and start the turbomolecular pump of the MED 020.
3. At a vacuum of  $3 \times 10^{-5}$  mbar (or higher vacuum), preheat the filament of the Pt/C electron gun. Adjust voltage and current applied to the electron gun to reach a constant platinum/carbon evaporation rate of 0.03–0.05 nm/s (*see* **Note 36**), detected on the QSG 100 thin film monitor (*see* **Note 23**). Open the tilting shutter and start the measure of platinum film thickness on the QSG 100. When the detected thickness is 0.4 nm, start the rotation of the specimen table at the minimum speed (about 20 rpm). During the evaporation time (4–5 min) keep adjusting voltage and current applied to the electron gun to maintain the platinum/carbon evaporation rate constant. When the thickness readout at the QSG 100 is 8–10 nm (1800–2000 Hz), close the shutter, stop rotation, and shut off the electron gun. After the gun is cooled down, start the machine again and evaporate another 4–6 nm platinum on your grids while rotating (*see* **Note 37**).



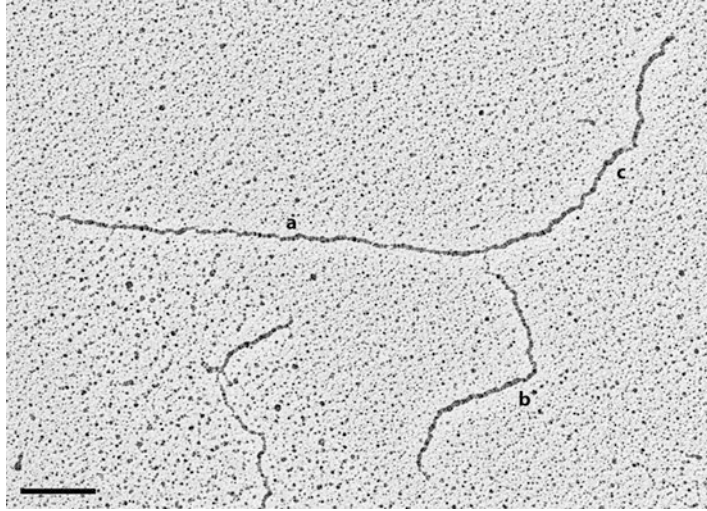


**Fig. 4** Series of photographs showing crucial steps in the platinum/carbon rotatory shadowing procedure (*see* Subheading 3.10 for details)

4. Remove the grids from the specimen table and store them properly for transportation. The grids can immediately be analyzed at the Transmission Electron Microscope or can be (re-)analyzed after unlimited storage periods.
1. The grids can be analyzed at any Transmission Electron Microscope. Duplex DNA in this technique is expected to appear as a 10 nm thick fiber, while ssDNA thickness should be 5–7 nm [20]. Both molecules should be clearly detectable on the homogenously granular background given by the platinum grains deposited on the carbon film (*see* Note 12).

**3.11 Visualization at the Transmission Electron Microscope, Contour Length Measures, and Statistics**

2. The plasmid DNA molecules added to the spreading mix can be used as an internal control of DNA absorption by the carbon film and as an internal size marker. Moreover, these circular molecules are often partially denatured (even in “native spreadings”) by the short incubation with Formamide and Glyoxal: their partial denaturation offers an easy opportunity to verify whether ssDNA and dsDNA in the same molecule are easily distinguishable by their thickness. At this step, if the concentration of the DNA, the quality of the shadowing, and/or the generally clean appearance of the grid background are not fully satisfactory, the experiment should be repeated, starting from the DNA spreading (or from earlier steps in the sample preparation, in case of recurrent problems, probably resulting from the DNA sample itself).
3. The extensive analysis of a satisfactory sample is usually performed at 5000–20,000 $\times$ . Despite the enrichment procedure described in Subheading 3.7, linear duplex DNA represents at least 90% of the genomic DNA visualized on the grids (it is retained on the QIAGEN-tip 20 probably because of local “breathing” of the DNA duplex and exposure of ssDNA). The identification of a replication fork requires the recognition of a “3 leg-junction,” i.e., a contact point from which three different DNA fibers depart (Fig. 5). These junctions need to be carefully analyzed at higher magnification (50,000–250,000 $\times$ ) to distinguish them from occasional overlaps of two linear molecules in proximity to one of the ends (“4-leg junctions”). At higher magnifications, some ssDNA is usually detectable at one or two sides of the junction, reinforcing the interpretation of the junction as a bona fide replication fork and supporting the identification of the daughter strands (Fig. 5). Furthermore, at least two of the three legs should be equal in length (as the genomic DNA has been digested, the elongation point should be equally distant from the two newly replicated restriction sites). The estimation of DNA molecule length is often complicated by the convoluted distribution of the fibers; this often requires the detailed analysis to be performed only later, once a digital file has been generated. Technical problems, such as partial restriction digestion or breakage of DNA molecules during sample preparation, account for the fraction of replication forks with all three legs different in length, i.e., asymmetric replicated duplexes. In standard conditions these should not represent more than 25–30% of the total replicating molecules, but the proportion of asymmetric forks can increase in conditions of DNA damage, associated with ssDNA accumulation.
4. Although photo-documentation can be obtained on any TEM generating photographic negatives (followed by developing



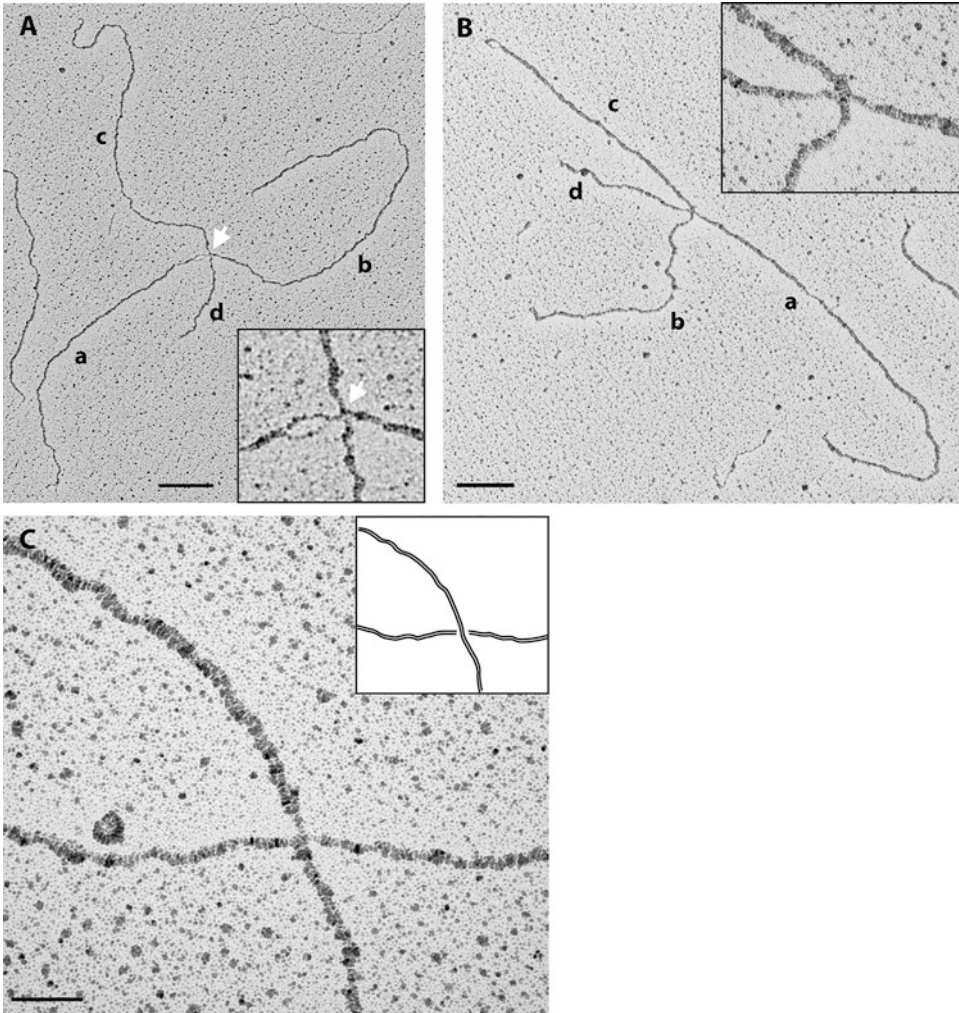
**Fig. 5** Normal replication fork visualized by in vivo psoralen crosslinking and TEM. Replicated duplexes (b, c) are of equal length. A ssDNA stretch is clearly visible in proximity to the fork on one of the two replicated duplexes (presumably the lagging strand)

and printing of the corresponding micrographs), the use of a microscope connected to a CCD camera is highly recommended: besides the reduction of the time required for photo-documentation, this procedure allows us to obtain digital files that can be promptly used for contour length measures through Imaging applications (*see* Subheading 3.11, step 6).

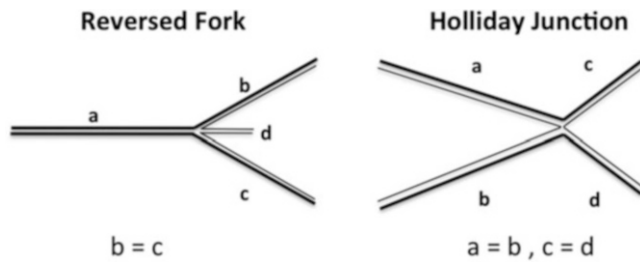
5. As a general rule, 70–100 replicating molecules should be collected in each experiment to obtain reliable data from the following statistical analysis. Observed trends need to be reproduced in two-three biological replicates. Should the observed differences in the frequency of specific intermediates be small and/or not fully reproducible, statistical assessment on data from three independent experiments is essential.
6. Once the digital files are converted in \*.tiff files, contour length measurements can be performed by standard image analysis applications. ImageJ has been successfully used, but other, more specific applications are being tested and could be more appropriate (and time-saving) for filament recognition and analysis. Once the measurements are performed, data can be analyzed by using standard statistical/graphical applications such as Microsoft Excel or Graphpad Prism.

### 3.12 Interpretation of Four-Way Junctions

Our most recent studies have identified reversed replication forks (Figs. 6a, b and 7) as frequent replication intermediates under certain conditions, i.e., upon genetic or pharmacological replication interference [11–17, 30]. Importantly, the regressed arm can either be connected to both or only one of the daughter duplexes,



**Fig. 6** Compilation of three molecules to assist the identification of reversed forks. (a) and (b) are representative examples of reversed forks, while (c) shows the accidental crossing of two independent DNA molecules (see Subheading 3.12 for details)



**Fig. 7** Drawing on reversed fork vs Holliday junction and their expected features in terms of contour length measurements



and it can be entirely or partially single-stranded. The unambiguous identification of these replication intermediates is based on contour length measurements in combination with several additional criteria. As reversed replication forks are four-way structures, accidental crossings of linear DNA molecules can easily be mistaken for reversed forks. The incidence of random crossings increases with the DNA concentration on the grids; therefore, it is necessary to work with samples with low DNA density to minimize the risk of misinterpretations.

The most important criterion for the assignment of a four-way structure as a reversed fork is the junction itself. If the junction is opened up and allows the clear identification of a rhomboid structure, the assignment as a reversed replication fork is usually straightforward (Fig. 6a). However, most reversed replication forks display a collapsed junction with little or no opening (Fig. 6b) and therefore require additional parameters for their interpretation. These parameters are the appearance of the junction, and the orientation and the length of the arms.

For a careful interpretation of a four-way junction, taking a high magnification image (135,000–250,000 $\times$ ) of the junction is recommended. All the arms of a reversed fork are connected to each other; therefore, they give the junction a “flat” appearance in one focal plane (Fig. 6b, inset). In contrast, in accidental crossings of linear DNA, one molecule is on the top of the other, resulting in a difference in focal plane that can help to discriminate DNA crossings from true DNA junctions already while taking pictures (Fig. 6c). Moreover, crossings are usually associated with a continuous shadowing of the molecule on the top and a discontinuous platinum deposition on the bottom molecule (Fig. 6c).

As in an unperturbed replication fork, the length of the two new replicated duplexes of a reversed fork should be identical ( $b = c$ ; Fig. 7). In contrast, as typical for all replication intermediates, the length of the parental arm ( $a$ ) is not defined, as is the length of the regressed ( $d$ ) arm ( $a \neq b = c \neq d$ ; Fig. 7). Conversely, Holliday junctions connect symmetrically two homologous fragments and will have by definition arms of equal length two by two ( $a = b, c = d$ ). Only in those rare cases in which a restriction site on the parental DNA is located in close proximity to the elongation point (late replication intermediate), the same restriction site will also be present on the regressed arm, giving rise to ( $a = d; b = c$ ) molecules. The appearance of such molecules would be identical to a late termination intermediate or to a Holliday Junction; therefore, they cannot be unambiguously assigned and should not be considered in the analysis.

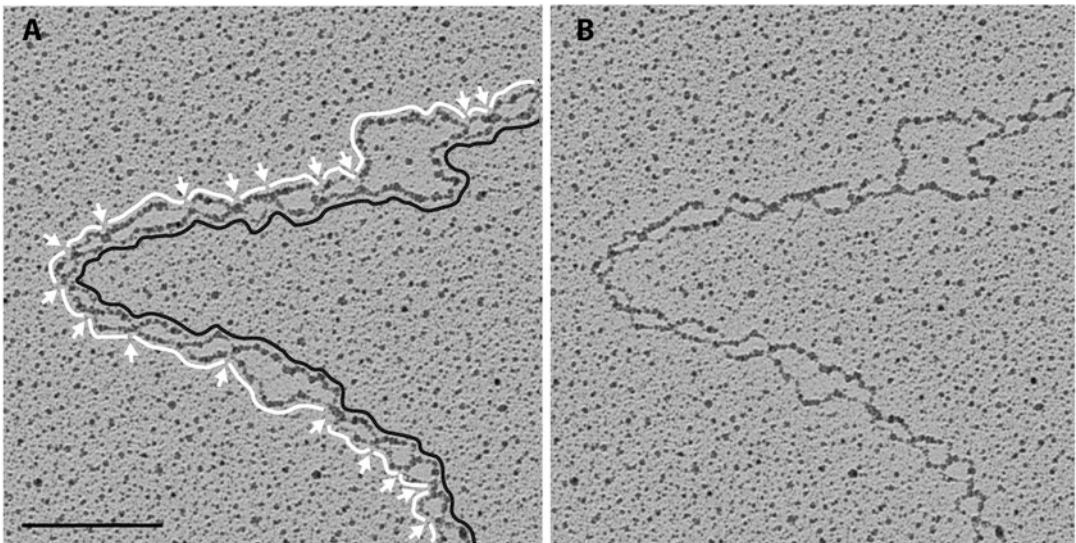
The final criterion for the identification of reversed forks is the orientation of the arms of the molecule. Reversed forks and other DNA junctions usually display different orientations of at least three of the four arms of the molecule, whereas the “arms” in

accidental crossings are often aligned along only two axes, sometimes even in a perpendicular fashion (compare Fig. 6b and c).

As for all other replication intermediates (*see* Subheading 3.11, **step 3**), partial digestion and DNA breakage can lead to reversed forks displaying no obvious symmetry in the length of the four arms. Particular attention should be paid to the analysis of the junction at high magnification (*see* above) to assign these molecules as reversed forks. As a general rule, the fraction of asymmetric reversed forks should never exceed the frequency of asymmetry observed in the same samples for the population of “normal” (three-way) replication forks.

### 3.13 Analysis of Replicating Chromatin Under Denaturing Conditions

Denaturing spreadings (Subheading 3.9.2) allow the analysis of nucleosome density on DNA. For this kind of analysis, a sufficient degree of psoralen crosslinking is required. 80–90% of the linkers should contain at least one crosslink, so that 80–90% of detected ssDNA bubbles (Figs. 1 and 8) represent mononucleosomes of 150–200 bp [2, 22–24]. Nucleosome density is usually expressed by the so-called  $r$ -value, which is calculated as the combined contour length of all nucleosome bubbles in a given stretch of DNA (Fig. 8a, white), divided by the overall contour length of the same DNA stretch (Fig. 8b, black). A reduced  $r$ -value indicates a reduction in nucleosome density.



**Fig. 8** Example of denaturing molecule (b) and system to measure it (a). The portion of the molecules organized in nucleosomes is calculated by adding all DNA stretches in the visible ssDNA bubbles (*white tracts*), thus excluding gaps of crosslinked DNA (*white arrows*). This combined length is then divided by the total length of the DNA tract analyzed (*black line*), giving rise to the “ $r$ -value” (*see* Subheading 3.13 for details)

---

## 4 Notes

1. A high efficiency of psoralen crosslinking can be obtained with shorter irradiation times, using monochromatic 365 nm Hg lamps [25]. However, with proper tuning of the irradiation time, this protocol has the advantage of using standard equipment (Biolinker) frequently available in molecular biology laboratories.
2. High quality genomic DNA for EM analysis has also been obtained reproducibly using QIAGEN Genomic-tip 20/G, following the manufacturer's instruction (vortexing steps are substituted by more gentle inversion of the tubes, to avoid shearing of genomic DNA). Nevertheless, the procedure described in detail is the one used for most of the EM data recently published on yeast replicating genomic DNA [5, 8, 31].
3. Help CTAB dissolution by heating the solutions to about 50 °C (do not overheat or boil!).
4. A number of different evaporators are available, with relevant differences in terms of vacuum chamber design and orientation of the electron guns in respect to the specimen. All evaporators can be used to produce carbon films and most of them can be adapted to perform extremely flat angle Pt/C rotary shadowing. The procedure described here allows us to perform Carbon evaporation (Subheading 3.8) and Pt/C rotary shadowing (Subheading 3.10) with the same machine, currently available on the market, with minor technical modifications to the apparatus (*see Note 12*). In our experience, the correct maintenance of the apparatus and the constant use/order of identical consumable parts (Wolfram cathodes, carbon rods, platinum insets) is of crucial importance to avoid unexpected problems during evaporation and to obtain reproducibly good carbon layers and rotary shadowing.
5. Supporting wire mesh stands, similar to the one shown in Fig. 2b, can be produced in any workshop. It is composed of two Teflon rings, the lower of which has holes that allow controlled water flow out of the central cavity. A wire mesh stand is placed in-between the two Teflon rings and can accommodate the round filter paper where the grids are placed. Teflon can be substituted by different materials (Plexiglas, plastic, etc.). However, Teflon offers the best resistance to acid washes of the apparatus, occasionally required to remove traces of the carbon films.
6. Excellent results have been obtained with re-distilled water. More recently, standard MilliQ water (produced by a conventional Millipore apparatus), with resistivity of about 18 MΩ cm and total organic content of less than eight parts per billion, has provided satisfactory results. This water is stored in a proper

flask solely dedicated to the EM work, which is directly filled at the MilliQ water apparatus.

7. Nickel grids can be used as an alternative to copper grids to make use of magnetic holders for the evaporation machine. This enables the use of for example the Leica BAF060 which speeds up the shadowing procedure since the machine has a permanent vacuum. The handling and procedures are identical but it is recommended to demagnetize the tweezers to prevent the nickel grids from sticking to them.
8. The original powder stock of BAC was kindly provided by Bayer, Leverkusen, Germany. The chemical is now available also from different suppliers (i.e., Sigma, cat. B6295), but has never been tested for this specific application.
9. Tissue culture dishes are preferred over standard Petri dishes, as their surface treatment allows covering the surface with smaller volumes. A reduced hypophase volume improves the stability of the detergent-DNA film during the spreading (*see also Note 32*).
10. Bent tweezers (Fig. 3c and f): to hold the grids as flat as possible while touching the spreading surface, it is of great help to slightly bend the tweezers tips; this can be done mechanically applying pressure on the tips with a proper tool. Particular care is necessary to avoid the breakage of the tips. Recently, we also obtained satisfactory results using Dumont pinzettes (*see Subheading 2.8*).
11. Better results have been obtained with grid tables that fix individual grids by small pins (Fig. 4a), compared with those currently available on the market, where multiple grids are fixed on the table by metal stripes. Besides the ease in fixing/removing individual grids, the former tables show reduced interference with low angle rotary shadowing, by maximizing the area of each grid effectively exposed to the correct amount of evaporated Pt. Leica is currently considering to restart production of these grid tables. Magnetic grid holders can also be efficiently used, if Nickel grids are used (*see Note 7*).
12. The accurate determination of the very flat angle between the electron gun and specimen table is crucial to observe an optimal difference in thickness between ssDNA and dsDNA. To facilitate the determination and the reproducibility of the optimal angle in independent experiments, the knob available on the MED 020 to tilt the rocking rotary stage can be substituted with a precision rotation platform, providing micrometric control of this angle (Fig. 4a and d, black arrow).
13. A standard EM experiment, consisting of 100 replicating molecules (by a standard 1k × 1k resolution CCD camera), corresponds to several Gigabytes of digital files!



14. For experiments in *S. cerevisiae*, it is recommended to extract genomic DNA from cells synchronized in S-phase, i.e., 30–180 min after release from an  $\alpha$ -factor arrest, depending on the treatment [5, 9]. While synchronization experiments have been performed also for EM analysis in mammalian cells—3–6 h after release from 16 h HU-block [32]—this does not seem to be strictly required. Although certainly more time consuming, it is possible to obtain 70–100 RIs from asynchronous, untreated U2OS cells or MEFs [15].
15. It is important to remove traces of growth media, which would otherwise absorb part of the monochromatic light in Subheadings 3.1, step 4 and 3.2, step 4.
16. Due to the increasing Ethanol concentration in the suspension, cells tend to aggregate and deposit to the bottom of the dish. This does not interfere with the following extraction of genomic DNA, but requires extensive resuspension of the sample prior to each irradiation.
17. Timing of irradiation and number of crosslinking cycles strictly depend on the irradiance of the 365 nm bulbs. The observed irradiance within the crosslinker, measured by UV meter, is  $6.2 \text{ mW/cm}^2$ . This needs to be monitored, as it changes considerably with the type of apparatus, the age of the bulbs, and the distance of the samples from the bulbs. The total irradiation time should be adjusted to apply the same total irradiation ( $\text{J/cm}^2$ ) to the samples as indicated here. Only doing so, it is possible to obtain optimal frequency of crosslinks on genomic DNA (1 crosslink in 80–90% of DNA linkers), particularly important for chromatin studies (*see* Subheading 3.13).
18. Incorporation of Cy3-dCTP can be monitored by visualization of nuclei using a standard fluorescence microscope. We consider an appropriate time for nuclear replication when 80–90% of the nuclei appear positive for Cy3 incorporation.
19. We have tested the efficiency of this RI enrichment protocol and compared it to standard BND cellulose-mediated enrichment [21], by performing a new round of digestion, RI enrichment, and EM analysis of several genomic DNA samples that had already been analyzed by EM [17], and for which sufficient amount of residual genomic DNA was still available in our stocks. For all the tested samples, we obtained practically identical results—in terms of ssDNA accumulation at forks and reversed fork frequency—to those previously obtained processing the same samples with the standard BND-cellulose enrichment (Fig. 6a in [17]). However, we reproducibly noticed that—despite similar density of total DNA on our EM grids—the identification of 70–100 RIs was significantly

(two- to threefold) faster than what typically observed with BND-enriched DNA. These observations strongly suggest that this optimized enrichment protocol allows “discarding” a higher fraction of linear, non-replicating DNA during the first elution step, yielding a higher enrichment of RI in the final elution, without introducing any bias on the type of replication structures and on ssDNA accumulation finally detected in the sample.

20. While choosing the restriction enzyme to digest genomic DNA, it is important to consider that: (a) frequent cutters will lead to smaller RIs (possibly complicating the analysis of phenomena happening at longer distance from the replication forks); (b) due to frequent methylation of CpG sequences, mammalian genomic DNA may be resistant to the action of a number of common restriction enzymes. It is also recommended to use enzymes releasing blunt DNA ends, to prevent ssDNA overhangs competition for binding to the BND cellulose.
21. It is important that residual RNA is removed prior to loading the sample on the QIAGEN-tip 20, to avoid interference with the enrichment procedure. This is particularly important for the extraction procedure described in Subheading 3.4, which normally leads to the extraction of large amounts of dsRNA, besides the expected genomic DNA. When residual RNA is detected on the gel, add to the restriction mix a few units of RNaseA (for ssRNA; 5  $\mu$ l of 10 mg/ml in 250  $\mu$ l reaction) and RNaseIII (for dsRNA; 1  $\mu$ l of 1:100 dilution in TE in a 250  $\mu$ l reaction; do not digest longer than 3 h to prevent DNA degradation). They are both usually active in the common restriction buffers.
22. The DNA recovered in Subheading 3.7, **step 12** is normally 1–15% of the total DNA loaded on the QIAGEN-tip 20 column, depending on the fraction of cells actively replicating their DNA. If the amount recovered at this step is far from this proportion, proper aliquots of the flow-through from Subheadings 3.7, **step 6** and 3.7, **step 7** can be loaded on a gel to figure out where the DNA got lost, recover it and repeat the procedure with appropriate modifications. The accurate quantification of the DNA recovered in Subheading 3.7, **step 12** requires an agarose gel, as the caffeine “contamination” in the RIs-enriched fraction interferes with standard spectrophotometric or fluorimetric measurements of DNA concentration.
23. The readout is calculated correcting for the different position of the quartz sensor in respect to the specimen (tooling factor). With the quartz crystal placed next to the sample (*see* Fig. 4b) the tooling factor in the QSG 100 should be set to 1. Different

settings will require the empirical determination of the correct tooling factor.

24. After two-to-three evaporation sessions, the electron gun overheats. Because of the high temperature, carbon films tend to strongly adhere to the mica sheet and are resistant to the floating procedure in Subheading 3.8, **step 7**.
25. Do not proceed lowering the mica in the water unless the carbon film is visibly detaching from the mica sheet. The proper coordination of mica sheet submersion and carbon film detachment is crucial for carbon film integrity and requires some manual skills and experience.
26. Generally, for each DNA spreading, grids covered with two-to-three different carbon films are used. As each carbon film has different absorption properties, this increases the chances of obtaining at least some grids with an optimal concentration of absorbed DNA molecules (*see Note 31*).
27. Formamide acts as a “partially-denaturing” agent, helping to disentangle and unfold DNA molecules during the spreading procedures. These conditions are optimal for psoralen-crosslinked DNA, which is inherently resistant to denaturation; if uncrosslinked DNA needs to be used, formamide concentration can be reduced to prevent DNA denaturation. If the DNA is only briefly exposed to formamide during the procedure (*see Note 30*), the vast majority of the psoralen-crosslinked DNA molecules is still visible as dsDNA, although some “breathing” is occasionally detected along the duplex and especially at DNA ends. Full denaturation of the sample requires longer exposure to these agents at higher temperatures (*see Subheading 3.9.2*).
28. The mix of the two droplets (by spinning briefly, *see Subheading 3.9.1, step 5*) should only be performed immediately before the spreading and having all necessary material ready, to minimize DNA exposure to formamide (*see Note 29*).
29. The final concentration of the DNA on the grid is crucial for a proper analysis of RIs. This depends on the DNA concentration in the sample, but also on the carbon absorption properties and the size of the spreading surface. Therefore, the spreading procedure often needs to be repeated to obtain grids with an optimal distribution of DNA molecules. For DNA samples that are particularly diluted, the DNA volume can be increased up to 5  $\mu\text{l}$ , provided that formamide/glyoxal volumes are changed accordingly. BAC/plasmid volumes are unchanged, so that the size of the spreading surface and the density of marker molecules are not changed, leading to a higher frequency of DNA molecules of interest on the EM grids.

30. The stability of the monolayer can be heavily affected by any local perturbation of the conditions (hitting the bench, colleagues walking by, etc.). It is therefore highly recommended to perform the BAC spreading procedure in a quiet laboratory, on an (at least transiently) isolated bench. It is helpful to protect the working area with a Plexiglas box, open only to the side facing the operator.
31. The DNA molecules accumulate in close proximity to the edge of the spreading surface (close to the graphite powder).
32. Ethanol traces on the tweezers heavily disturb the spreading surface!
33. It is unavoidable that the monolayer surface will shrink and be perturbed while repeatedly touching it with the grids. The concentration of the DNA can also vary accordingly. Nonetheless, if the spreading surface is clearly unstable from the beginning, it is recommended to repeat the spreading procedure (*see also* **Notes 9** and **32**, on how to improve the spreading surface stability).
34. Denatured DNA is more difficult to absorb to the carbon films. It may happen that carbon-coated grids that proved proficient in absorbing dsDNA do not perform equally well absorbing denatured (mostly single-stranded) DNA (*see* **Note 37**). Due to the lower thickness, denatured DNA is also more difficult to visualize at the EM and usually requires optimal contrast both from uranyl acetate staining and from Pt/C rotary shadowing.
35. It is important that the plasmid marker is added only after denaturation. Preserving its duplex status, the marker will serve as an absorption control for the carbon films and will help assessing the quality of the contrast obtained on the grids (*see* Subheading **3.11**, **step 2**).
36. Standard settings for platinum shadowing of a MED020 evaporation machine are 1.70 kV, 0.40 mA and tooling factor 1.0 and for carbon evaporation 1.70 kv, 0.70 mA and tooling factor 0.5. These values can differ from machine to machine but can be used as a starting point to determine your machine's settings.
37. A second round of platinum shadowing provides increase contrast to DNA molecules on carbon-coated grids. It is essential to wait that the Pt/C gun cools down, before the second round of platinum coating is performed.

---

## Acknowledgments

We are grateful to José M. Sogo for his crucial support while learning all technicalities of this approach. We wish to thank the whole team at

the ZMB (Center for Microscopy and Image Analysis of the University Zurich) for consistently excellent technical assistance, while running our EM experiments. We are grateful to Arnab Ray Chaudhuri, Yoshitami Hashimoto, Fabio Puddu, and Vincenzo Costanzo for their assistance in optimizing this EM approach on *Xenopus* egg extracts. We also wish to thank Petr Cejka for suggesting the use of QIAGEN columns as possible valuable alternative to BND cellulose for the enrichment of a subpopulation of DNA molecules based on differential ssDNA content. We are also grateful to Sebastian Ursich for his recent efforts optimizing this approach and for careful reading of the manuscript. Work in the Lopes lab is currently financed by the SNF grant 31003A\_169959, the ERC Consolidator grant 617102 (ReStreCa) and the Swiss Cancer League grant KFS-3967-08-2016.

## References

- Inciarte MR, Salas M, Sogo JM (1980) Structure of replicating DNA molecules of *Bacillus subtilis* bacteriophage phi 29. *J Virol* 34:187–199
- Lucchini R, Sogo JM (1995) Replication of transcriptionally active chromatin. *Nature* 374:276–280
- Sogo JM, Stahl H, Koller T, Knippers R (1986) Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. *J Mol Biol* 189:189–204
- Avemann K, Knippers R, Koller T, Sogo JM (1988) Camptothecin, a specific inhibitor of type I DNA topoisomerase, induces DNA breakage at replication forks. *Mol Cell Biol* 8:3026–3034
- Lopes M, Foiani M, Sogo JM (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol Cell* 21:15–27
- Engels K, Giannattasio M, Muzi-Falconi M et al (2011) 14-3-3 proteins regulate exonuclease I-dependent processing of stalled replication forks. *PLoS Genet* 7:e1001367. doi:10.1371/journal.pgen.1001367
- Giannattasio M, Follonier C, Tourriere H et al (2010) Exo1 competes with repair synthesis, converts NER intermediates to long ssDNA gaps, and promotes checkpoint activation. *Mol Cell* 40:50–62. doi:10.1016/j.molcel.2010.09.004
- Hashimoto Y, Chaudhuri AR, Lopes M, Costanzo V (2010) Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. *Nat Struct Mol Biol* 17:1305–1311. doi:10.1038/nsmb.1927
- Sogo J, Lopes M, Foiani M (2002) Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science* (New York, NY) 297:599–602
- Berti M, Chaudhuri AR, Thangavel S et al (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat Struct Mol Biol* 20:347–354. doi:10.1038/nsmb.2501
- Follonier C, Oehler J, Herrador R, Lopes M (2013) Friedreich's ataxia-associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nat Struct Mol Biol*. doi:10.1038/nsmb.2520
- Neelsen KJ, Zanini IMY, Mijic S et al (2013) Deregulated origin licensing leads to chromosomal breaks by rereplication of a gapped DNA template. *Genes Dev* 27:2537–2542. doi:10.1101/gad.226373.113
- Neelsen KJ, Zanini IMY, Herrador R, Lopes M (2013) Oncogenes induce genotoxic stress by mitotic processing of unusual replication intermediates. *J Cell Biol* 200:699–708. doi:10.1128/MCB.24.16.7140-7150.2004
- Neelsen KJ, Lopes M (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nat Rev Mol Cell Biol* 16:207–220. doi:10.1038/nrm3935
- Ray Chaudhuri A, Hashimoto Y, Herrador R et al (2012) Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol* 19:417–423. doi:10.1038/nsmb.2258
- Ray Chaudhuri A, Ahuja AK, Herrador R, Lopes M (2015) Poly(ADP-ribose) glycohydrolase prevents the accumulation of unusual replication structures during unperturbed S

- phase. *Mol Cell Biol* 35:856–865. doi:[10.1128/MCB.01077-14](https://doi.org/10.1128/MCB.01077-14)
17. Zellweger R, Dalcher D, Mutreja K et al (2015) Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *J Cell Biol* 208:563–579. doi:[10.1083/jcb.201406099](https://doi.org/10.1083/jcb.201406099)
  18. Thangavel S, Berti M, Levikova M et al (2015) DNA2 drives processing and restart of reversed replication forks in human cells. *J Cell Biol* 208:545–562. doi:[10.1083/jcb.201406100](https://doi.org/10.1083/jcb.201406100)
  19. Fugger K, Mistrik M, Neelsen KJ et al (2015) FBH1 catalyzes regression of stalled replication forks. *Cell Rep*. doi:[10.1016/j.celrep.2015.02.028](https://doi.org/10.1016/j.celrep.2015.02.028)
  20. Lopes M (2009) Electron microscopy methods for studying in vivo DNA replication intermediates. *Methods Mol Biol* 521:605–631
  21. Neelsen KJ, Chaudhuri AR, Follonier C et al (2014) Visualization and interpretation of eukaryotic DNA replication intermediates in vivo by electron microscopy. *Methods Mol Biol* 1094:177–208. doi:[10.1007/978-1-62703-706-8\\_15](https://doi.org/10.1007/978-1-62703-706-8_15)
  22. Gasser R, Koller T, Sogo JM (1996) The stability of nucleosomes at the replication fork. *J Mol Biol* 258:224–239
  23. Gruss C, Wu J, Koller T, Sogo JM (1993) Disruption of the nucleosomes at the replication fork. *EMBO J* 12:4533–4545
  24. Mejlvang J, Feng Y, Alabert C et al (2014) New histone supply regulates replication fork speed and PCNA unloading. *J Cell Biol* 204:29–43. doi:[10.1083/jcb.201305017](https://doi.org/10.1083/jcb.201305017)
  25. Wellinger RE, Lucchini R, Dammann R, Sogo JM (1999) In vivo mapping of nucleosomes using psoralen-DNA crosslinking and primer extension. *Methods Mol Biol* 119:161–173
  26. Vollenweider HJ, Sogo JM, Koller T (1975) A routine method for protein-free spreading of double- and single-stranded nucleic acid molecules. *Proc Natl Acad Sci U S A* 72:83–87
  27. Sogo JM, Thoma F (1989) Electron microscopy of chromatin. *Methods Enzymol* 170:142–165
  28. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning, a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press., Cold Spring Harbor, NY
  29. Trenz K, Errico A, Costanzo V (2008) Ptx1 is required for chromosomal DNA replication under stressful conditions. *EMBO J* 27:876–885. doi:[10.1038/emboj.2008.29](https://doi.org/10.1038/emboj.2008.29)
  30. Ahuja AK, Jodkowska K, Teloni F et al (2016) A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nat Commun* 7:10660. doi:[10.1038/ncomms10660](https://doi.org/10.1038/ncomms10660)
  31. Sogo JM (2002) Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science* 297:599–602. doi:[10.1126/science.1074023](https://doi.org/10.1126/science.1074023)
  32. Mojas N, Lopes M, Jiricny J (2007) Mismatch repair-dependent processing of methylation damage gives rise to persistent single-stranded gaps in newly replicated DNA. *Genes Dev* 21:3342–3355

## A Molecular Toolbox to Engineer Site-Specific DNA Replication Perturbation

Nicolai B. Larsen, Ian D. Hickson, and Hocine W. Mankouri

### Abstract

Site-specific arrest of DNA replication is a useful tool for analyzing cellular responses to DNA replication perturbation. The *E. coli* Tus-*Ter* replication barrier can be reconstituted in eukaryotic cells as a system to engineer an unscheduled collision between a replication fork and an “alien” impediment to DNA replication. To further develop this system as a versatile tool, we describe a set of reagents and a detailed protocol that can be used to engineer Tus-*Ter* barriers into any locus in the budding yeast genome. Because the Tus-*Ter* complex is a bipartite system with intrinsic DNA replication-blocking activity, the reagents and protocols developed and validated in yeast could also be optimized to engineer site-specific replication fork barriers into other eukaryotic cell types.

**Key words** Replication fork barrier, DNA replication stress, Tus-*Ter*

---

### 1 Introduction

The *E. coli* Tus-*Ter* barrier consists of the high-affinity binding of the DNA replication-blocking protein, Tus, to a specific 23-bp “*Ter*” sequence. The Tus-*Ter* barrier is polar in *E. coli*, and DNA replication forks are arrested only when encountering the “restrictive” face of the Tus-*Ter* complex [1]. We demonstrated previously that short arrays of either 3×Tus-*Ter* or 7×Tus-*Ter* barriers can function as a polar replication pausing system when engineered into the budding yeast genome [2, 3]. Our recent data indicate that the strength of replication fork pausing within a defined restriction fragment is saturated at 14×Tus-*Ter* barriers. However, because yeast can apparently overcome replication fork pausing at Tus-*Ter* barriers [3], further increasing the number of *Ter* sites, or engineering additional tandem >14×Tus-*Ter* barriers in *cis*, may be used to additively delay DNA replication at specific chromosomal loci. Here, we describe a detailed method for how to create and analyze budding yeast strains with inducible Tus-*Ter* barriers. Although these protocols were developed for engineering Tus-*Ter* barriers



into the yeast genome, the core reagents could easily be adapted for use in other cell types by utilization of other host-cell specific selectable markers. Indeed, the Tus-*Ter* system has also been established and independently validated in both mouse and human cell lines by other groups [2, 4, 5], suggesting that the Tus-*Ter* complex can be used as a versatile tool to engineer site-specific DNA replication perturbation in any cell type. Combined with the recent advances in genome engineering using CRISPR-Cas9 and short-homology targeting sequences [6], this could permit detailed analyses of DNA replication perturbation at specific regions of the genome in eukaryotic cells in a manner analogous to those developed in budding yeast and fission yeast [7–9].

The plasmids for engineering Tus-*Ter* barriers into the yeast genome contain  $14 \times Ter$  or  $21 \times Ter$  sites, in either Leftwards ( $L \rightarrow R$ ) or Rightwards ( $L \leftarrow R$ ) fork-blocking orientation, and are flanked by either the *LEU2* or *URA3* selectable markers on either side. Because all 16 of these *Ter* vectors harbor common primer sequences flanking the *Ter* modules, a number of configurations of *Ter* sites and selectable marker(s) can simultaneously be amplified using the same set of locus-specific primers. These *Ter* modules can then be integrated into the yeast genome by homologous recombination, and the validated strains transformed with a plasmid that allows the controlled expression of HA-tagged Tus (or a non-blocking allele of Tus as a control [2]). The choice of two different selectable markers (*LEU2* or *URA3*) in the *Ter* vectors allows two discrete barriers to be engineered/maintained in a given yeast strain. These Tus-*Ter* barriers can either be constructed in *trans* to allow in vivo comparisons between different locations/chromosomes, or else constructed in *cis* to additively disrupt DNA replication on the same chromosome. Additional combinations/configurations of Tus-*Ter* barriers are also theoretically possible through marker switching or recycling, if required. Furthermore, the *URA3* marker can also be used as a counter-selectable reporter to assay for replication-fork-arrest-induced mutagenesis at Tus-*Ter* barriers (by positive selection for *ura3* mutations on plates containing 5-FOA [10]). The positioning of the *URA3* reporter on either side of the Tus-*Ter* barrier permits the two regions either side of a Tus-*Ter* barrier to be probed for mutagenic outcomes. For example, an origin proximal *URA3* configuration permits positive selection for mutagenic outcomes either directly at, or behind, stalled replication forks. Conversely, an origin-distal *URA3* configuration permits positive selection for mutagenic events associated with fork resumption, or fork merging within this reporter gene.

Combined with the well-established methodologies to study DNA replication in yeast [11], this system serves as a “molecular toolbox” that can be used to create, and then monitor the phenotypic consequences of various challenges to the DNA replication machinery. This will permit the detailed analysis of localized



responses to DNA replication perturbation at discrete regions of the genome (e.g., early/late replicating regions, fragile sites, sub-telomeric loci, etc.), as well as analyzing the fate of unreplicated regions of the genome that persist beyond S-phase [12].

## 2 Materials

### 2.1 DNA Amplification for Construction of Yeast Strains with Inducible Tus-Ter Barriers

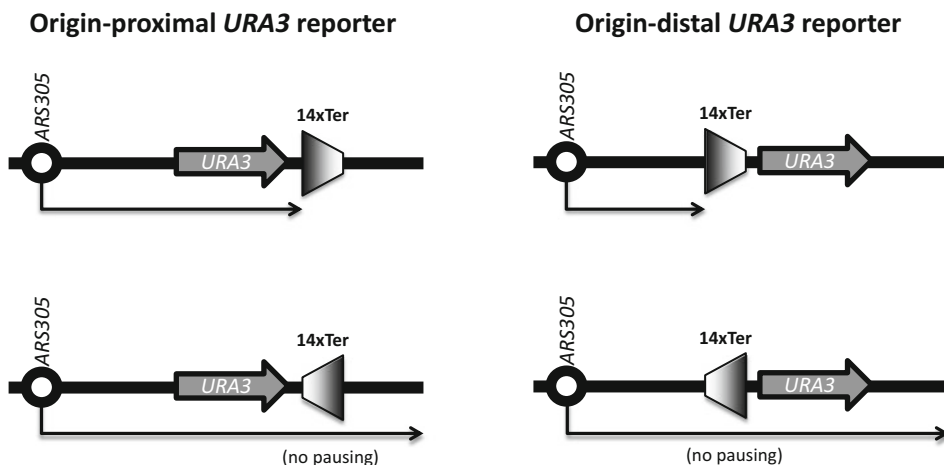
1. High fidelity PCR polymerase (e.g., PrimeSTAR HS DNA polymerase from Takara-Clontech).
2. Locus-specific targeting primers (*see Note 1* and Subheading 3.1).
3. *Ter* plasmid to use as a PCR template (*see Note 2* and Fig. 1).

### 2.2 Yeast Transformation

1. YEP Broth + 2% Glucose.
2. LiAc-PEG<sub>3350</sub> mix (Combine 1.5 ml of 1 M LiAc with 10 ml 50% PEG).
3. Carrier DNA (2 mg/ml herring sperm DNA).
4. CSM-Ura plates.

### 2.3 Validation of Successful *Ter* Module Integrants

1. 200 mM LiAc, 1% SDS solution.
2. 96–99% Ethanol.
3. Frozen-EZ Yeast Transformation II kit from Zymo Research.



**Fig. 1** Diagram of the Tus-*Ter* barrier constructs engineered adjacent to the early-firing origin, *ARS305*, on ChrIII. The four configurations are shown: permissive, or restrictive, *Ter* sites, with an origin-proximal, or origin-distal, *URA3* marker. A number of other configurations are possible if using different *Ter* vectors as PCR templates. A 21× *Ter* module can be amplified instead of 14× *Ter*, or a *LEU2* marker can be amplified instead of *URA3*. In total, 16× [*Ter* + marker] configurations are possible

**2.4 Synchronization of Cells in the G<sub>1</sub> Phase and Induction of Tus**

1. YEP Broth + 3% Lactate (autoclaved).
2. Alpha-factor mating pheromone (Synthetic peptide of the following sequence: WHWLQLKPGQPMY, >90% purity).
3. Galactose.

**2.5 Flow Cytometric Analysis**

1. RNase A (10 mg/ml).
2. Proteinase K (20 mg/ml).
3. Propidium Iodide solution (1 mg/ml).

**2.6 DNA Extraction for Two-Dimensional Gel Electrophoresis**

1. 10% Sodium azide solution (keep at 4 °C).
2. Spheroplast buffer (1 M sorbitol, 100 mM EDTA pH 8, 0.1% β-mercaptoethanol, 100 U/ml zymolyase (AMSBIO, UK)).
3. Solution I (2% w/v CTAB, 1.4 M NaCl, 100 mM Tris-HCl pH 7.6, 25 mM EDTA pH 8.0).
4. Solution II (1% CTAB, 50 mM Tris-HCl pH 7.6, 10 mM EDTA).
5. Solution III (1.4 M NaCl, 10 mM Tris-HCl pH 7.6, 1 mM EDTA).
6. Chloroform/isoamylalcohol 24:1.
7. KIMBLE HS Borosilicate glass centrifuge tubes.
8. Isopropanol.
9. 99% ethanol.
10. 10 mM Tris-HCl pH 8.0.

**2.7 Two-Dimensional gel Electrophoresis (2DGE) of DNA Extracts**

1. 2.5 M potassium acetate, pH 6.
2. TE buffer.
3. Low EEO agarose (e.g., USBiological).
4. Owl A2 Large Gel System (ThermoFisher Scientific).
5. 2DGE Loading Dye (0.83% Bromophenol blue, 50% glycerol).
6. 10 mg/ml Ethidium bromide.
7. TBE buffer.
8. Depurination solution (8 ml of 37% HCL per 1 L water).
9. Denaturing solution (0.5 M NaOH, 1.5 M NaCl).
10. Neutralization solution (1 M Ammonium acetate, 0.02 M NaOH).
11. 10× SSC.
12. Genescreen Hybridization Transfer Membranes (Perkin Elmer, USA).
13. UVC crosslinker.

## 2.8 Hybridization of Membranes

1. Hybridization solution (6× SSC, 1× Denhardt's, 1% sarcosyl; For Pre-Hybridization solution, add BSA to 0.1% final concentration).
2. DNA probe.
3. Rediprime II DNA labeling system (GE Healthcare Amersham).
4. <sup>32</sup>P dCTP (6000 Ci/mmol).
5. Quick Spin Column (Roche).
6. Wash Solution I (2× SSC, 1% SDS; incubate at 65 °C).
7. Wash Solution II (0.1× SSC, 0.1% SDS; incubate at 37 °C).

---

## 3 Methods

The protocol detailed below describes how to engineer site-specific replication fork stalling at any desired location in the yeast genome. To date, we have used this protocol to integrate *Ter* modules into five discrete locations, and subsequently observed DNA replication fork pausing when Tus is induced. For the purposes of brevity, we describe how to integrate and validate DNA replication fork pausing at a number of different modules inserted adjacent (to the right) of *ARS305* on ChrIII: 14× *Ter-URA3* and *URA3-14×Ter* (Fig. 1). In this scenario, the L → R barrier will be the “restrictive” (blocking) orientation, whereas the R ← L barrier will serve as the “permissive” (non-blocking) control, relative to the *ARS305*-borne forks. Note that the restrictive/permissive configurations of Tus-*Ter* barriers will always depend on the activity of local DNA replication origins, which determine the timing and directionality of DNA replication in the region of interest. Relative origin usage and efficiency have been extensively characterized in budding yeast [13, 14], and these studies can therefore be used as the basis for strategic engineering of Tus-*Ter* barriers.

### 3.1 DNA Amplification for Construction of Yeast Strains with Inducible Tus-Ter Barriers

1. All 16 *Ter* plasmids (*see Note 2*) have the same common flanking sequences that can be used as “TMA” (*Ter* module amplification) PCR primers:

TMA-1 Fwd primer: cgactcactatagggcggaattgg.

TMA-2 Rev. primer: gccgctctagaactagtggatc.

To target the region to the right of *ARS305*, 65-bp homology arms were included (indicated in upper case; *see Note 1*) to give the final primer sequences as:

305-TMA-1 Fwd:

GAGCAAGACAAACAGGGCCAGCTGAT

GCATATGTTTTGTGTTGCTTTCCTACGATCAGCTAA

TGCcgactcactatagggcggaattgg.

305-TMA-2 Rev.:

AAATGAGTTTTGTCCCACCTTCCCTTTGG  
GAAAAGGCAATGTAAATCTTAGAGGCAAGAACCAC  
Agccgctctagaactagtgatc.

2. The [*Ter* + marker] modules should be amplified using a high-fidelity PCR enzyme (using ~10–20 pg of *Ter* vector as a PCR template), at a high annealing temperature. Typically, we use PrimeSTAR HS DNA polymerase from Takara-Clontech, with the following PCR parameters:
  - (a) 98 °C for 90 s.
  - (b) [98 °C for 10 s, 60 °C for 15 s, 72 °C for 1 min/kb] × 35 cycles.
  - (c) 72 °C for 10 min.
3. In the example shown above, the entire *Ter* cassette, selectable marker, and homology arms are amplified as a single PCR product. However, it is also possible to use “split-*URA3* PCR” to generate two 428-bp overlapping PCR products that can only create a functional *URA3* gene by recombining into the desired locus. If *Ter* sequences flank the split-*URA3* PCR products on either side, then co-transformation of these (see Subheading 3.2 below) permits up to 42 × Tus-*Ter* barriers (i.e.,  $21 \times \text{Ter-}URA3^{(URA3 \text{ promoter} + 1-533)} + (105-804 + URA3 \text{ terminator})$  *URA3-21 × Ter* =  $21 \times \text{Ter-}URA3-21 \times \text{Ter}$ ) to be engineered within a ~3 kb fragment in a single round of yeast transformation.

The sequences of the “split-*URA3*” primers are as follows:

Split-*URA3*-Fwd: (to be paired with the locus-specific TMA-2 Rev. primer): TTGGATGTTTCGTACCACCAAGGAAT.

Split-*URA3*-Rev.: (to be paired with the locus-specific TMA-1 Fwd primer): GAGCAATAAAGCCGATAACAAAATCTT TGTCG.

### 3.2 Yeast Transformation

The transformation procedure is adapted from Gietz and Schiestl [15] and has been optimized for use in BY4741 strains. This protocol can also be used for making gene disruptants (see Note 3).

1. Grow cells overnight in YPD.
2. The next day, dilute cells ~1:40 in YPD and grow for 5–6 h at 30 °C/200 rpm. Each transformation reaction requires 15 ml of exponentially growing cells (OD<sub>600</sub> reading of 1.0–1.2).
3. Boil 2 mg/ml carrier DNA aliquots for 5 min, and then keep them on ice.
4. Prepare labeled tubes with 900 µl of LiAc-PEG mix.
5. Harvest cells from 15 ml aliquots of cultures by centrifuging at 900 × *g* for 5 min. Resuspend cells in 1 ml water, transfer the

mixture to 1.5 ml microfuge tubes, and briefly centrifuge. Remove the water to leave a cell pellet.

6. To each cell pellet, add 150  $\mu$ l of carrier DNA and (up to) 15  $\mu$ l of deletion cassette (or approximately 1  $\mu$ g of PCR products). Mix by pipetting and transfer the entire mixture to the Falcon tubes containing 900  $\mu$ l LiAc-PEG mix.
7. Incubate the transformation reactions at 30 °C/200 rpm for 10 min.
8. Transfer to 42 °C/180 rpm for 10–12 min in a water bath shaker.
9. Transfer the tubes to an ice-water bath to instantly chill them.
10. Add 3 ml of appropriate prechilled selection medium to the tubes, gently mix, and then centrifuge at 1000 rpm for 1 min.
11. After pouring off the supernatant, gently resuspend the cells in the ~300  $\mu$ l of residual media. Plate cells onto CSM–Ura plates (*see* **Note 3**).
12. After 3 days, Ura<sup>+</sup> colonies should be detectable. Restreak single colonies onto fresh CSM–Ura plates.

### **3.3 Validation of Successful *Ter* Module Integrants**

This DNA extraction procedure [16] can also be used to verify successful integration of *Ter* modules (or gene disruptions) by PCR using primers that flank the targeting regions. For those strains that give correct band sizes indicative of successful integration, it is recommended to confirm the DNA sequence is correct. In our experience, the success rate (i.e., no sequence alterations) is usually >66%.

1. Harvest cells from 200 to 300  $\mu$ l of stationary phase cultures, and resuspend in 300  $\mu$ l of 200 mM LiAc, 1% SDS solution.
2. Incubate for 10 min at 70 °C/600 rpm.
3. Add 1 ml of 99% ethanol, and then vortex the samples.
4. Centrifuge the DNA and cell debris at full speed for 5 min.
5. Wash the pellet with 70% ethanol.
6. Dissolve the pellet in 200  $\mu$ l of water, and incubate samples for a few hours at 37 °C to facilitate resuspension of DNA.
7. Clarify the samples by centrifugation and use 1–2  $\mu$ l of the supernatant as a PCR template.

Successfully transformed (and sequence validated) strains can then be transformed with a plasmid that allows the controlled expression of plasmid-borne HA-tagged Tus (or a non-blocking control) from the *GALI*-promoter [2]. To transform yeast with plasmids, we routinely use the Frozen-EZ Yeast Transformation II kit from Zymo Research, and follow the manufacturer's instructions.

### 3.4 Synchronization of Cells in G1 and Induction of Tus

This protocol has been optimized for BY4741 strains. If using another strain background, or if a particular deletion mutant is slow growing, some optimization of this protocol may be required.

1. Grow cultures in 4 ml of selective medium overnight.
2. The next day, add 1 ml yeast per 25 ml medium. OD<sub>600</sub> readings should be ~0.35. Incubate cultures at 200 rpm, 30 °C for ~7 h until OD<sub>600</sub> reading reaches ~1.30.
3. Set up overnight cultures in baffled flasks. The culture should comprise no more than 1/4th of the total volume. Add cells to 0.04 OD units/ml. Grow cultures overnight at 200 rpm, 30 °C.
4. After ~17–18 h, the OD<sub>600</sub> value should be around ~1.5–1.6. Add 5 µg/ml alpha factor (*see Note 4*), and galactose to 2% final concentration.
5. Incubate yeast cultures for 2.5 h at 200 rpm, 30 °C.
6. After arrest/induction, harvest cells by centrifugation (2000 × *g* for 3 min).
7. Wash cells twice in 100 ml YEP + 3% lactate + 2% Gal (2000 × *g* for 3 min), and resuspend the cells in double the starting volume (in **step 4**) of YEP + 3% lactate + 2% Gal.
8. Incubate cultures at 30 °C, 200 rpm. Harvest cells at fixed time points for flow cytometry and 2DGE analysis.

### 3.5 Flow Cytometric Analysis

Bulk DNA replication by flow cytometry is detectable at approximately 20 min after release from alpha-factor arrest, and is usually completed around 60 min later (i.e., at 80 min).

1. Harvest 1 ml of cells by centrifugation and fix in 70% ethanol overnight.
2. Wash cells, and resuspend in 1 ml of 50 mM sodium citrate (pH 7.0).
3. Briefly sonicate cells (*see Note 5*), and then treat with 0.25 mg/ml RNase A for 1 h at 50 °C.
4. Add Proteinase K to a final concentration of 1 mg/ml, and incubate for a further 1 h at 50 °C.
5. Dilute samples in 50 mM sodium citrate containing 16 µg/ml of propidium iodide, and incubate at room temperature for a minimum of 30 min. Remaining samples can be stored at 4 °C.
6. Analyze samples using a flow cytometer. We use a Becton Dickinson FACSCalibur machine, and CellQuest software for analysis.

### 3.6 DNA Extraction for 2DGE

This protocol is based on the hexadecyltrimethylammonium bromide (“CTAB”) method of DNA extraction [11]. For the analysis of *ARS305*-borne replication forks, the optimal time point for

2DGE analysis is approx. 30–35 min following release from G<sub>1</sub> arrest.

1. Add 0.1% (final w/v) sodium azide to 200 ml aliquots of cells at predefined time points, and harvest the cells by centrifugation ( $900 \times g$ , 5 min).
2. Prepare spheroplasts by incubating cells in 5 ml of spheroplast buffer (1 M sorbitol, 100 mM EDTA pH 8, 0.1%  $\beta$ -mercaptoethanol, 100 U/ml zymolyase (AMSBIO, UK)) for 1 h at 30 °C, 200 rpm.
3. Harvest spheroplasts by centrifugation ( $3500 \times g$ , 10 min), briefly wash in 5 ml water, and then lyse the spheroplasts in 2 ml water (*see Note 6*).
4. Add 2.5 ml of Solution I (2% w/v CTAB, 1.4 M NaCl, 100 mM Tris-HCl pH 7.6, 25 mM EDTA pH 8.0) and 200  $\mu$ l of 10 mg/ml RNase A, and incubate samples at 50 °C for 30 min.
5. Add 200  $\mu$ l of 20 mg/ml Proteinase K, and incubate samples for a further 90 min at 50 °C.
6. Add a further 100  $\mu$ l of 20 mg/ml Proteinase K, and incubate samples overnight at 30 °C.
7. Separate the supernatant (S; *see steps 8–10* below) and pellet (P; *see steps 11–13* below) by centrifugation ( $3500 \times g$  for 10 min) and treat as follows.
8. (S) Combine the supernatant with 2.5 ml of 24:1 chloroform: isoamylalcohol in 15 ml Falcon tubes, mix (*see Note 7*), and clarify by centrifugation ( $3500 \times g$  for 5 min).
9. (S) Transfer the upper phase to borosilicate glass centrifuge tubes containing 10 ml of Solution II (1% CTAB, 50 mM Tris-HCl pH 7.6, 10 mM EDTA), and incubate for 1–2 h, with occasional gentle mixing.
10. (S) Precipitate the DNA by centrifugation at  $8500 \times g$  for 10 min. Resuspend the DNA in 2.5 ml of Solution III (1.4 M NaCl, 10 mM Tris-HCl pH 7.6, 1 mM EDTA).
11. (P) Resuspend the pellet from **step 7** in 2 ml of Solution III and incubate at 50 °C for 1 h. Gently mix samples at regular intervals to facilitate resuspension.
12. (P) Combine with 2.5 ml of 24:1 chloroform:isoamylalcohol in 15 ml Falcon tubes, mix (*see Note 7*), and clarify by centrifugation ( $3500 \times g$  for 5 min).
13. (P) Combine the upper phase with the resuspended DNA from **step 10**.
14. Add an equivalent volume (5 ml) of isopropanol, mix gently, and then centrifuge at  $8500 \times g$  for 10 min.

15. Briefly wash the pellet in 75% ethanol ( $8500 \times g$  for 1 min).
16. Resuspend the DNA in 10 mM Tris-HCl pH 8.0 (*see Note 8*).

### 3.7 Two-Dimensional gel Electrophoresis (2DGE) of DNA Extracts

For optimal 2DGE images, ~15–20  $\mu\text{g}$  DNA should be cut with DNA restriction enzymes that liberate a ~5 kb restriction fragment for Southern blot analysis. Restriction fragments of <4 kb or >6 kb generally give poorer-quality images in our experience. For other considerations regarding restriction enzymes, *see Note 9*. The basic principles of this technique are that the first dimension gel separates DNA restriction fragments on the basis of their size (low voltage; no ethidium bromide), whereas the second dimension gel separates DNA restriction fragments on the basis of their shape (high voltage; + ethidium bromide). The types of DNA structures that can be identified using this technique have been described extensively [11].

1. For each sample, digest 15–20  $\mu\text{g}$  of DNA with appropriate restriction enzymes for a minimum of 5 h. Typical restriction digests contain 100 U of enzymes, and a total volume of 650  $\mu\text{l}$ .
2. Add 83  $\mu\text{l}$  of 2.5 M Potassium Acetate solution (pH 6), and 730  $\mu\text{l}$  of isopropanol. Gently mix by inversion, and then centrifuge at  $16,000 \times g$  for 10 min.
3. Wash pellets in 75% ethanol, and centrifuge at  $16,000 \times g$  for 5 min.
4. Air-dry the pellets to remove all residual ethanol, and then resuspend the DNA in 20  $\mu\text{l}$  of TE buffer. Allow a minimum of 1 h for resuspension of DNA. In the meantime, prepare the first dimension gel.
5. Prepare a 0.35% low EEO agarose (US Biological, USA) gel. The agarose should be 55 °C when pouring, and the gels should be poured in a cold room (*see Note 10*).
6. Add 5  $\mu\text{l}$  of 2DGE loading dye to each of the DNA samples, gently mix, and then carefully load onto the first dimension gel. Leave a spare lane between each sample to allow cutting of gel slices the next day. Load DNA markers in the final lane.
7. Run first dimension gels overnight at 50 V for 18–21 h in TBE buffer. The length of time depends on the gel apparatus, and the size of restriction fragment being analyzed. In our standard conditions, the linear form of the restriction fragment (~5 kb) has usually migrated ~10 cm.
8. Once the first dimension is complete, stain the gel with TBE + 0.3  $\mu\text{g}/\text{ml}$  ethidium bromide for 30 min.
9. Cut 6.5 cm gel strips from first dimension gels that contain the 1N  $\rightarrow$  2N restriction fragments (i.e., cut from ~4  $\rightarrow$  12 kb). Rotate gel slices 90° and arrange (six gel slices per second



dimension gel) in two horizontal lines of three, separated vertically by a minimum of 10 cm. Use a large knife and flexible plastic ruler to help maneuver the fragile gel slices. Pour a second dimension 0.90% agarose gel (+0.3  $\mu\text{g}/\text{ml}$  ethidium bromide) around these.

10. Run second dimension gels at 180 V (in TBE buffer containing 0.3  $\mu\text{g}/\text{ml}$  ethidium bromide) for 8 h at 4 °C. The DNA from each sample will be visible as an arc on the second dimension gels.
11. Cut second dimension gels into 20 cm  $\times$  10 cm slices, and transfer the DNA samples to membranes by Southern blotting. The following wash procedures should be performed prior to assembling the Southern blot apparatus:

Depurination solution	15 min
Denaturing solution	20 min
Neutralization solution	20 min

12. The Southern blot apparatus (from bottom to top) should be arranged as follows: Whatman paper wick soaked in 10 $\times$  SSC, second dimension Gel, Pre-soaked membrane, 2 $\times$  soaked Whatman paper, 2 $\times$  dry Whatman paper, paper towels (~10–15 cm), plastic tray, weight. Take care to remove any air bubbles between the gel and membrane, and use parafilm to cover any exposed parts of the paper wick.
13. After 16–24 h, the DNA is (faintly) detectable on membranes using a UV transilluminator. After drying the membranes for 30 min on filter paper, expose both the sides of the membranes to 2 $\times$  cycles of 120,000  $\mu\text{J}$  UVC. Membranes can then be enclosed in filter paper until they are ready to hybridize.

### 3.8 Hybridization of Membranes

DNA replication intermediates can be detected using unique locus-specific  $^{32}\text{P}$  dCTP (6000 Ci/mmol) radiolabeled probes (*see Note 11*).

1. Wash membranes three times in water and add (up to 4) to hybridization tubes.
2. Add 80 ml of pre-hybridization buffer to each tube, and incubate at 65 °C with rotation for >1 h.
3. To make the radiolabeled probe, denature 100 ng of the DNA probe (in 47.5  $\mu\text{l}$  TE buffer) at 100 °C for 10 min, and then transfer the denatured DNA to an ice:water bath for 5 min.
4. Combine the DNA from **step 3** with 2.5  $\mu\text{l}$  of  $^{32}\text{P}$  dCTP (6000 Ci/mmol) in the Rediprime II DNA labeling system. Incubate at 37 °C for >15 min.

5. Remove the unincorporated nucleotides using a Quick Spin column.
6. Boil the flow-through at 100 °C for 10 min. Also boil the carrier DNA aliquots (500 µl of 4 mg/ml herring sperm DNA).
7. For each hybridization tube, prepare 20 ml of 65 °C pre-warmed hybridization buffer, and combine the following:
  - (a) >5 µl of radiolabelled probe (split the flow-through evenly between the hybridization tubes).
  - (b) Denatured carrier DNA.
8. Pour off the pre-hybridization buffer, and replace with the hybridization solution from **step 7**.
9. Incubate overnight at 65 °C in a hybridization oven (with rotation).
10. The next day, pour off the hybridization buffer, and add 50 ml Wash Solution I. Incubate at 65 °C (with rotation) for 10 min.
11. Transfer all of the membranes to a plastic box, and wash them as follows (at room temperature, with gentle agitation):
  - 500 ml Wash Solution I for 15 min.
  - 500 ml Wash Solution II for 10 min (×3 washes).
12. Dry the membranes on filter paper for 30 min, and then wrap in a cling film. Transfer the membranes to phosphorscreen cassettes, and store overnight.
13. Scan the phosphorscreens using a phosphorimaging system (e.g., Typhoon, Storm, or Phosphorimager). Replication intermediates are usually detectable after 1–2 days.
14. If required, stripping of membranes for subsequent reprobings can be achieved by incubating the membranes in a boiling solution of 0.1% SDS for >1 h at 65 °C (with rotation). Membranes can then be probed again using the above procedure.

---

## 4 Notes

1. *Ter* targeting primers should be PAGE-purified following synthesis to ensure they are the correct sequence. If particular combinations of 65-bp targeting sequences are inefficient, then the length of the homology arms can be increased through an additional round of PCR.
2. *Ter* vectors should be propagated using Tus-negative bacteria. Small aliquots of the desired vectors, as well as their complete sequences, are available on request. The HA-Tus expression plasmids (with either *LEU2*, *MET15*, or *HIS3* markers) are also available on request.

3. When plating the transformed cells, the cell density is important. We recommend plating 1/3rd and 2/3rd of the cells, respectively, onto two large selection plates to ensure differing cell densities. If using this protocol for making gene disruptants requiring drug selection, resuspend the cells in YPD for 3 h at 30 °C, 200 rpm prior to plating. After this time, harvest the cells by centrifugation at 1000 rpm for 1 min, and then plate the cells onto large YPD plates containing the desired drug. We typically use the following markers/drug doses for gene disruptants: *kanMX* (400 mg/L G418), *hphMX* (300 mg/L Hygromycin B), or *natMX* (100 mg/L CloNAT).
4. When starting the alpha-factor arrest, the cell cycle distribution should be approximately 50:50 1C:2C, as judged by flow cytometry. If the majority of cells in the culture have a 1C content, modifying the growth conditions to improve aeration is recommended. Alpha-factor should also be batch-tested to find the optimal concentration for efficient G<sub>1</sub>-arrest, and subsequent release.
5. Over-sonication of cells can cause DNA fragmentation (detectable by flow cytometry as a sub-1C peak). We have observed that different strain backgrounds require different amounts of sonication to disrupt cell clumps. For BY4741, we typically use ~13 W output for 10 s. However, we recommend that the sonication conditions are optimized for each strain/sonicating apparatus using trial runs of logarithmic, G<sub>1</sub>-arrested, and G<sub>2</sub>/M-arrested cells, for validation.
6. During the 5 ml wash step, carefully invert the tubes to wash the sides of the tubes. This step helps remove any residual sorbitol that can interfere with later steps. Following addition of 2 ml water, incubate the samples at 200–300 rpm until a homogenous mixture is obtained. This can usually take up to an hour.
7. Shake the tubes three times, and release any pressure due to gas build-up. Do not shake the samples too vigorously or violently, but just enough to get a (transiently) even suspension. After the separation of the phases by centrifugation, take care not to accidentally take up any of the lower phase while pipetting.
8. Resuspend DNA in a minimum of 500 µl of 10 mM Tris pH 8. Incubate samples at room temperature for 1–2 h to facilitate resuspension of DNA, and then store the samples at 4 °C. Do not freeze samples at –20 °C, as this can destroy fragile replication intermediates. Always use cut-tips when transferring the DNA and avoid vortexing or any mechanical stress.
9. The same membranes can be stripped and reprobed multiple times to analyze different loci. Therefore, if strains contain two Tus-*Ter* barriers, one should choose compatible enzymes that

give ~5 kb fragments in each case. Other genomic loci may also be considered controls for early/late-replicating regions. We generally find that addition of multiple restriction enzymes often improves the quality of the 2DGE images, as this helps remove background signals.

10. First dimension gels are very fragile and should be handled carefully in all subsequent steps. Take care in particular to remove the gel combs carefully while the gel is submerged in TBE buffer. It is recommended that the wells are stained with a small amount of loading dye prior to loading the DNA samples to (1) ensure that the wells are intact, and (2) facilitate loading of samples.
11. This should be a unique and pure PCR product for the region of the genome to be analyzed. This can be made by doing a two-step PCR, using “nested” primers for the second step and low amounts (10–20 pg) of the primary PCR product as template.

---

## Acknowledgments

Work in the authors’ laboratory is funded by the Danish National Research Foundation (DNRF115), The European Research Council, The Novo Nordisk Foundation, and The Nordea Foundation.

## References

1. Hill TM, Marians KJ (1990) Escherichia coli Tus protein acts to arrest the progression of DNA replication forks in vitro. *Proc Natl Acad Sci U S A* 87(7):2481–2485
2. Larsen NB, Hickson ID, Mankouri HW (2014) Tus-Ter as a tool to study site-specific DNA replication perturbation in eukaryotes. *Cell Cycle* 13(19):2994–2998. doi:10.4161/15384101.2014.958912
3. Larsen NB, Sass E, Suski C et al (2014) The Escherichia coli Tus-Ter replication fork barrier causes site-specific DNA replication perturbation in yeast. *Nat Commun* 5:3574. doi:10.1038/ncomms4574
4. Willis NA, Chandramouly G, Huang B et al (2014) BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature* 510(7506):556–559. doi:10.1038/nature13295
5. Willis NA, Scully R (2016) Spatial separation of replisome arrest sites influences homologous recombination quality at a Tus/Ter-mediated replication fork barrier. *Cell Cycle*:1–9. doi:10.1080/15384101.2016.1172149
6. Natsume T, Kiyomitsu T, Saga Y et al (2016) Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep* 15(1):210–218. doi:10.1016/j.celrep.2016.03.001
7. Calzada A, Hodgson B, Kanemaki M et al (2005) Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes Dev* 19(16):1905–1919. doi:10.1101/gad.337205
8. Ahn JS, Osman F, Whitby MC (2005) Replication fork blockage by RTS1 at an ectopic site promotes recombination in fission yeast. *EMBO J* 24(11):2011–2023. doi:10.1038/sj.emboj.7600670
9. Lambert S, Watson A, Sheedy DM et al (2005) Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. *Cell* 121(5):689–702. doi:10.1016/j.cell.2005.03.022
10. Iraqi I, Chekkal Y, Jmari N et al (2012) Recovery of arrested replication forks by homologous recombination is error-prone. *PLoS Genet* 8(10):e1002976. doi:10.1371/journal.pgen.1002976

11. Liberi G, Cotta-Ramusino C, Lopes M et al (2006) Methods to study replication fork collapse in budding yeast. *Methods Enzymol* 409:442–462
12. Mankouri HW, Huttner D, Hickson ID (2013) How unfinished business from S-phase affects mitosis and beyond. *EMBO J* 32(20): 2661–2671. doi:[10.1038/emboj.2013.211](https://doi.org/10.1038/emboj.2013.211)
13. Raghuraman MK, Winzeler EA, Collingwood D et al (2001) Replication dynamics of the yeast genome. *Science* 294(5540):115–121. doi:[10.1126/science.294.5540.115](https://doi.org/10.1126/science.294.5540.115)
14. Muller CA, Hawkins M, Retkute R et al (2014) The dynamics of genome replication using deep sequencing. *Nucleic Acids Res* 42(1):e3. doi:[10.1093/nar/gkt878](https://doi.org/10.1093/nar/gkt878)
15. Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2(1):31–34. doi:[10.1038/nprot.2007.13](https://doi.org/10.1038/nprot.2007.13)
16. Looke M, Kristjuhan K, Kristjuhan A (2011) Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques* 50 (5):325–328. doi:[10.2144/000113672](https://doi.org/10.2144/000113672)

# Chapter 21

## Single Cell Gel Electrophoresis for the Detection of Genomic Ribonucleotides

Barbara Kind, Christine Wolf, Kerstin Engel, Alexander Rapp,  
M. Cristina Cardoso, and Min Ae Lee-Kirsch

### Abstract

Single cell gel electrophoresis or comet assay enables the quantification of DNA damage such as single-strand or double-strand breaks on a single cell level. Here, we describe a variant of this method for the detection of ribonucleotides embedded in genomic DNA. Briefly, cells are embedded in agarose on a microscopic slide, lysed under high salt and alkaline conditions and then subjected to in situ treatment with *E. coli* RNase HII which nicks 5' to a ribonucleotide within the context of a DNA duplex thereby converting genomic ribonucleotides into strand breaks. After unwinding of genomic DNA using a highly alkaline buffer, electrophoresis under mild alkaline conditions is performed resulting in formation of comets due to migration of fragmented DNA toward the anode. Following SYBR Gold staining comets can be visualized by fluorescence microscopy. In this setting, the length and the intensity of comets formed reflect the level of genomic ribonucleotides present in a given cell.

**Key words** Comet assay, Single cell electrophoresis, Alkaline lysis, RNase H2, Ribonucleotides, Genomic DNA

---

## 1 Introduction

Ribonucleotides misincorporated into genomic DNA by replicative polymerases represent the most frequent DNA base lesion in replicating mammalian cells [1]. In vivo, genomic ribonucleotides are removed by ribonucleotide excision repair, which is initiated by ribonuclease H2 (RNase H2) through cleavage 5' of a ribonucleotide followed by strand displacement synthesis by Pol  $\delta$ , flap removal by FEN1, and ligation by DNA ligase I [2]. If left unrepaired, ribonucleotides render the DNA backbone susceptible to DNA strand breaks leading to genome instability [1, 3–5]. Thus, RNase H2 plays an essential role in the maintenance of genome integrity. In humans, mutations in the genes encoding the three

---

Barbara Kind and Christine Wolf are co-first authors.

Marco Muzi-Falconi and Grant W. Brown (eds.), *Genome Instability: Methods and Protocols*, Methods in Molecular Biology, vol. 1672, DOI 10.1007/978-1-4939-7306-4\_21, © Springer Science+Business Media LLC 2018

RNase H2 subunits (*RNASEH2A*, *RNASEH2B*, *RNASEH2C*) cause the neuroinflammatory disorder Aicardi-Goutières syndrome [6]. While complete RNase H2 deficiency in mice is embryonic lethal due to massive DNA damage [1, 5], RNase H2 mutations found in patients with Aicardi-Goutières syndrome were shown to be hypomorphic [7, 8]. In patient cells, increased levels of genomic ribonucleotides cause chronic low level DNA damage leading to the activation of the antiviral type I interferon axis [8–10].

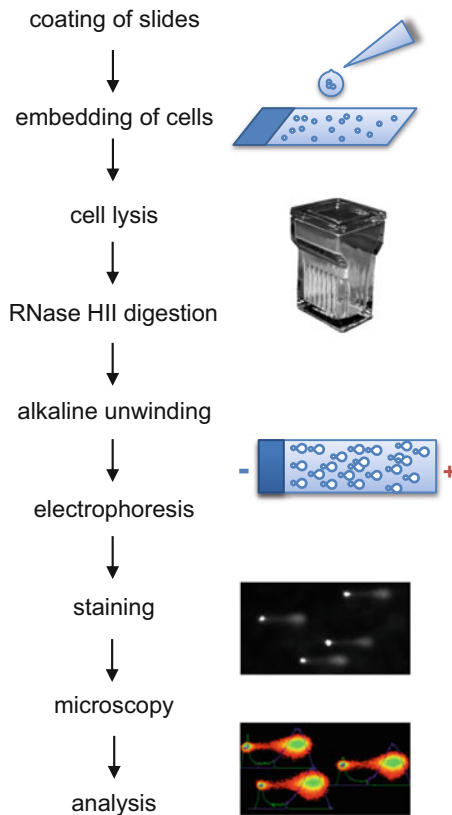
Here, we describe a method for the detection and quantification of ribonucleotides embedded in genomic DNA of human fibroblasts based on single cell gel electrophoresis or comet assay [8]. The basic principle of the comet assay rests on the migration of negatively charged DNA fragments out of agarose-embedded nuclei toward the anode during electrophoresis resulting in the formation of comet tails [11–13]. While electrophoresis under neutral pH conditions predominantly exposes DNA double-strand breaks, electrophoresis under highly alkaline denaturing conditions (pH >13) further exposes single-strand breaks in DNA [12]. Furthermore, the use of specific endonucleases that convert base lesions into single-strand breaks enables detection of lesion-specific DNA damage. For example, treatment of cells with 8-oxoguanine DNA glycosylase prior to electrophoresis allows detection of oxidized bases [14, 15]. In our protocol *E. coli* RNase HII is used to nick genomic DNA at sites of embedded ribonucleotides to create single-strand breaks. The steps involved in the comet assay for the detection of genomic ribonucleotides are depicted in Fig. 1. Cells are first embedded in agarose on a microscopic slide and lysed to remove membranes, cytoplasm, and nucleoplasm including nucleosomes. The remaining nucleoids are subjected to RNase HII treatment followed by electrophoresis using a mildly alkaline buffer. Migration of fragmented DNA leads to the formation of comets, the sizes of which reflect the extent of DNA fragmentation generated by RNase HII treatment. The representative images of comets obtained with this method in human fibroblasts deficient in RNase H2 are shown in Fig. 2.

---

## 2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water, to attain a sensitivity of  $18 \text{ M}\Omega \times \text{cm}$  at  $25^\circ \text{C}$ ) and analytical grade reagents. Prepare and store all reagents at room temperature unless indicated otherwise.

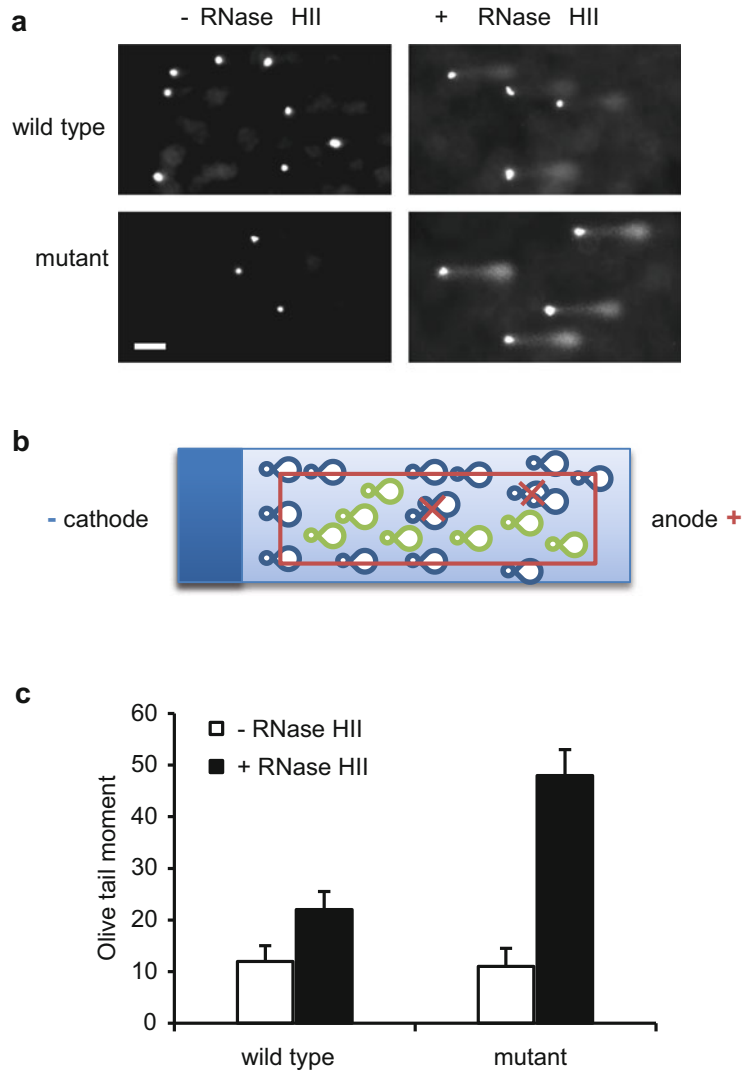
1. Low melting Agarose (1%): Dissolve 1 g agarose (type VII, low gelling temperature; Sigma Aldrich) in 100 ml  $1 \times \text{PBS}$ . Make aliquots and store at  $-20^\circ \text{C}$ .



**Fig. 1** Flow chart depicting steps involved in the modified comet assay for the detection of genomic ribonucleotides.

2. Normal melting Agarose (0.5%): Dissolve 0.5 g agarose (type II; Sigma Aldrich) in 100 ml 1× PBS. Make aliquots and store at  $-20^{\circ}\text{C}$ .
3. Normal melting Agarose (1%): Dissolve 1 g agarose (type II; Sigma Aldrich) in 100 ml 1× PBS. Make aliquots and store at  $-20^{\circ}\text{C}$ .
4. Antifade solution with SYBR Gold: Dissolve 2.33 g DABCO (1,4-diazabicyclo[2.2.2]octane; 20 mM) in 2 ml 1 M Tris-HCl (pH 8.0) and 80 ml glycerol by heating to  $70^{\circ}\text{C}$ . Add distilled water to a volume of 100 ml. Add 1  $\mu\text{l}$  of SYBR Gold stock solution (Thermo Fisher Scientific) to 10 ml Antifade solution and store in a dark vial at  $-20^{\circ}\text{C}$ .
5. Slides (Dakin Fully Frosted Slides, 71876-01; Electron Microscopy Sciences).
6. Phosphate-buffered saline (1× PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 2 mM  $\text{KH}_2\text{PO}_4$ : Dissolve 8 g NaCl, 0.2 g KCl, 1.44 g  $\text{Na}_2\text{HPO}_4$  and 0.24 g  $\text{KH}_2\text{PO}_4$  in 800 ml distilled  $\text{H}_2\text{O}$ . Adjust pH to 7.4 with HCl. Add distilled water to a volume of 1 L.





**Fig. 2** Modified comet assay of RNase H2-deficient human fibroblasts. **(a)** Images of comets obtained from primary human fibroblasts of a healthy control (wild type) and a patient with RNase H2 deficiency (mutant) with and without RNase HII (-/+ RNase HII) treatment. Scale bar 20  $\mu$ m. **(b)** Schematic of slide after electrophoresis and SYBR Gold staining. The red box indicates the area from which comets should be sampled for analysis. Superimposed comets should be excluded from analysis. **(c)** Olive tail moments as determined with CASP software. At least 50 comets were analyzed per slide. Shown are the means and SD.

7. Tris-borate-EDTA (10 $\times$  TBE): 0.9 M Tris base, 0.9 M boric acid, 20 mM EDTA): Dissolve 108 g Tris, 55 g boric acid and 9.3 g Na<sub>2</sub>H<sub>2</sub>EDTA in 800 ml distilled water. Add distilled water to a volume of 1 L.

8. Alkaline lysis buffer: 2.5 M NaCl, 0.1 M EDTA, 10 mM Tris. Dissolve 73.05 g NaCl, 18.61 g Na<sub>2</sub>H<sub>2</sub>EDTA and 0.606 g Tris in 400 ml distilled water. Adjust to pH 10 with 1 N NaOH. Adjust volume to 500 ml with distilled water. Add 1 ml Triton X-100 per 100 ml buffer (1%) prior to lysis.
9. RNase H2 buffer: 20 mM Tris-HCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 2 mM MgSO<sub>4</sub>. Mix 20 ml 1 M Tris-HCl (pH 8.0), 10 ml 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 ml 1 M KCl and 2 ml 1 M MgSO<sub>4</sub>. Adjust to pH 8.8 with 1 N NaOH. Add distilled water to a volume to 1 L. Add 1 ml Triton X-100 per 1 L buffer (0.1%) prior to use.
10. *E. coli* RNase HIII (M0288; NEB).
11. 0.1% SDS: Dissolve 1 g SDS in 1 L distilled water.
12. Alkaline unwinding buffer: 0.3 M NaOH, 1 mM Na<sub>2</sub>H<sub>2</sub>EDTA. Mix 60 ml 5 N NaOH and 2 ml 0.5 M Na<sub>2</sub>H<sub>2</sub>EDTA with 800 ml distilled water. Adjust pH to 13.1 with 10 N NaOH and add distilled water to a volume of 1 L.
13. Neutral electrophoresis buffer (0.5× TBE).

---

### 3 Methods

Carry out all the procedures at room temperature unless otherwise specified.

#### 3.1 Coating of Slides

Pre-warm slides on a hot plate (IKAMAG Rec-G) at 50 °C. First layer: Disperse 50 µl of warm (50 °C) 0.5% normal melting agarose per slide using a second slide to spread the agarose. Let agarose solidify. Second layer: Disperse 400 µl of warm (50 °C) 1% normal melting agarose per slide. Quickly cover the agarose with a 24 × 60 mm coverslip to smoothen the surface (*see Note 1*). Agarose-coated slides can be stored in a moist chamber at 4 °C for up to 1 week.

#### 3.2 Embedding of Cells

Melt aliquots of 1% low melting agarose in a water bath at 42 °C. For cooling of slides, prepare ice box and place metal plate on ice (*see Note 2*). Prepare alkaline lysis buffer with 1% Triton X and cool to 4 °C. Trypsinize fibroblasts and centrifuge at 200 × *g* for 5 min and 4 °C. Discard the supernatant, resuspend cells in 1 ml 1× PBS. Dilute cell suspension to 1 × 10<sup>6</sup> cells per ml in 1× PBS. Pre-warm agarose-coated slides to 42 °C on a hot plate and remove coverslip shortly prior to adding the agarose-cell suspension. Add 150 µl of cell suspension to 600 µl of 1% low melting agarose and mix carefully using a pipette with a cut tip to prevent shearing of cells. Pipette 150 µl of agarose-cell suspension per precoated slide and quickly cover with a 24 × 60 mm coverslip. Transfer the slide immediately onto a cold plate on ice.

**3.3 Cell Lysis**

Allow the agarose-cell suspension to solidify for 5 min, carefully remove the coverslip, and let slides air-dry for a few minutes. Incubate slides in a glass staining jar filled with precooled alkaline lysis solution at 4 °C overnight (*see Note 3*). Wash slides 10× for 10 min each with 1× PBS.

**3.4 RNase HII Digestion**

Equilibrate slides 3× for 20 min each in RNase H2 buffer, drain excess buffer by placing slides upright on a paper towel. Dilute *E. coli* RNase HII in RNase H2 buffer at 100 U per 50 µl. Add 50 µl of RNase HII solution per slide. Add 50 µl RNase H2 buffer without enzyme per no enzyme control slide. Cover slides with a 24 × 60 mm coverslip and incubate them in a wet chamber at 37 °C overnight. Stop RNase HII digestion by washing the slides in 0.1% SDS for 10 min. Wash slides 10× for 10 min each with 1× PBS.

**3.5 Alkaline Unwinding**

Wash slides 2× using alkaline unwinding buffer for a few seconds. Equilibrate cells for 20 min in alkaline unwinding buffer. Wash slides 10× for 10 min each in 0.5× TBE.

**3.6 Electrophoresis Under Neutral Conditions**

Precool neutral electrophoresis buffer to 4 °C and pour into an electrophoresis chamber (Owl Scientific, model 5) just to cover the tray holder. Place slides next to each other onto tray holder with the short edges facing the electrodes. Fill up electrophoresis chamber with buffer. Conduct electrophoresis at 1 V/cm and 15 mA for 25 min (*see Note 4*).

**3.7 Staining of Comets**

Place slides upright on paper towel for 10 min to drain buffer. Place slides horizontally on a paper towel, add 200 µl Antifade solution with SYBR Gold per slide and cover with a coverslip. Store slides in a dark wet chamber at 4 °C for at least 30 min before microscopic analysis.

**3.8 Microscopic Analysis**

Use a fluorescence microscope with an excitation filter at 495 nm and an emission filter at 537 nm and a 10× objective. Image 50–100 randomly selected comets per slide using identical imaging settings. The fluorescence intensity of the comet tail relative to the head reflects the number of DNA breaks. Quantify the degree of DNA migration by assessing the Olive tail moment using comet assay analysis software such as CASP (CaspLab) [16] (*see Note 5*). The Olive tail moment is defined as the product of the tail length and the fraction of total DNA in the tail [11].

---

**4 Notes**

1. Avoid formation of bubbles in agarose, which could interfere with electrophoresis. Per cell line to be analyzed prepare three slides with RNase HII digestion and three slides without RNase HII as no enzyme control.

2. We use the stainless-steel cover of an instrument tray, which is large enough to fit the number of slides needed for the experiment. One may also use a glass plate.
3. For a Coplin staining jar (Thermo Fisher Scientific) holding eight slides, 70 ml lysis buffer are needed. Remove and add washing buffers very slowly to avoid agarose from detaching off the slides. We use a glass funnel pointed toward the wall of the staining jar when adding buffer.
4. To eliminate effects of electric field inhomogeneity during electrophoresis, place replicate sample slides and controls in alternating positions into the electrophoresis chamber avoiding the outer rims of the tray holder. Make sure to run enzyme treated and no enzyme controls in the same electrophoresis run. Two to three rows of slides can be placed into the electrophoresis chamber we use. Run electrophoresis in a 4 °C cold room.
5. Quantification of comets can be done either by visual scoring or by using image analysis software tools available as commercial packages or freeware (for an overview *see* [www.cometassay.com](http://www.cometassay.com)).

---

## Acknowledgment

This work was supported by grants by the Deutsche Forschungsgemeinschaft (KFO 249; LE 1074/4-1 and LE 1074/4-2 to M.L.-K.; KI 1956/2-1 to B.K.) and the Bundesministerium für Forschung und Bildung (grants 02S8355 and 02NUK017D to M.C.C.; grant 02NUK036D to A.R.). B.K. is a recipient of a Maria Reiche fellowship of TU Dresden.

## References

1. Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F et al (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* 149:1008–1022
2. Sparks JL, Chon H, Cerritelli SM, Kunkel TA, Johansson E, Crouch RJ, Burgers PM (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* 47:980–986
3. Nick McElhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundstrom EB, Johansson E, Chabes A, Kunkel TA (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6:774–781
4. Kim N, Huang SN, Williams JS, Li YC, Clark AB, Cho JE, Kunkel TA, Pommier Y, Jinks-Robertson S (2011) Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science* 332:1561–1564
5. Hiller B, Achleitner M, Glage S, Naumann R, Behrendt R, Roers A (2012) Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *J Exp Med* 209:1419–1426
6. Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R et al (2006) Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic

- congenital viral brain infection. *Nat Genet* 38:910–916
7. Kind B, Muster B, Staroske W, Herce HD, Sachse R, Rapp A, Schmidt F, Koss S, Cardoso MC, Lee-Kirsch MA (2014) Altered spatio-temporal dynamics of RNase H2 complex assembly at replication and repair sites in Aicardi-Goutieres syndrome. *Hum Mol Genet* 23:5950–5960
  8. Gunther C, Kind B, Reijns MA, Berndt N, Martinez-Bueno M, Wolf C, Tungler V, Chara O, Lee YA, Hubner N et al (2015) Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. *J Clin Invest* 125:413–424
  9. Mackenzie KJ, Carroll P, Lettice L, Tarnauskaitė Ž, Reddy K, Dix F, Revuelta A, Abbondati E, Rigby RE, Rabe B et al (2016) Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response. *EMBO J* 35:831–844. doi:[10.15252/embj.201593339](https://doi.org/10.15252/embj.201593339)
  10. Pokatayev V, Hasin N, Chon H, Cerritelli SM, Sakhuja K, Ward JM, Morris HD, Yan N, Crouch RJ (2016) RNase H2 catalytic core Aicardi-Goutières syndrome-related mutant invokes cGAS-STING innate immune-sensing pathway in mice. *J Exp Med* 213:329–336. doi:[10.1084/jem.20151464](https://doi.org/10.1084/jem.20151464)
  11. Ostling O, Johanson KJ (1984) Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells. *Biochem Biophys Res Commun* 123:291–298
  12. Singh NP, McCoy MT, Tice RR, Schneider EL (1988) A simple technique for quantitation of low levels of DNA damage in individual cells. *Exp Cell Res* 175:184–191
  13. Olive PL, Banath JP (2006) The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc* 1:23–29
  14. Collins AR, Duthie SJ, Dobson VL (1993) Direct enzymic detection of endogenous oxidative base damage in human lymphocyte DNA. *Carcinogenesis* 14:1733–1735
  15. Kushwaha S, Vikram A, Trivedi PP, Jena GB (2011) Alkaline, Endo III and FPG modified comet assay as biomarkers for the detection of oxidative DNA damage in rats with experimentally induced diabetes. *Mutat Res* 726:242–250. doi:[10.1016/j.mrgentox.2011.10.004](https://doi.org/10.1016/j.mrgentox.2011.10.004)
  16. Końca K, Lankoff A, Banasik A, Lisowska H, Kuszewski T, Góźdź S, Koza Z, Wojcik A (2003) A cross-platform public domain PC image-analysis program for the comet assay. *Mutat Res* 534:15–20

## Measuring the Levels of Ribonucleotides Embedded in Genomic DNA

Alice Meroni, Giulia M. Nava, Sarah Sertic, Paolo Plevani, Marco Muzi-Falconi, and Federico Lazzaro

### Abstract

Ribonucleotides (rNTPs) are incorporated into genomic DNA at a relatively high frequency during replication. They have beneficial effects but, if not removed from the chromosomes, increase genomic instability. Here, we describe a fast method to easily estimate the amounts of embedded ribonucleotides into the genome. The protocol described is performed in *Saccharomyces cerevisiae* and allows us to quantify altered levels of rNMPs due to different mutations in the replicative polymerase  $\epsilon$ . However, this protocol can be easily applied to cells derived from any organism.

**Key words** DNA replication, DNA repair, DNA polymerase, Ribonucleotides incorporation, RNase H, Genome stability, Genomic rNMPs

---

### 1 Introduction

During evolution, DNA was selected as the principal molecule to preserve genetic information likely due to its greater stability compared to RNA, whose 3' hydroxyl group increases its susceptibility to hydrolysis.

At every cell cycle, genomic DNA is duplicated by DNA polymerases, enzymes that are specialized to copy a single-stranded DNA template and polymerize deoxyribonucleotides (dNTPs) accordingly, forming a complementary DNA strand. Given the much greater abundance of rNTPs compared to dNTPs in the nucleus, DNA polymerases evolved a steric gate to help preventing rNTPs from entering the active site [1].

However, recent data revealed that large amounts of rNTPs are incorporated in genomic DNA during replication [2]. The presence of rNMPs in the chromosomes has physiological roles [3–5] and is

---

M. Muzi-Falconi and F. Lazzaro are co-last authors.

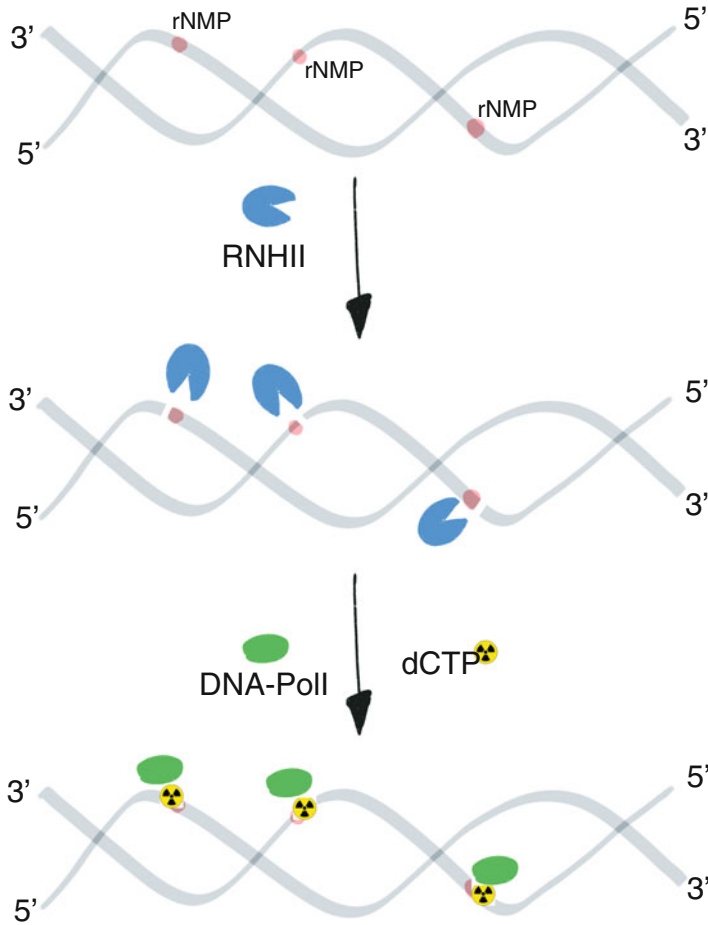
normally transient: specific RNase H-based pathways excise them before mitosis [6]. Failure to remove genomic rNMPs causes replication stress and genome instability in yeast and human cells [7–12]. Mutations in the genes coding for RNase H2 in humans are responsible for the rare Aicardi-Goutieres Syndrome (AGS) [13]. Intriguingly, cells derived from AGS patients accumulate rNMPs in their chromosomes and exhibit constitutively activated DNA damage response and post-replication repair mechanisms [8, 10].

To investigate the mechanisms underlying incorporation and removal of ribonucleotides in chromosomes and to determine their effect on genome integrity, it is important to determine rapidly and semiquantitatively the amounts of ribonucleotides present genomic DNA. Here, we describe an experimental strategy based on the approach originally described by Hiller and colleagues [12] and then in [8]. Briefly, genomic DNA is extracted and treated *in vitro* with bacterial RNase HII, which introduces nicks at every ribonucleotides site. These nicks are radioactively labeled taking advantage of the DNA Polymerase I nick translation capability (Fig. 1). In this chapter, we describe the procedure starting from the preparation of genomic DNA from yeast cells and compare the effect of two mutations affecting the steric gate of pol  $\epsilon$ , M644G and M644L, that respectively increase and decrease rNTPs incorporation [9].

---

## 2 Materials

1. Eppendorf tubes 1.5 and 2 mL.
2. Pipettes and tips.
3. Glass Pasteur pipette.
4. MilliQ water.
5. 250 mL glass flasks.
6. Stirrer.
7. Gel electrophoresis apparatus.
8. Power supply.
9. UV transilluminator and digital camera.
10. Plastic wrap.
11. Tape.
12. Scalpel.
13. Thermomixer.
14. Geiger counter.
15. Gel dryer.



**Fig. 1** Representative scheme for ribonucleotides incorporation assay. RNHII recognizes and cleaves ribonucleotides embedded into genomic DNA (*red dot*) leaving 5' P-ribose ends. The DNA-Poll enzyme, through nick translation, marks RNHII-induced nicks with radiolabeled dCTP

16. 3 MM Whatman blotting papers.
17. Towel papers.
18. Weight ( $\geq 400$  g).
19. Phosphorimager and screen.
20. YDER Yeast DNA Extraction kit, materials and reagents listed in the kit instructions (Thermo Scientific).
21. RNase A 10 mg/mL.
22. Phenol:Chloroform:Isoamyl Alcohol 25:24:1 v/v/v (Saturated with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA). Stored at 4 °C.
23. 100% ethanol. Stored at -20 °C.
24. 3 M sodium acetate, pH 7.0. Stored at 4 °C.
25. 70% ethanol. Stored at -20 °C.



26. Agarose powder.
27. 10 mg/mL ethidium bromide (EtBr).
28. TAE 1×: 40 mM Tris-HCl, pH 8.0, 20 mM acetic acid, and 1 mM EDTA. Stored at room temperature.
29. Lambda DNA marker.
30. RNase HII 5000 U/mL.
31. 10× ThermoPol<sup>®</sup> Reaction Buffer: 200 mM Tris-HCl, 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM KCl, 20 mM MgSO<sub>4</sub>, 1% Triton<sup>®</sup> X-100 (New England Biolabs).
32. 10× dNTPs mix (without dCTP); 200 μM dATP, 200 μM dGTP and 200 μM dTTP.
33. NEBuffer 2; 500 mM NaCl, 100 mM Tris-HCl, 100 mM MgCl<sub>2</sub>, 10 mM DTT (New England Biolabs).
34. DNA Polymerase I 10,000 U/mL.
35. α<sup>32</sup>P-dCTP, 3000 Ci/mmol.
36. STOP solution: 30% glycerol, 200 mM EDTA, bromophenol blue, in MilliQ water.
37. TCA 30%.
38. Software for quantification and analysis: ImageQuant and Microsoft Excel.

---

### 3 Methods

#### 3.1 Genomic DNA Preparation

1. Isolate yeast genomic DNA using the Y-DER extraction Kit according to the manufacturer's instructions. All the steps are performed as described in the kit's instructions, with the following modifications:
  - (a) Use 50 mL cultures with an OD<sub>600</sub> between 0.3 and 0.8.
  - (b) RNase A 10 mg/mL is diluted 1:1000 in the Y-PER reagent.
2. Resuspend DNA in 200 μL of MilliQ water and add an equal volume of Phenol:Chloroform:Isoamyl Alcohol 25:24:1 v/v/v saturated with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA (*see Note 1*).
3. Vortex vigorously for 15 s.
4. Centrifuge at maximum speed for 10 min at RT.
5. Carefully transfer only the aqueous phase (upper phase) to a new 1.5 mL eppendorf tube. That phase contains DNA. Do not transfer material from the interface or the lower phase. If so, repeat the procedure from **step 2**, adding an equal volume of Phenol:Chloroform:Isoamyl Alcohol 25:24:1 v/v/v saturated with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA.

6. Precipitate DNA by adding three volumes of ice-cold 100% ethanol and 1/10 of the volume of sodium acetate 3 M, pH 7.0. Mix well and keep overnight at  $-20\text{ }^{\circ}\text{C}$ .
7. Spin down the precipitate at maximum speed for 45' at  $4\text{ }^{\circ}\text{C}$ . Draw off the supernatant.
8. Wash adding 1 mL of ice-cold 70% ethanol. Spin at maximum speed for 30 min at  $4\text{ }^{\circ}\text{C}$ . Draw off the supernatant and dry the pellet at  $42\text{--}44\text{ }^{\circ}\text{C}$  (*see Note 2*).
9. Resuspend gently the pellet with 50  $\mu\text{L}$  of MilliQ water.

**3.2 Nicking DNA at Ribonucleotide Sites**

1. Quantify genomic DNA by loading 2  $\mu\text{L}$  on a 1% agarose gel in TAE (with 0.67  $\mu\text{g}/\text{mL}$  EtBr) next to 1  $\mu\text{L}$  of Lambda DNA marker. Run for 10 min at 8–10 V/cm. The genomic DNA band should be compact and easily quantifiable (*see Note 3*).
2. Normalize DNA in each sample to 25  $\text{ng}/\mu\text{L}$  by adding MilliQ water.
3. Prepare two new 1.5 mL eppendorf tubes per sample. Label the tubes with the sample name plus “-” and “+” (e.g., Sample 1- and Sample 1+) (*see Note 4*), and transfer 20  $\mu\text{L}$  (500 ng) of normalized genomic DNA to each tube.
4. Dilute 1:10 the RNase HII in ThermoPol Buffer 1 $\times$ . Prepare the two reaction mixtures (mix “-” and mix “+”) in new 1.5 mL tubes in excess with respect to the number of samples. Keep the mixtures on ice.  
1 $\times$  reaction mix recipe (Note that the final reaction volume is 50  $\mu\text{L}$ ):

	Mix “-”	Mix “+”
ThermoPol Buffer 10 $\times$	5 $\mu\text{L}$	5 $\mu\text{L}$
RNase HII	/	1 $\mu\text{L}$
ThermoPol Buffer 1 $\times$	1 $\mu\text{L}$	/
H <sub>2</sub> O MilliQ	24 $\mu\text{L}$	24 $\mu\text{L}$
Total volume	30 $\mu\text{L}$	30 $\mu\text{L}$

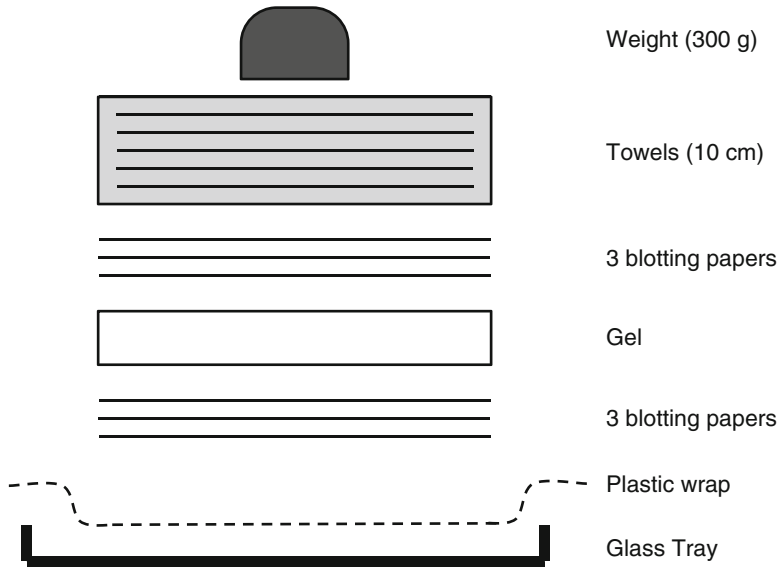
5. Vortex briefly and add 30  $\mu\text{L}$  of each mixture to the appropriate labeled tube containing genomic DNA.
6. Incubate at  $37\text{ }^{\circ}\text{C}$  with 550 rpm agitation in a thermomixer for 2.30 h.
7. Add 50  $\mu\text{L}$  of MilliQ water and precipitate DNA following **steps 6–8** of section 3.1, consider 100  $\mu\text{L}$  of total volume.
8. Resuspend gently the pellet in 20  $\mu\text{L}$  of MilliQ water.

### 3.3 Radioactive Labeling of Nicks

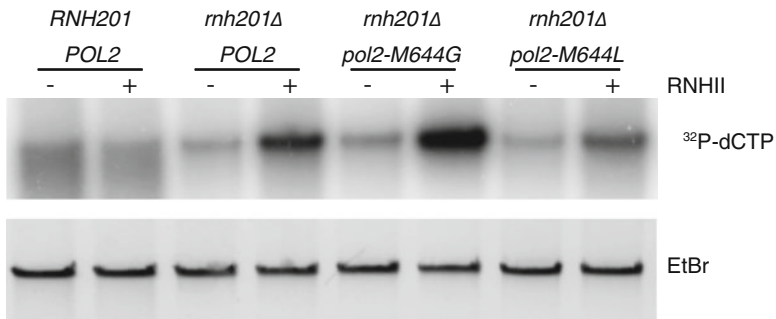
1. Quantify and normalize DNA by loading 2  $\mu\text{L}$  on a 1% agarose gel in TAE (with 0.67  $\mu\text{g}/\text{mL}$  EtBr) next to 1  $\mu\text{L}$  of Lambda DNA marker. Run for 10 min at 9–10 V/cm (*see* Note for gel preparation).
2. Transfer 300 ng of DNA to a new 1.5 mL tube and add MilliQ water to 15  $\mu\text{L}$  of total volume.
3. Prepare the common DNA Polymerase I reaction mixture, keep it on ice.  
1  $\times$  DNA Polymerase I reaction mix recipe:

10 $\times$ dNTPs mix (without dCTP):	2 $\mu\text{L}$
DNA Pol I 10 $\times$ buffer NEBuffer 2	2 $\mu\text{L}$
DNA Pol I 10,000 U/mL	0.5 $\mu\text{L}$
$\alpha^{32}\text{P}$ -dCTP 3000 Ci/mmol	0.3 $\mu\text{L}$
H <sub>2</sub> O MilliQ	0.2 $\mu\text{L}$

4. Add 5  $\mu\text{L}$  of the DNA Polymerase I reaction mix to the each sample and incubate at 16  $^{\circ}\text{C}$  for 30 min (*see* Note 5).
5. Add 4  $\mu\text{L}$  of the STOP solution (*see* Notes 5 and 6).
6. Load 20  $\mu\text{L}$  on a 1% agarose gel in TAE (with 0.67  $\mu\text{g}/\text{mL}$  EtBr). Here the DNA size marker is dispensable.
7. Run at 7 V/cm for 1.20 h.
8. Cut the gel immediately under the bromophenol blue line (*see* Note 7).
9. Examine the gel by UV light and photograph it digitally. This will allow normalization of the radioactive signal with respect to the DNA loaded in the gel.
10. Soak the gel in TCA 30% for half an hour to precipitate the DNA. The bromophenol blue turns yellow.
11. Assemble the sandwich on a glass tray: cover completely the inner part with plastic wrap. Layer three pieces of 3MM Whatman blotting paper larger than the gel. Take out the gel from TCA 30% and place it on the top of the blotting paper. Place in order: three more sheets of blotting paper on the gel, a stack of paper towels, and a weight (0.3–0.5 kg) (Fig. 2).
12. Let the gel dry overnight at room temperature.
13. Remove the paper towels and the blotting papers above the gel. Transfer the desiccated gel on a new 3MM blotting paper.
14. Dry the gel on the blotting paper using a gel dryer for 20 min at 80  $^{\circ}\text{C}$ . Use more than one blotting paper and cover the gel with plastic wrap to avoid radioactive contamination of the instrument.



**Fig. 2** Scheme to assemble paper sandwich to dry the agarose gel



**Fig. 3** Visualization of ribonucleotides incorporation assay results. The strains tested are derivatives of a W303 background (*MATA ade2-1 trp1-1 leu2-3112 his3-11,15 ura3-1 can1-100 RAD5*) with a deletion of gene coding for the catalytic subunit of RNase H2 (*rnh201Δ*) combined with wt or mutated *POL2* gene. The *RNH201 POL2* wt strain is used as control. The radiolabeled signal represents the nicks labeled by Poll. The signal dependent upon RNHII treatment is proportional to the genomic ribonucleotides levels. The EtBr panel represents the loading control, acquired before gel drying and necessary for radioactive quantification

15. Expose the dried gel on a phosphorimager screen for 5–15 min (*see Note 8*).
16. Scan the screen in a phosphorimager. To quantify the result *see Note 9*. An Example is shown in Fig. 3.

---

## 4 Notes

1. DNA has to be clean and as little nicked as possible to achieve the best resolution. For these reasons it is strongly recommended to clean DNA through phenol:chloroform extraction and ethanol precipitation. We also found that the YDER preparation yields a genomic DNA with fewer nicks compared with other methods.
2. Draw off as much supernatant as possible with a glass Pasteur pipette, this would ensure removing the majority of the ethanol, and then place the eppendorf tubes in a heater at 42–44 °C. Drying time depends on how much ethanol is left in the samples. Check the samples after 20 min, if the ethanol is still there, leave them in the heater and check later. Note that excessive drying would damage the sample, for this reason keep checking samples every 20 min until the pellet is completely dry.
3. To normalize DNA in the samples take a digital image and use quantifying tools such as ImageQuant or ImageLab (BioRad).
4. Here, each sample is split into two: one half is digested with purified bacterial RNase HII, the other half is left untreated. This allows discrimination between the ribonucleotide-dependent nicks and the nicks generated during DNA preparation.
5. Labeling the nicks is the key step of this protocol. For this reason the procedure needs strict standardization. We suggest proceeding sample by sample adding the reaction mix to each tube every 15–30 s. After 30 min of incubation, repeat this procedure with the STOP solution. In this way, all the samples would be incubated for exactly 30 min at 16 °C.
6. As most DNA polymerases, DNA polymerase I needs magnesium ions. In this case the high concentration of EDTA in the STOP solution will stop the reaction, while the glycerol and the bromophenol blue make the samples ready to be loaded on a gel.
7. Labeled genomic DNA is loaded on the agarose gel together with the DNA Polymerase I reaction mix, including the unincorporated radioactive nucleotides. The nucleotides migrate faster than genomic DNA and the long run ensures complete separation. Free nucleotides migrate immediately below the bromophenol blue; therefore, the gel is cut immediately after the run to avoid their diffusion through the gel. The bromophenol blue is also a pH-indicator, used to monitor the change in the gel pH while soaking in 30% TCA. When it turns yellow

the agarose gel pH has become acid so DNA would be precipitated into it.

8. The exposure time could vary depending on the  $\alpha^{32}\text{P}$ -dCTP activity. With fresh and fully active  $\alpha^{32}\text{P}$ -dCTP the signal is saturated in 15 min.
9. Signal quantification. To quantify the resulting signal proceed with the following step. Quantify the bands corresponding to the genomic DNA, including the ones from the EtBr capture. Use the Volume Tool of ImageQuant software drawing a rectangle around the right band. Normalize each radioactive value to the corresponding EtBr value. This would ensure the correct interpretation of the radioactive signal. Then for each sample subtract the “-” signal from the “+” signal as a background normalization.

---

## Acknowledgments

This work was supported by grants from AIRC (n.15631) and Telethon (GGP15227) to M.M.-F., from AIRC and Fondazione Cariplo (TRIDEO 2014 Id. 15724) to F.L., and from Fondazione Cariplo (grant number 2013-0798) to P.P.

## References

1. Joyce CM (1997) Choosing the right sugar: how polymerases select a nucleotide substrate. *Proc Natl Acad Sci U S A* 94:1619–1622
2. McElhinny SAN, Watts BE, Kumar D et al (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci U S A* 107:4949–4954
3. Ghodgaonkar MM, Lazzaro F, Olivera-Pimentel M et al (2013) Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell* 50:323–332
4. Lujan SA, Williams JS, Clausen AR et al (2013) Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol Cell* 50:437–443
5. Dalgaard JZ (2012) Causes and consequences of ribonucleotide incorporation into nuclear DNA. *Trends Genet* 28:592–597
6. Sparks JL, Chon H, Cerritelli SM et al (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* 47:980–986
7. Lazzaro F, Novarina D, Amara F et al (2012) RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol Cell* 45:99–110
8. Pizzi S, Sertic S, Orcesi S et al (2015) Reduction of hRNase H2 activity in Aicardi-Goutières syndrome cells leads to replication stress and genome instability. *Hum Mol Genet* 24:649–658
9. Nick McElhinny SA, Kumar D, Clark AB et al (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6:774–781
10. Günther C, Kind B, Reijns MAM et al (2014) Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. *J Clin Invest* 125(1):413–424
11. Reijns MAM, Rabe B, Rigby RE et al (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* 149:1008–1022
12. Hiller B, Achleitner M, Glage S et al (2012) Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *J Exp Med* 209:1419–1426
13. Crow YJ, Manel N (2015) Aicardi-Goutières syndrome and the type I interferonopathies. *Nat Rev Immunol* 15(7):429–440

## Mapping Ribonucleotides Incorporated into DNA by Hydrolytic End-Sequencing

Clinton D. Orebaugh, Scott A. Lujan, Adam B. Burkholder, Anders R. Clausen, and Thomas A. Kunkel

### Abstract

Ribonucleotides embedded within DNA render the DNA sensitive to the formation of single-stranded breaks under alkali conditions. Here, we describe a next-generation sequencing method called hydrolytic end sequencing (HydEn-seq) to map ribonucleotides inserted into the genome of *Saccharomyce cerevisiae* strains deficient in ribonucleotide excision repair. We use this method to map several genomic features in wild-type and replicase variant yeast strains.

**Key words** HydEn-seq, DNA replication, DNA repair, DNA polymerase, Ribonucleotide excision repair, RNase H2, Next-generation sequencing, Bioinformatics, Genomics

---

### 1 Introduction

Ribonucleotides are abundantly incorporated into DNA during eukaryotic DNA replication [1–4]. These ribonucleotides are normally removed by RNase H2-dependent ribonucleotide excision repair (RER) [5, 6]. The presence of ribonucleotides in DNA can benefit the cell [7–9], and ribonucleotides left unrepaired can lead to replication stress [10, 11], DNA strand breaks [5, 12], and autoimmune disease in humans [13–15]. Deletion of the gene encoding the catalytic subunit of the RNase H2 enzyme, *RNH201*, eliminates RER, resulting in persistent genomic ribonucleotides [2, 5, 7, 16–18]. The presence of a 2'-OH group of the ribose sugar imbedded within a DNA polynucleotide increases the chances of spontaneous DNA hydrolysis, especially in the presence of increased pH [19, 20]. Treatment of DNA containing ribonucleotides with alkali results in a strand break on the 3' side of the ribonucleotide and produces DNA ends containing a 2'-3'-cyclic phosphate and 5'-OH ends [5, 19, 20]. This provides an opportunity to monitor ribonucleotide incorporation by wild-type forms of

the three major yeast replicative DNA polymerases, Pol  $\alpha$  (*Pol1*), Pol  $\delta$  (*Pol3*), and Pol  $\epsilon$  (*Pol2*), and by variant polymerase derivatives with altered ribonucleotide incorporation propensity [1, 5, 7, 16, 18, 21–23]. This is evident by the detection of abundant alkali-sensitive sites in the nuclear genome of yeast strains harboring wild type and the variant replicases in combination with the deletion of the *RNH201* gene. Here, we describe Hydrolytic End-sequencing (HydEn-seq), a method that enables detection and mapping of ribonucleotides across the yeast genome using next-generation sequencing on an Illumina platform [21].

---

## 2 Materials

### 2.1 Genomic DNA Purification and Alkaline Hydrolysis

1. Sterile toothpicks.
2. YPDA Agar: 20 g dextrose, 20 g Bacto-peptone, 10 g Bacto-yeast extract, 20 g Bacto-agar, and 2 mL of a 0.5% adenine sulfate solution brought up to 1 L with deionized water in a graduated cylinder and sterilized for 20 min in an autoclave. Pour 30 mL into 10 cm petri dishes and cool at room temperature to solidify.
3. YPDA Broth: 20 g dextrose, 20 g Bacto-peptone, 10 g Bacto-yeast extract, and 2 mL of a 0.5% adenine sulfate solution brought up to 1 L with deionized water in a graduated cylinder and sterilized for 20 min in an autoclave.
4. Erlenmeyer flask.
5. MasterPure Yeast DNA Purification kit (Epicentre).
6. Isopropanol.
7. 70% Ethanol.
8. Microcentrifuge.
9. 1.5 mL microcentrifuge tubes (sterilized).
10. Laboratory oven.
11. 3 M KOH.
12. 3 M Na Acetate pH 5.2.

### 2.2 Library Preparation

1. PCR thermal cycler.
2. HighPrep PCR magnetic beads (MagBio Genomics).
3. Magnetic-ring stand (96-well) (ThermoFisher Scientific).
4. Phosphorylation reagents:
  - (a) T4 Polynucleotide Kinase (10,000 U/mL) (3' phosphatase minus) (NEB).
  - (b) 10 $\times$  Polynucleotide Kinase buffer (NEB).
  - (c) ATP (10 mM).



## 5. Adapter ligation reagents:

- (a) ARC 140 Oligonucleotide with 5' amino terminated C6 spacer (5'-/5AmMC6/ACACTCTTCCCTACAC-GACGCTCTCCGATCT-3') (100  $\mu$ M).
- (b) T4-RNA ligase (10,000 U/mL) (NEB).
- (c) T4-RNA Ligase 10 $\times$  buffer (NEB).
- (d) ATP (2 mM).
- (e) PEG 8000 (50% solution).
- (f)  $\text{CoCl}_3(\text{NH}_3)_6$  (10 mM).

## 6. Second strand synthesis reagents:

- (a) ARC 76/ARC 77 annealed primers (ARC 76: 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNN\*N\*N -3', ARC 77: 5'-AGATCGGAA-GAGCACACGTCTGAACTCCAGTC\*A\*C - 3', ["\*" indicates a phosphorothioate bond]) (10  $\mu$ M).
- (b) T7 DNA Polymerase (unmodified) (NEB).
- (c) T7 DNA Polymerase 10 $\times$  buffer (NEB).
- (d) Deoxynucleotide solution mix (2 mM each of dATP, dCTP, dGTP, and dTTP).
- (e) Bovine serum albumin (1 mg/mL).

## 7. PCR amplification reagents:

- (a) ARC 49 oligonucleotide (5'-AATGATACGGCGAC-CACCGAGATCTACACTCTTCCCTACAC-GACGCTCTCCGATCT-3') (2  $\mu$ M).
- (b) Barcoded oligonucleotide (5'-CAAGCAGAAGACGG-CATACGAGATX<sub>(6-8)</sub>GTGACTGGAGTTCA-GACGTGTGCTCTTCCGATCT-3' (X<sub>(6-8)</sub> is a unique barcode sequence for each library to be sequenced in the same pool)) (2  $\mu$ M).
- (c) KAPA HiFi HotStart DNA polymerase ReadyMix (2 $\times$ ) (KAPA Biosystems).

**2.3 Quality Check**

1. 2100 bioanalyzer (Agilent).
2. DNA 7500 or 1000 Reagents (Agilent).

**2.4 DNA Sequencing**

1. HiSeq2500 high-throughput DNA sequencer (Illumina).

---

**3 Methods****3.1 Yeast Growth and Genomic DNA Purification**

1. Using a sterile toothpick, streak yeast strain(s) on YPDA agar plates to produce single colonies, and incubate about 48 h at 30 °C.

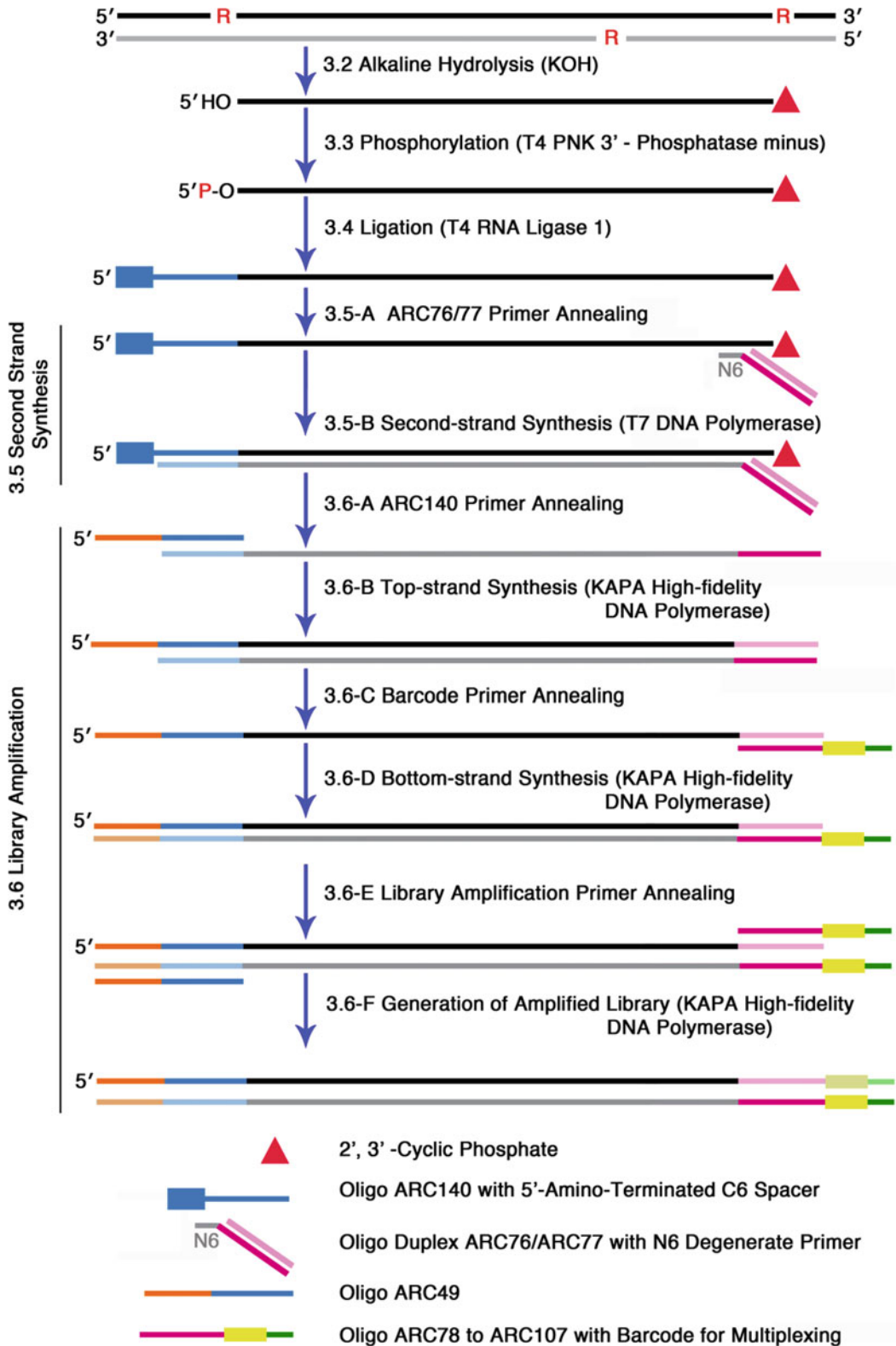
2. Resuspend about five colonies in 5 mL of YPDA broth and add a sufficient volume to 50 mL of YPDA broth (about 50–2000  $\mu\text{L}$ ) so that an overnight growth at 30 °C with shaking at 160 RPM produces a culture in exponential growth phase ( $\sim\text{OD}_{600} = 0.3\text{--}0.8$ ).
3. After overnight growth, pellet 15  $\text{OD}_{600}$  units of each culture by centrifugation, remove all growth media, and purify genomic DNA according to the Epicentre MasterPure Yeast DNA Purification kit instructions with the omission of the RNase A treatment step (*see* **Note 1**).

### **3.2 Alkaline Hydrolysis of Yeast Genomic DNA**

1. Treat genomic DNA with 300 mM KOH for 2 h at 55 °C in a laboratory oven in a total volume of 50  $\mu\text{L}$  (Fig. 1, **Step 1**).
2. Precipitate genomic DNA by combining with 0.2 volumes of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice-cold 100% ethanol. Incubate on ice for 5 min, and spin full speed in a microcentrifuge for 5 min to pellet precipitated DNA. Wash with 70% ethanol, spin again at full speed for 5 min, aspirate any remaining ethanol, and allow pellet to air-dry if necessary. Resuspend DNA pellet in 20  $\mu\text{L}$  of Qiagen buffer EB.

### **3.3 Phosphorylation of DNA 5' Termini**

1. Warm MagBio beads at room temperature during this step.
2. Assemble the phosphorylation reaction without DNA:
  - (a) 1  $\mu\text{L}$  T4 Polynucleotide Kinase (3' phosphatase minus)
  - (b) 2.5  $\mu\text{L}$  10 $\times$  Polynucleotide Kinase buffer, 2.5  $\mu\text{L}$  ATP (10 mM)
3. Denature 19  $\mu\text{L}$  of DNA resuspended from the previous step at 85 °C for 3 min, then cool on ice.
4. Add 6  $\mu\text{L}$  of reaction mix to denatured DNA and mix by pipetting up and down six to eight times.
5. Incubate the reaction for 30 min at 37 °C, then heat inactivate for 20 min at 65 °C.
6. Purify phosphorylated DNA with magnetic PCR beads (*see* **Note 2**), add 36  $\mu\text{L}$  (1.8 reaction volumes) of magnetic beads to the reaction, and mix by pipetting up and down six to eight times. Incubate at room temperature for 5 min and then pellet beads with a magnetic rack for 2–3 min or until the solution has cleared. Remove the supernatant and wash 2 $\times$  with 170  $\mu\text{L}$  of 70% ethanol. Dry beads at room temperature for  $\sim$ 5 min (It is important to remove all traces of ethanol from the beads but not overdry the beads as both the situations will result in reduced yield). Add 14  $\mu\text{L}$  Qiagen Buffer EB and incubate for 5 min. Transfer 13  $\mu\text{L}$  of the supernatant to a new tube.



**Fig. 1** HydEn-seq library construction and amplification. The individual steps of library preparation, from alkaline hydrolysis of genomic DNA through final amplification, are presented as a flowchart. Complementary sequences are indicated by *light* and *dark shades* of the same color

### 3.4 Adapter Ligation to the 5' End

1. Assemble the ligation reaction without DNA or enzyme:
  - (a) 5  $\mu\text{L}$  T4-RNA ligase 10 $\times$  buffer
  - (b) 25  $\mu\text{L}$  50% PEG 8000
  - (c) 5  $\mu\text{L}$   $\text{CoCl}_3(\text{NH}_3)_6$  (10 mM)
  - (d) 0.5  $\mu\text{L}$  ATP (2 mM)
  - (e) 0.5  $\mu\text{L}$  ARC 140 oligonucleotide (100  $\mu\text{M}$ ) (Fig. 1, Step 3) (Table 1).
2. Denature 13  $\mu\text{L}$  of DNA eluate from the previous step at 85  $^\circ\text{C}$  for 3 min, and then cool on ice.
3. Add 36  $\mu\text{L}$  of reaction mix to denatured DNA and mix by pipetting up and down six to eight times.
4. Add 1  $\mu\text{L}$  of T4 RNA ligase to each reaction and mix by pipetting up and down six to eight times.
5. Allow the reaction to incubate overnight at room temperature.
6. Purify ligation product with two magnetic bead washes. For the first, add 45  $\mu\text{L}$  of magnetic beads, wash twice with 70% ethanol, and elute with 20  $\mu\text{L}$  of buffer EB. For the second, add 36  $\mu\text{L}$  of magnetic beads, incubate and wash as described in the previous step, and elute with 14  $\mu\text{L}$  of buffer EB.

### 3.5 Second Strand Synthesis

1. Assemble the second strand synthesis reaction without DNA or enzyme:
  - (a) 2  $\mu\text{L}$  10 $\times$  T7 DNA polymerase buffer
  - (b) 2  $\mu\text{L}$  dNTP mix (2 mM)
  - (c) 0.8  $\mu\text{L}$  BSA (1 mg/mL)
  - (d) 2  $\mu\text{L}$  ARC 76/ARC 77 annealed oligonucleotides (10  $\mu\text{M}$ ) (Fig. 1, Step 4), (Table 1).
2. Denature 13  $\mu\text{L}$  of DNA eluate from the previous step at 85  $^\circ\text{C}$  for 3 min, and then cool on ice.
3. Add 6.8  $\mu\text{L}$  of the reaction mix to denatured DNA and mix by pipetting up and down six to eight times.
4. Incubate DNA and reaction mix at room temperature for 5 min and then add 0.4  $\mu\text{L}$  of T7 DNA polymerase, mix by pipetting up and down six to eight times with a 10  $\mu\text{L}$  pipette.
5. Incubate the reaction at room temperature for 5 min and then stop the reaction with the addition of 16  $\mu\text{L}$  of magnetic beads, mix by pipetting up and down six to eight times.
6. Incubate and wash beads as described in the previous step, and elute with 11  $\mu\text{L}$  of buffer EB.

### 3.6 Library Amplification

1. Assemble the library amplification reaction without DNA or barcoded oligonucleotide:

**Table 1**  
**HydEn-seq oligonucleotides**

Oligo name	Sequence (5' -3')
ARC 49	AATGATACGGCGACCACCAGAGATCTACACTCTTTCCCTACACGACGGCTCTTCCGGATCT
ARC 76	GTGACTGGAGTTCAGACCGTGTGCTCTTCCGATCTNNNN*N*N
ARC 77	AGATCGGAAAGAGCACACGCTCTGAACTCCAGTC*A*C
ARC 78	CAAGCAGAAAGACGGGCATACGAGATCGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 84	CAAGCAGAAAGACGGGCATACGAGATACATCGGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 85	CAAGCAGAAAGACGGGCATACGAGATGCCCTAAGTGAAGTGAAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 86	CAAGCAGAAAGACGGGCATACGAGATTTGGTCAAGTGAAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 87	CAAGCAGAAAGACGGGCATACGAGATCACTGTGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 88	CAAGCAGAAAGACGGGCATACGAGATATTTGGCGTGAAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 89	CAAGCAGAAAGACGGGCATACGAGATGATCTGTTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 90	CAAGCAGAAAGACGGGCATACGAGATTCAAAGTGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 91	CAAGCAGAAAGACGGGCATACGAGATCTGATCGTGAAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 93	CAAGCAGAAAGACGGGCATACGAGATAAGTGAAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 94	CAAGCAGAAAGACGGGCATACGAGATGTAGCCGTTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 95	CAAGCAGAAAGACGGGCATACGAGATTCAAAGGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 96	CAAGCAGAAAGACGGGCATACGAGATTTATTGACTGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 97	CAAGCAGAAAGACGGGCATACGAGATACGGAACTGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 98	CAAGCAGAAAGACGGGCATACGAGATTTCTGACATGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 99	CAAGCAGAAAGACGGGCATACGAGATCGGGACGGGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT
ARC 100	CAAGCAGAAAGACGGGCATACGAGATGTGGGACGTGACTGGAGTTCAGACCGTGTGCTCTTCCGGATCT

(continued)

**Table 1**  
**(continued)**

Oligo name	Sequence (5'–3')
ARC 101	CAAGCAGAAAGACGGGCATACGAGATCGTTTCAOCTGACTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 102	CAAGCAGAAAGACGGGCATACGAGATAAGGCCACCGTGA CTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 103	CAAGCAGAAAGACGGGCATACGAGATTCCGAAAACCGTGA CTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 104	CAAGCAGAAAGACGGGCATACGAGATTACGTACGGTGA CTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 105	CAAGCAGAAAGACGGGCATACGAGATATCCACTCGTGA CTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 106	CAAGCAGAAAGACGGGCATACGAGATATATCATCA GTGTGACTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 107	CAAGCAGAAAGACGGGCATACGAGATAAAGGAAATGTGA CTGGAGTTCAGACCGTGTGCTCTTCCCGATCT
ARC 140	/5AmMC6/ACACTCTTTCCCTACACGACGCTCTTCCCGATCT

Oligonucleotides used for HydEn-seq library construction and amplification. Index sequences are in *bold* for oligonucleotides ARC (78, 84–91, 93–107). Phosphorothioate bonds, which provide resistance to nuclease degradation, are indicated by (*asterisk*). 5' Amino modifier C6, which prevents formation of oligonucleotide concatamers during adapter ligation, is indicated by (/5AmMC6/). Oligos ARC 76 and ARC 77 are annealed and gel purified prior to use

- (a) 25  $\mu\text{L}$  KAPA HiFi HotStart DNA polymerase ReadyMix (2 $\times$ )
  - (b) 7.5  $\mu\text{L}$  ARC 49 oligonucleotide (2  $\mu\text{M}$ ) (Fig. 1, Step 5) (Table 1).
2. Add 32.5  $\mu\text{L}$  of reaction mix to 10  $\mu\text{L}$  of DNA eluate from the previous step.
  3. Add 7.5  $\mu\text{L}$  of barcoded oligonucleotide (ARC 78, 84–91, 93–107), (*see* Note 3) to the mixture and mix by pipetting up and down six to eight times.
  4. Run the library amplification reaction in a PCR thermal cycler using the following conditions: 98  $^{\circ}\text{C}$  hot start for 45 s, then repeat the following for a total of 18 cycles (melt DNA at 98  $^{\circ}\text{C}$  for 15 s, anneal primers at 65  $^{\circ}\text{C}$  for 30 s, and extend primers at 72  $^{\circ}\text{C}$  for 30 s), followed by a final extension at 72  $^{\circ}\text{C}$  for 60 s, cool reaction to 4  $^{\circ}\text{C}$  and hold.
  5. Purify amplified library by adding 40  $\mu\text{L}$  of magnetic beads, incubate and wash as described in the previous step, and elute in 20  $\mu\text{L}$  buffer TE.

### 3.7 Quality Check

1. Measure the concentration of each library using a fluorometric dsDNA assay. The typical range of concentrations is 1–100 ng/ $\mu\text{L}$ , although lower or higher concentrations may still work.
2. Run 1  $\mu\text{L}$  of each library on an Agilent Bioanalyzer using the DNA 7500 or DNA 1000 kit according to the manufacturer's instructions to determine average fragment size.
3. Determine the molarity of each library by dividing the concentration of each library (g/L) by the average fragment molecular weight (327 g/nucleotide  $\times$  fragment length in nucleotides).
4. Pool libraries in equimolar amounts for analysis on an Illumina HiSeq2500 high-throughput DNA sequencer.

### 3.8 Sequencing

1. Run pooled libraries on a single lane on an Illumina HiSeq2500 high-throughput DNA sequencer using paired-end mode. Save resultant FASTQ files.

### 3.9 Alignment and Visualization

1. Trim appropriate adapter sequence from the 3' ends of End1 and End2 reads using cutadapt, excluding both mates when either is trimmed shorter than 15 nt, and utilizing quality trimming to enhance ability to recognize adapter when it is present (`-m 15 -q 10`).
2. Concatenate the sequences of all oligos used during library preparation in a FASTA format, and use this file to construct a bowtie1 index (bowtie-build).
3. Align trimmed End1 reads to the oligo index using bowtie1, allowing two mismatches. Utilize `-un` option to have

unmapped reads written to a separate file. Single- rather than paired-end alignment is utilized here to achieve a higher alignment rate, resulting in a more stringent filter.

4. Extract the mates of all End1 reads passed to the “unhit” file from the trimmed End2 FASTQ by matching read ID.
5. Align paired “unhit” FASTQs to an appropriate reference genome using bowtie1, allowing two mismatches and a fragment size of up to 10 kb, retaining unique alignments only, passing all unmappable and multi-mapped reads to separate files, and trimming one additional nucleotide from the 3' end (`-m1 -v2 -un -max -3 1 -best -X10000`). Utilizing the `-best` option ensures multi-mapped reads are not erroneously categorized as unmapped, and 3' trimming allows the successful alignment of 100% complementary pairs.
6. Perform single-end alignment of remaining unmapped End1 reads to reference genome, allowing two mismatches, retaining only unique alignments, and passing unmappable and multi-mapped reads to separate files, as before (`-m1 -v2 -un -max`). This is included to recover the 5' ends of fragments rendered unmappable by errors in the End2 read alone.
7. Extract the 5' mapping position of all End 1 reads for alignments performed in **steps 5** and **6**. Shift positions 1 nt upstream (subtract 1 from coordinate for plus strand, add 1 for minus strand) and determine total read count per position, per strand, genome-wide. From this information, generate bedGraph or bigWig files for visualization on the UCSC or other genome browsers. These counts will be utilized for all downstream analyses other than normalization.
8. Perform paired-end alignment of multi-mapped reads from **step 5**, and single-end alignment of multi-mapped reads from **step 6**, reporting the single best alignment, but otherwise utilizing identical parameters.
9. Extract the 5' mapping positions, shift upstream, and determine counts genome-wide as in **step 7**. Combine with counts derived from unique alignments, and use resulting values when determining telomere-based normalization factors.

### **3.10 Generating Heatmaps/Metagene Plots**

1. For a list of strand-specific genomic features of interest, select an appropriate single-nucleotide “anchor” (e.g., Transcription start sites (TSS), start of Autonomously Replicating Sequence Consensus Sequences (ACS)).
2. Determine start and end coordinates of desired number of equally sized bins over a fixed size region relative to each anchor point, for example, a  $-1000$  to  $+999$  window, separated into 40 bins of 50 nt each. Bins should be arranged in a strand-specific manner, such that bin 1 of a plus strand feature



should be positioned at anchor  $-1000$  to anchor  $-951$ , and bin 1 of a minus strand feature should be positioned at anchor  $+951$  to anchor  $+1000$ .

3. For each bin, determine the total counts of intersecting HydEn-seq 5' ends, and arrange these values in a matrix, with each row representing a feature, and each column consisting of all bins with the same position relative to the anchor. Separate matrices may be generated for same- and opposite-strand reads, and this may be performed using raw or normalized counts.
4. Heatmaps may be generated from these matrices using Partek Genomics Suite, or other appropriate visualization tools (Fig. 2a).
5. Perform Meta-analyses by summing or averaging the counts in each column, and plotting these values, utilizing the bin center as the x-coordinate (Fig. 2b).
6. An additional context may be provided to meta-analyses with a paired plot of nucleotide frequencies. Extract the strand-specific genomic sequence, and identify the total count of A, T, G, and C for each bin. Arrange these counts in a matrix similar to those described above, determine the aggregate frequency of each nucleotide per column, and plot as in **step 5**.

### 3.11 Assessing Ribonucleotide Base Identity

1. Identify start and end coordinates of genomic regions of interest, for example, those determined to have the highest strand bias for a given variant replicase.
2. Extract the strand-specific genomic sequence for the selected regions. Determine the total counts of A, T, G, and C to provide genomic background frequencies.
3. For each position examined, determine the number of HydEn-seq 5' ends mapped to the same location and strand. Add each value to the total for the corresponding ribonucleotide.
4. Ribonucleotide frequencies may be visualized and compared to genomic background and across HydEn-seq data sets by plotting as stacked column/bar charts.

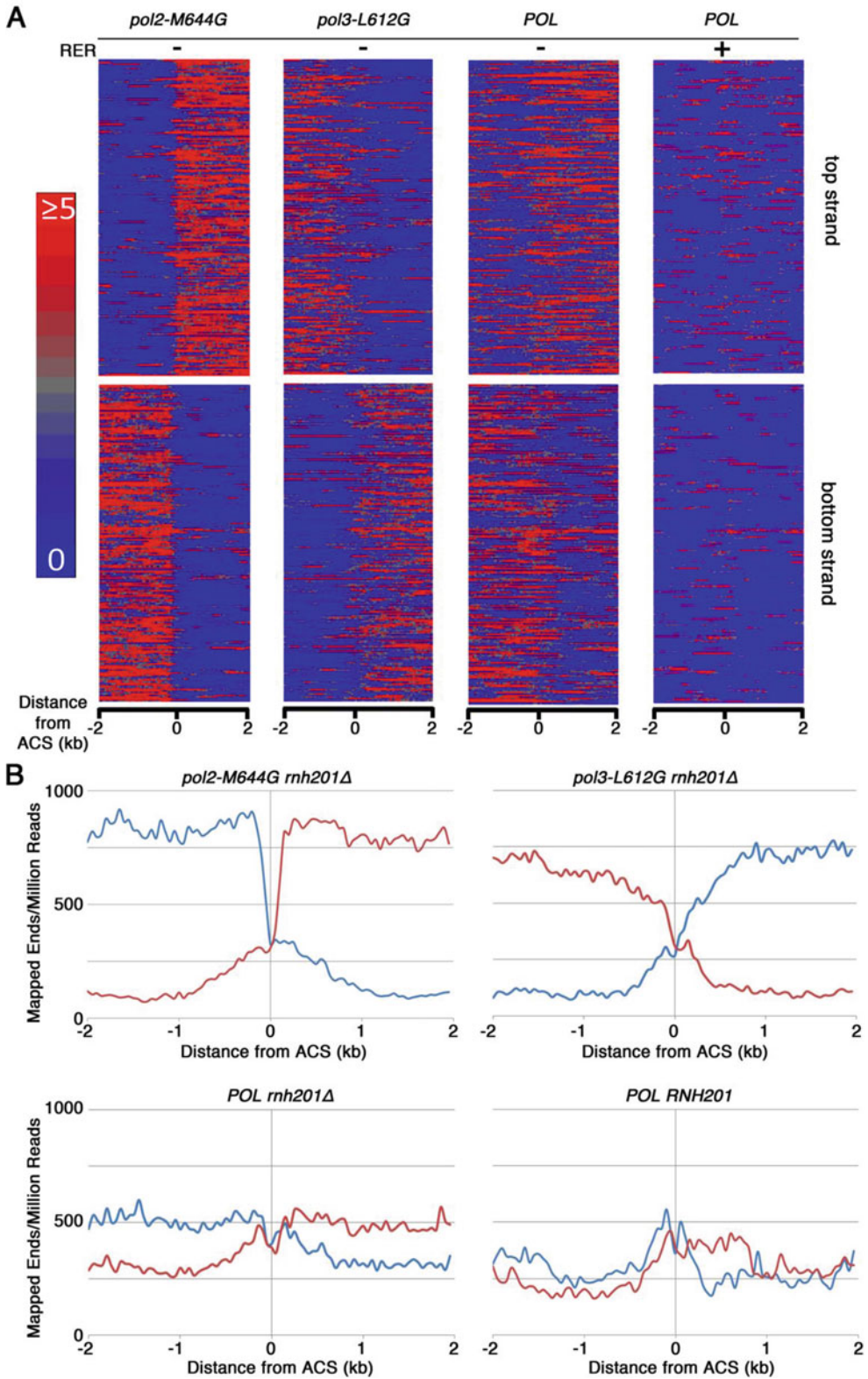
### 3.12 Normalization Versus Chromosomal 5' Ends

1. Define telomeric regions as sections of the reference genome that consist of only telomeric repeats and at least one read length from no non-telomeric sequence. The total length of these regions is

$$L_{\text{telomere}} = \sum_{i=1}^{2 \times N_C} \max(0, L_{i, \text{telomere}} - L_{\text{read}}),$$

where  $N_C$  is the chromosome count (16 in *S. cerevisiae*).

2. Count single- and multi-mapped reads (from Subheading 3.9, **step 8**) whose 5'-ends (shifted as per Subheading 3.9, **step 9**) map within telomeric regions ( $N_{\text{telomere}}$ ) and single-mapped reads outside of telomeric regions ( $N_{\text{bulk}}$ ).



**Fig. 2** Representative stacked ACS Intensity Plot and ACS Average Meta-Analysis. (a) Intensity plots show mapped ribonucleotides/million reads 2 kilobases upstream and downstream of 200 vertically stacked

3. The HydEn-seq end density, excluding telomeric regions is

$$R_{\text{bulk}} = \frac{N_{\text{bulk}}}{2 \times L_{\text{genome}} - L_{\text{telomere}}},$$

where  $L_{\text{genome}}$  is the genome length ( $\sim 12$  Mbp in *S. cerevisiae*).

4. The mean number of 5' chromosome ends per telomere is

$$N_{\text{telomere}} = \frac{N_{\text{telomere}} - L_{\text{telomere}} \times R_{\text{bulk}}}{2 \times N_C}.$$

5. Division by  $N_{\text{telomere}}$  converts end counts per bin into end counts per position per genome. The inverse of this yields the estimate of the mean fragment size ( $L_{\text{fragment}}$ ), where the bin is the whole genome, is

$$L_{\text{fragment}} = \frac{N_{\text{telomere}}}{R_{\text{bulk}}},$$

and the mean number of ends per genome ( $N_{\text{ends}}$ ) is

$$N_{\text{ends}} = \frac{2 \times L_{\text{genome}}}{L_{\text{fragment}}}.$$

### 3.13 Predicting Replication Origins from HydEn-Seq Maps

1. Calculate the fraction of replication events ( $F$ ) in which the bottom strand ( $b$ ) is replicated as the leading strand in genomic bin  $i$ . This requires two HydEn-seq data sets with opposite strand biases (*lead* and *lag* below). As per [22],

$$F_{i \text{ lead}, b} = \frac{1}{1 + \sqrt{\frac{O_{i \text{ lag}, b}}{O_{i \text{ lead}, b}}}}$$

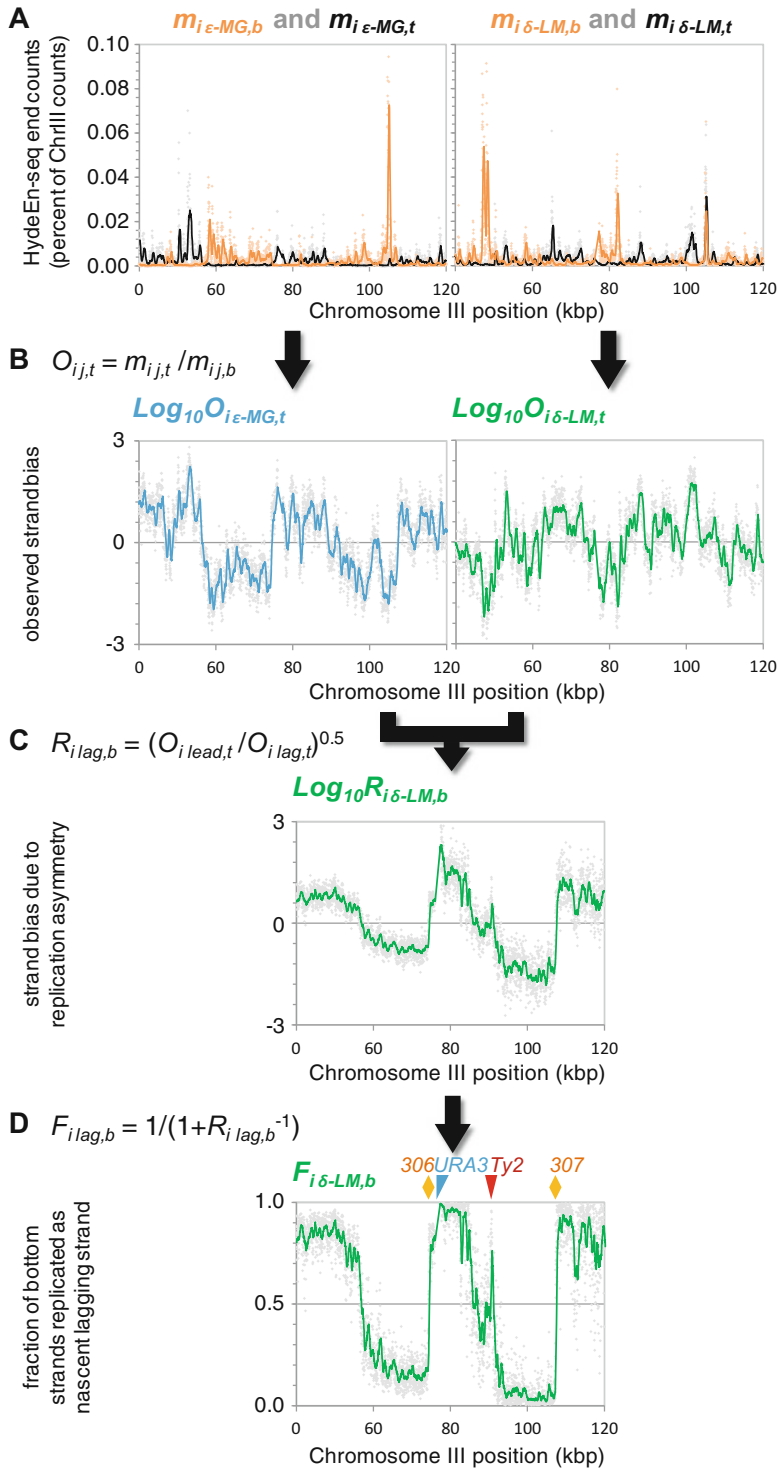
where

$$O_{ij, b} = \frac{m_{ij, b}}{m_{ij, t}}$$

and where  $m_{ij, k}$  is the normalized (Subheading 3.12) and background-subtracted (specific to chosen strains) end count in the  $k$ -strand (*bottom* or *top*) in the  $j$ -biased data set. Estimates improve when using data sets with more extreme strand biases and when using an average of normalized counts ( $\overline{m}_{ij}$ ) from replicate experiments (Fig. 3).

←

**Fig. 2** (continued) Autonomously Replicating Sequence Consensus Sequences (ACS). Reads mapping to the *right* of the ACS on the *top strand* and to the *left* of the ACS on the *bottom strand* are on the leading strand. Reads mapping to the *left* of the ACS on the *top strand* and to the *right* of the ACS on the *bottom strand* are on the lagging strand. The *pol2-M644G rnh201Δ* strain exhibits a bias toward ribonucleotides mapped to the leading strand. The *pol3-L612G rnh201Δ* exhibits a bias toward ribonucleotides mapped to the lagging strand. Wild-type replicase backgrounds are presented with (*RNH201*) and without (*rnh201Δ*) RER. Images were generated with Partek Genomics Suite. **(b)** Meta-analysis of mapped ribonucleotides/million reads 2 kilobases upstream and downstream of 200 ACS. The *red line* is the *top strand* and the *blue line* is the *bottom strand*



**Fig. 3** Calculating the fraction of replication events in which the *bottom strand* is replicated as the leading strand. These panels illustrate the data transformations in Subheading 3.13, **step 1**. (a) Normalized

2. Predicted origins were defined as regions where  $F_{i \text{ lead}, b}$  changed abruptly. Origin calling parameters were set based on the results from a training set of confirmed *Saccharomyces cerevisiae* replication origins and data sets from modified yeast strains (200 bp bins; smoothed over 9 bins) [21]. As such, individual parameters may be more or less suitable for other data sets and especially for other taxa, but general procedures should be widely applicable. In the original yeast experiments, origins were defined as positions where either the average slope (the derivative) of  $F_{i \text{ lead}, b}$  exceeded 0.00011 fractional units per bp in an 11-bin window (2.2 kbp) or 0.00016 fractional units per base pair in at least three of five surrounding bins ( $\geq 600$  bp out of 1 kbp). Based on the training set, these parameters were selected to define a sufficiently abrupt change over a region wide enough to exclude random noise.

---

## 4 Notes

1. The RNase a treatment step is omitted from genomic DNA purification to avoid contaminant nuclease activity in the supplied RNase A. The alkali treatment sufficiently degrades any residual RNA.
2. While the enzymatic steps are performed in PCR tubes or plates, we recommend transferring the reactions to 96-well culture plates for incubation and capturing the magnetic beads with a magnetic ring plate. This will spread the beads over a larger surface area and they will dry faster. We recommend drying at 30 °C, but any temperature between 23 °C and 37 °C will work. It is important not to over dry the beads, as this will reduce yield.
3. Use a unique barcode sequence for each library that will be run on the same lane of the sequencer. We run as many as 24 libraries on each lane, often with six biological replicates of each of two isolates of the same genotype. In this fashion we can generate 12 individual measurements of each of two

---

**Fig. 3** (continued) (Subheading 3.12) end counts ( $m_{ij, k}$ ), by strand (bottom = *b*, orange; top = *t*, black), for two data sets with opposite strand biases (here, leading =  $\epsilon$ -MG, from *pol2-M644G rnh201Δ* yeast; lagging =  $\delta$ -LM, from *pol3-L612 M rnh201Δ* yeast). **(b)** Apparent strand biases ( $O_{ij, k}$ ) are calculated for each data set. **(c)** The apparent biases are combined, removing correlated noise and leaving an estimate of true bias ( $R_{ij, k}$ ). **(d)** This is used to estimate the fraction of replication events in which the *bottom strand* is replicated as the leading strand ( $F_{i \text{ lead}, b}$ ). The abrupt transitions used to predict origins (Subheading 3.13, **step 2**) are obvious at origins *ARS306* and *ARS307* (orange). Predictions may be confounded where strains differ (e.g., *URA3* orientation at the *AGP1* locus; blue) and at certain other features (transposon Ty2; red). End counts averaged over multiple data sets ( $\overline{m_{ij}}$ ) may be used

genotypes on a single run. This is especially useful for establishing statistically significant differences between genotypes.

## Acknowledgment

We thank Dr. Jessica Williams and Dr. Kin Chan for helpful comments on the manuscript. This work was supported by the Division of Intramural Research of the US National Institutes of Health (NIH), National Institute of Environmental Health Sciences (project Z01 ES065070 to T.A.K.).

## References

1. Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundstrom EB, Burgers PM, Johansson E, Chabes A, Kunkel TA (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci U S A* 107(11):4949–4954. doi:[10.1073/pnas.0914857107](https://doi.org/10.1073/pnas.0914857107). 0914857107 [pii]
2. Clausen AR, Zhang S, Burgers PM, Lee MY, Kunkel TA (2013) Ribonucleotide incorporation, proofreading and bypass by human DNA polymerase  $\delta$ . *DNA Repair* 12(2):121–127. doi:[10.1016/j.dnarep.2012.11.006](https://doi.org/10.1016/j.dnarep.2012.11.006). S1568-7864(12)00285-6 [pii]
3. Goksenin AY, Zahurancik W, LeCompte KG, Taggart DJ, Suo Z, Pursell ZF (2012) Human DNA polymerase epsilon is able to efficiently extend from multiple consecutive ribonucleotides. *J Biol Chem* 287(51):42675–42684. doi:[10.1074/jbc.M112.422733](https://doi.org/10.1074/jbc.M112.422733)
4. Williams JS, Lujan SA, Kunkel TA (2016) Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat Rev Mol Cell Biol* 17(6):350–363. doi:[10.1038/nrm.2016.37](https://doi.org/10.1038/nrm.2016.37)
5. Nick McElhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundstrom EB, Johansson E, Chabes A, Kunkel TA (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6(10):774–781. doi:[10.1038/nchembio.424](https://doi.org/10.1038/nchembio.424). nchembio.424 [pii]
6. Sparks JL, Chon H, Cerritelli SM, Kunkel TA, Johansson E, Crouch RJ, Burgers PM (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* 47(6):980–986. doi:[10.1016/j.molcel.2012.06.035](https://doi.org/10.1016/j.molcel.2012.06.035). S1097-2765(12)00599-0 [pii]
7. Lujan SA, Williams JS, Clausen AR, Clark AB, Kunkel TA (2013) Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol Cell* 50(3):437–443. doi:[10.1016/j.molcel.2013.03.017](https://doi.org/10.1016/j.molcel.2013.03.017). S1097-2765(13)00221-9 [pii]
8. Ghodgaonkar MM, Lazzaro F, Olivera-Pimentel M, Artola-Boran M, Cejka P, Reijns MA, Jackson AP, Plevani P, Muzi-Falconi M, Jiricny J (2013) Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell* 50(3):323–332. doi:[10.1016/j.molcel.2013.03.019](https://doi.org/10.1016/j.molcel.2013.03.019). S1097-2765(13)00223-2 [pii]
9. Vengrova S, Dalgaard JZ (2006) The wild-type *Schizosaccharomyces pombe* mat1 imprint consists of two ribonucleotides. *EMBO Rep* 7(1):59–65. doi:[10.1038/sj.embor.7400576](https://doi.org/10.1038/sj.embor.7400576). 7400576 [pii]
10. Caldecott KW (2014) Molecular biology. Ribose—an internal threat to DNA. *Science* 343(6168):260–261. doi:[10.1126/science.1248234](https://doi.org/10.1126/science.1248234)
11. Pizzi S, Sertic S, Orcesi S, Cereda C, Bianchi M, Jackson AP, Lazzaro F, Plevani P, Muzi-Falconi M (2015) Reduction of hRNase H2 activity in Aicardi-Goutieres syndrome cells leads to replication stress and genome instability. *Hum Mol Genet* 24(3):649–658. doi:[10.1093/hmg/ddu485](https://doi.org/10.1093/hmg/ddu485)
12. Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F, Devenney PS, Sexton D, Grimes G, Holt IJ, Hill RE, Taylor MS, Lawson KA, Dorin JR, Jackson AP (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* 149(5):1008–1022. doi:[10.1016/j.cell.2012.04.011](https://doi.org/10.1016/j.cell.2012.04.011). S0092-8674(12)00508-9 [pii]
13. Coffin SR, Hollis T, Perrino FW (2011) Functional consequences of the RNase H2A subunit mutations that cause Aicardi-Goutieres syndrome. *J Biol Chem* 286(19):16984–16991. doi:[10.1074/jbc.M111.228833](https://doi.org/10.1074/jbc.M111.228833)



14. Shaban NM, Harvey S, Perrino FW, Hollis T (2010) The structure of the mammalian RNase H2 complex provides insight into RNA:NA hybrid processing to prevent immune dysfunction. *J Biol Chem* 285(6):3617–3624. doi:10.1074/jbc.M109.059048
15. Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R, Baumann C, Baxter P, Bertini E, Chandler KE, Chitayat D, Cau D, Dery C, Fazzi E, Goizet C, King MD, Klepper J, Lacombe D, Lanzi G, Lyall H, Martinez-Frias ML, Mathieu M, McKeown C, Monier A, Oade Y, Quarrell OW, Rittney CD, Rogers RC, Sanchis A, Stephenson JB, Tacke U, Till M, Tolmie JL, Tomlin P, Voit T, Weschke B, Woods CG, Lebon P, Bonthron DT, Ponting CP, Jackson AP (2006) Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. *Nat Genet* 38(8):910–916. doi:10.1038/ng1842
16. Lujan SA, Williams JS, Pursell ZF, Abdulovic-Cui AA, Clark AB, Nick McElhinny SA, Kunkel TA (2012) Mismatch repair balances leading and lagging strand DNA replication fidelity. *PLoS Genet* 8(10):e1003016. doi:10.1371/journal.pgen.1003016. PGENETICS-D-12-00710 [pii]
17. Miyabe I, Kunkel TA, Carr AM (2011) The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet* 7(12):e1002407. doi:10.1371/journal.pgen.1002407. PGENETICS-D-11-01459 [pii]
18. Williams JS, Smith DJ, Marjavaara L, Lujan SA, Chabes A, Kunkel TA (2013) Topoisomerase I-mediated removal of ribonucleotides from nascent leading-strand DNA. *Mol Cell* 49(5):1010–1015. doi:10.1016/j.molcel.2012.12.021. S1097-2765(13)00033-6 [pii]
19. Lipkin D, Talbert PT, Cohn M (1954) The mechanism of the alkaline hydrolysis of ribonucleic acids. *J Am Chem Soc* 76(11):2871–2872. doi:10.1021/Ja01640a004
20. Li YF, Breaker RR (1999) Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group. *J Am Chem Soc* 121(23):5364–5372. doi:10.1021/Ja990592p
21. Clausen AR, Lujan SA, Burkholder AB, Orebaugh CD, Williams JS, Clausen MF, Malc EP, Mieczkowski PA, Fargo DC, Smith DJ, Kunkel TA (2015) Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol* 22(3):185–191. doi:10.1038/nsmb.2957. nsmb.2957 [pii]
22. Williams JS, Clausen AR, Lujan SA, Marjavaara L, Clark AB, Burgers PM, Chabes A, Kunkel TA (2015) Evidence that processing of ribonucleotides in DNA by topoisomerase I is leading-strand specific. *Nat Struct Mol Biol* 22(4):291–297. doi:10.1038/nsmb.2989. nsmb.2989 [pii]
23. Koh KD, Balachander S, Hesselberth JR, Storici F (2015) Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nat Methods* 12(3):251–257. 253 p following 257. doi: 10.1038/nmeth.3259, nmeth.3259 [pii]

## Detection of DNA-RNA Hybrids In Vivo

María García-Rubio, Sonia I. Barroso, and Andrés Aguilera

### Abstract

DNA-RNA hybrids form naturally during essential cellular functions such as transcription and replication. However, they may be an important source of genome instability, a hallmark of cancer and genetic diseases. Detection of DNA-RNA hybrids in cells is becoming crucial to understand an increasing number of molecular biology processes in genome dynamics and function and to identify new factors and mechanisms responsible for disease in biomedical research. Here, we describe two different procedures for the reliable detection of DNA-RNA hybrids in the yeast *Saccharomyces cerevisiae* and in human cells: DNA-RNA Immunoprecipitation (DRIP) and Immunofluorescence.

**Key words** DNA-RNA hybrids, R loop, Genome instability, S9.6 antibody

---

### 1 Introduction

DNA-RNA hybrids are highly stable structures formed during essential cellular functions such as transcription and replication. Hybrids covering few nucleotides (around 10) are essential for Okazaki fragment synthesis during DNA replication as well as for RNA synthesis by RNA polymerases. In a different context, however, an RNA molecule may invade a double-stranded DNA (dsDNA) to form a DNA-RNA hybrid and the displaced DNA strand. These structures are known as R loops.

R loops have been shown to have a physiological role in initiation of replication of mitochondrial DNA or some bacterial plasmids, as well as in Immunoglobulin class switch recombination, a vital process for the generation of the genetic diversity that the defense systems of vertebrates require, as reviewed in [1]. In these cases, DNA-RNA hybrids may expand from tens of nucleotides to a few hundreds, the S regions involved in Ig class switching having been reported to form hybrids that can exceed even 1 kb in size [2]. However, since a first report showing that yeast cells lacking the

---

María García-Rubio and Sonia I. Barroso contributed equally to this work.



RNA biogenesis and export factor THO accumulate R loops responsible for high levels of transcription-dependent genome instability [3], different contributions have provided evidence that DNA-RNA hybrids or R loops can form at higher levels than previously foreseen, becoming an important threat to genome integrity [1]. The increasing abundance of factors identified to be involved in prevention or resolution of DNA-RNA hybrids have been shown to map all over the genome with particular hotspot regions, and different reports have demonstrated that they may play positive and negative roles in transcription. Importantly, however, it is becoming evident that cells have a large number of proteins involved in the prevention or resolution of R loops.

In most physiological cases, so far, it has not been either reported or explored whether DNA-RNA hybrids require in addition the participation of specific proteins for their formation or stability. In this sense, we know that the action of the CRISPR-Cas9 system of genome editing is based on the formation of a DNA-RNA hybrid indeed [4]. Thus, the possibility that R loops may be facilitated or stabilized by specific proteins [5] needs to be further explored in the future. However, any *in vivo* study on R loops, whether naturally formed or artificially induced, requires to be completed with the *in vivo* detection of DNA-RNA hybrids in cells. This is so far a bottle neck step in all studies. Even though there have been several methods used in different studies, the most extended ones are based on the use of the monoclonal anti-DNA-RNA hybrid antibody of the S9.6 hybridoma [2, 3, 6–13]. Here, we provide the step-by-step protocol of the two more-extended techniques used with this antibody, DNA-RNA immunoprecipitation (DRIP) [14] and immunofluorescence (IF) analyses to detect hybrids both in yeast and in human cells [15, 16]. As can be seen, it is essential in any DNA-RNA hybrid detection techniques to validate that the signals detected correspond unequivocally to DNA-RNA hybrids by eliminating them by treatment with RNase H *in vitro* or by overexpressing RNase H *in vivo*.

---

## 2 Materials

### 2.1 DNA-RNA Immunoprecipitation (DRIP) in Human Cells

1. Cell cultures.
2. 10 cm plates.
3. PBS Tablets (GIBCO).
4. 1× TE buffer: 10 mM Tris-HCl pH 7.6, EDTA 1 mM pH 8.0.
5. 20% SDS.
6. Proteinase K, recombinant (20 mg/ml, Roche).
7. Phenol:Chloroform:Isoamyl Alcohol 25:24:1 Saturated with 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA.

8. Phase Lock Gel Heavy 2 ml tubes (% Prime GmbH).
9. 3 M Sodium Acetate.
10. Isopropanol.
11. Glass rod.
12. 70% Ethanol.
13. Restriction enzymes (*Hind*III 2.5  $\mu$ l, *Eco*RI 2.5  $\mu$ l, *Xba*I 2.5  $\mu$ l, *Ssp*I 10  $\mu$ l, *Bst*GI 5  $\mu$ l), BSA 2.5  $\mu$ l and NEB Buffer 2 (New England BioLabs).
14. StrataClean Resin (Stratagene).
15. 50T5E buffer: 50 mM Tris-HCl pH 7.5, 5 mM EDTA pH 8.0.
16. 50% v/v slurry of Sephadex G-50 (Sigma) in 50T5E buffer (kept at 4 °C).
17. Empty micro Bio-spin chromatography columns (Biorad).
18. RNase H and appropriate 10 $\times$  buffer (New England BioLabs).
19. Fluorometer apparatus and appropriate cuvettes.
20. Appropriate DNA fragments.
21. anti-DNA-RNA hybrid antibody (S9.6).
22. 1 M Na-Phosphate pH 7.0: 39 ml 2 M monobasic sodium phosphate NaH<sub>2</sub>PO<sub>4</sub>, 61 ml 2 M dibasic sodium phosphate Na<sub>2</sub>HPO<sub>4</sub>, 100 ml H<sub>2</sub>O.
23. 10 $\times$  Binding Buffer: 100 mM NaPO<sub>4</sub> pH 7.0, 1.4 M NaCl, 0.5% Triton X-100, stored at room temperature.
24. Elution Buffer: 50 mM Tris pH 8.0, 10 mM EDTA, 0.5% SDS.
25. Protein A-coated magnetic beads (Dynabeads, Novex by Life Technologies).
26. Magnetic rack (Life Technologies).
27. Proteinase K, recombinant (20 mg/ml, Roche).
28. Rotator mixer.
29. Thermomixer.
30. PCR Purification Kit (Quiagen).
31. 1 $\times$  TE buffer: 10 mM Tris-HCl pH 7.6, EDTA 1 mM pH 8.0.
32. Power SYBR green PCR master mix (Applied Biosystems).
33. Appropriate oligonucleotides (*see* **Note 10**).
34. Fast Real-Time PCR System.

## **2.2 DNA-RNA Immunoprecipitation in Yeast Cultures**

1. Cell cultures.
2. Zymolyase 20 T (USB).
3. Spheroplasting Buffer: 1 M sorbitol, 2 mM Tris-HCl pH 8.0, 100 mM EDTA pH 8.0, 0.1% v/v beta-mercapto-ethanol.

4. Solution I: 0.8 mM GuHCl, 30 mM Tris-HCl pH 8.0, 30 mM EDTA pH 8.0, 5% Tween 20, 0.5% Triton X-100.
5. RNase A stock solution: 10 mg/ml (Roche; *see* **Note 11**).
6. Proteinase K, recombinant (20 mg/ml, Roche).
7. Chloroform-Isoamyl alcohol (24:1).
8. Isopropanol (at room temperature).
9. 70% Ethanol (at room temperature).
10. Restriction enzymes (*Hind*III 2.5  $\mu$ l, *Eco*RI 2.5  $\mu$ l, *Xba*I 2.5  $\mu$ l, *Ssp*I 10  $\mu$ l, *Bsr*GI 5  $\mu$ l), BSA 2.5  $\mu$ l and NEB Buffer 2 (New England BioLabs).
11. StrataClean Resin (Stratagene).
12. 50T5E buffer: 50 mM Tris-HCl pH 7.5, 5 mM EDTA pH 8.0.
13. 50% v/v slurry of Sephadex G-50 (Sigma) in 50T5E buffer (kept at 4 °C).
14. Empty micro Bio-spin chromatography columns (Biorad).
15. RNase H and appropriate 10 $\times$  buffer (New England BioLabs).
16. Fluorometer apparatus and appropriate cuvettes.
17. Appropriate DNA fragments.
18. anti-DNA-RNA hybrid antibody (S9.6).
19. 1 M Na-Phosphate pH 7.0 (39 ml 2 M monobasic sodium phosphate NaH<sub>2</sub>PO<sub>4</sub>, 61 ml 2 M dibasic sodium phosphate Na<sub>2</sub>HPO<sub>4</sub>, 100 ml H<sub>2</sub>O).
20. 10 $\times$  Binding Buffer (100 mM NaPO<sub>4</sub> pH 7.0, 1.4 M NaCl, 0.5% Triton X-100, stored at room temperature).
21. Elution Buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5% SDS).
22. Protein A-coated magnetic beads (Dynabeads, Novex by Life Technologies).
23. Magnetic rack (Life Technologies).
24. Proteinase K, recombinant (20 mg/ml, Roche).
25. Rotator mixer.
26. Thermomixer.
27. PCR Purification Kit (Quiagen).
28. 1 $\times$  TE buffer: 10 mM Tris-HCl pH 7.6, EDTA 1 mM pH 8.0.
29. Power SYBR green PCR master mix (Applied Biosystems).
30. Appropriate oligonucleotides (*see* **Note 10**).
31. Fast Real-Time PCR System.

**2.3 S9.6**  
**Immunofluorescence**  
**in Mammalian Cells**

1. DMEM-Dulbecco's Modified Eagle Medium (GIBCO).
2. 24-well plates.
3. Tweezers.
4. Round Coverslips.
5. Bovine serum albumin (BSA).
6. PBS Tablets (GIBCO).
7. Blocking Solution: 2% BSA in PBS (Phosphate-Buffered Saline).
8. anti-DNA-RNA hybrid antibody (S9.6).
9. Nucleolin antibody (Abcam).
10. Ice-cold methanol:  $-20\text{ }^{\circ}\text{C}$  100% methanol (*see Note 16*).
11. Alexa Fluor 488 goat anti-rabbit (Invitrogen).
12. Alexa Fluor 594 goat anti-mouse (Invitrogen).
13. Vacuum line.
14. Immu-Mount Mounting Medium (Thermo Scientific).
15. Microscope Slides.
16. Confocal microscope.
17. DAPI solution (1  $\mu\text{g}/\text{ml}$  in PBS).

**2.4 S9.6**  
**Immunofluorescence**  
**in Yeast**

1. Yeast cultures.
2. Formaldehyde 37%.
3. Phosphate buffer 0.1 M pH 6.4.
4. Sorbitol-citrate 1.2 M: 218.6 g/l sorbitol, 7 g/l citric acid, 17.418 g/l  $\text{K}_2\text{HPO}_4$ .
5. Poly L Lysine Treated Slides.
6. Glusulase (Perkin Elmer).
7. Zymoliase 20T (USBiological): 50 mg/ml in sorbitol 1 M.
8. SDS.
9. PBS-BSA Solution: 1% BSA, 0.04 M  $\text{K}_2\text{HPO}_4$ , 0.01 M  $\text{KH}_2\text{PO}_4$ , 0.15 M NaCl, 0.1% NaN<sub>3</sub>. Filter sterilized and stored at 4  $^{\circ}\text{C}$ .
10. Digestion Mix: 200  $\mu\text{l}$  of 1.2 M sorbitol-citrate, 20  $\mu\text{l}$  of glusulase, 2  $\mu\text{l}$  of zymolyase 20T.
11. Alexa Fluor 594 goat anti-mouse (Invitrogen).
12. anti-DNA-RNA hybrid antibody (S9.6).
13. PBS Tablets (GIBCO).
14. DAPI.
15. Fluorescence microscope.

### 3 Methods

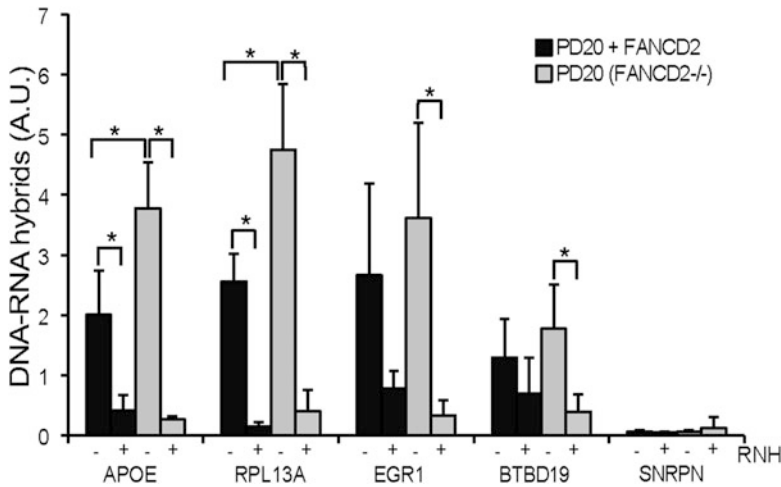
#### 3.1 DNA-RNA Immunoprecipitation in Human Cells

1. Transfer the content of a 10 cm plate of 90% confluent cells ( $3 \times 10^6$  cells) to a 15 ml falcon tube and pellet them at  $400 \times g$ ,  $4^\circ\text{C}$  in a swing rotor for 5 min. Remove the supernatant.
2. Wash and resuspend cells with 10 ml  $1\times$  cold PBS and pellet them again. Remove the supernatant.
3. Resuspend cells in 1.6 ml of TE  $1\times$  and split into two tubes with a cut-off pipette tip (*see Note 1*). Add  $21\ \mu\text{l}$  of 20% SDS, and  $2.5\ \mu\text{l}$  of Proteinase K to each sample (0.8 ml) and invert the tube a few times. Incubate overnight at  $37^\circ\text{C}$  (*see Note 2*).
4. Pour the DNA from the previous step into a phase lock 2 ml tube (*see Note 3*). Add 1 volume (0.8 ml) of Phenol:Chloroform:Isoamyl Alcohol, invert gently a few times, spin down 5 min at  $13,000 \times g$  (*see Note 4*).
5. Pour the DNA (top aqueous phase) in a new phase lock tube and repeat **step 4**.
6. Add  $1/10$  volume 3 M NaOAc ( $160\ \mu\text{l}$ ) and 1 volume (1.6 ml) isopropanol to a 15 ml tube and pour in the DNA (top aqueous phase, rejoin the samples in one tube) from two phase lock tubes.
7. Invert gently until the DNA begins to precipitate. Spool DNA on a glass rod (*see Note 5*).
8. Wash DNA with 70% Ethanol by allowing the EtOH run down glass rod.
9. Allow to air dry. Break off the tip with DNA and put it in a 1.5 ml tube with  $150\ \mu\text{l}$  TE  $1\times$ . Do not resuspend the DNA by over-pipeting/vortexing. Mix several times by gentle shaking (*see Note 6*).
10. Digest the DNA overnight using cocktail restriction enzymes according to supplier's instructions (add BSA  $1\times$ , final volume is  $250\ \mu\text{l}$ ) (*see Note 7*).
11. Add  $3\ \mu\text{l}$  StrataClean Resin, pipet with cut tips, and spin down at full speed in a microcentrifuge for 1 min (*see Note 8*).
12. Purify over Sephadex g-50 column (in 50T5E buffer) (*see Note 9*).
13. Split the flow-through containing DNA into two new tubes ( $125\ \mu\text{l}$ ).
14. Treat the half of the DNA (one tube) with  $5\ \mu\text{l}$  RNase H (NEB) overnight at  $37^\circ\text{C}$ . In the meantime, keep the other half at  $4^\circ\text{C}$ .
15. Repeat **step 12**.

16. Check DNA concentration (samples treated and untreated with RNase H).
17. Take 1  $\mu\text{g}$  of digested DNA (treated and untreated with RNase H) (INPUT) and process immediately as described in **steps 25** and **26**.
18. Dilute 5  $\mu\text{g}$  of digested DNA (treated and untreated with RNase H) in 450  $\mu\text{l}$  TE1 $\times$ .
19. Add 51  $\mu\text{l}$  of 10 $\times$  Binding Buffer.
20. Add 10  $\mu\text{l}$  of anti-DNA-RNA hybrid antibody (Stock 1 mg/ml).
21. Incubate overnight at 4  $^{\circ}\text{C}$  on a rotator mixer.
22. Wash 40  $\mu\text{l}$  of Protein A magnetic beads per DRIP sample, twice with 1 ml of 1 $\times$  Binding Buffer (10 $\times$  diluted in TE1 $\times$ ). Resuspend in the original volume (40  $\mu\text{l}$ ).
23. Add the prepared magnetic beads to the extracts (40  $\mu\text{l}$  per reaction) and incubate for 2 h at 4  $^{\circ}\text{C}$  on a rotator mixer.
24. Place the precooled magnetic rack on ice, insert the tubes, aspirate all liquid off, and resuspend the beads in 1 ml of 1 $\times$  Binding Buffer. Repeat the same procedure to achieve three washes with 1 $\times$  Binding Buffer.
25. Add 120  $\mu\text{l}$  Elution Buffer and 7  $\mu\text{l}$  Proteinase K and incubate at 55  $^{\circ}\text{C}$  for 45 min in a shaker. Dilute 1  $\mu\text{g}$  of INPUT samples (**step 17**) in 47  $\mu\text{l}$  of TE1 $\times$  and add 3  $\mu\text{l}$  of Proteinase K, incubate at 55  $^{\circ}\text{C}$  for 45 min.
26. Purify (INPUT and PRECIPITATE) over PCR purification columns (e.g., from Quiagen) according to manufacturer's protocol and elute in 150  $\mu\text{l}$  TE1 $\times$ . Store samples at  $-20^{\circ}\text{C}$ .
27. Quantify the enriched DNA fragments in the PRECIPITATE by real-time quantitative PCR (qPCR). We perform qPCR using Fast SYBR Green dye and the absolute quantification protocol in the 7500 Fast Real-Time PCR System (Applied Biosystems). Standard curves for all pairs of primers are performed for each analysis. All PCR reactions are performed in triplicate. The enrichment for each PCR of interest is normalized with respect to the corresponding ratios of the INPUT (*see* Fig. 1).

### **3.2 DNA-RNA Immunoprecipitation in Yeast Cultures**

1. Grow 100 ml yeast cultures in the appropriate medium and temperature to an absorbance at 600 nm of about 0.8.
2. Harvest cells by centrifugation at  $4000 \times g$ , 4  $^{\circ}\text{C}$  for 5 min.
3. Wash cells twice with 20 ml cold  $\text{H}_2\text{O}$ .
4. Harvest cells by centrifugation at  $4000 \times g$ , 4  $^{\circ}\text{C}$  for 5 min.



**Fig. 1** Relative amount of R loops in the patient  $FANCD2^{-/-}$  human PD20 cell line and the corrected PD20  $FANCD2^{+/+}$  control at four different genes. Results shown correspond to DRIP-qPCR assays using the S9.6 monoclonal antibody with and without RNase H (RNH) treatment. Signal values of DNA-RNA hybrids immunoprecipitated in each region are normalized to input values. *A.U.* arbitrary units. (Reproduced from [16] with modifications)

5. Resuspend the cell pellet in 2.4 ml Spheroplasting Buffer (freshly made) (add 2 mg Zymolyase to 2.4 Spheroplasting Buffer). Split the sample into two tubes (*see Note 1*).
6. Incubate at 30 °C for 30 min with rotation.
7. Spin down for 5 min at 4500 × *g* in a microcentrifuge and discard the supernatant. Remove all the supernatant.
8. Carefully wash the spheroplast pellet by adding 500 µl of cold H<sub>2</sub>O without resuspending.
9. Spin briefly and discard the water.
10. Break the spheroplasts by resuspending in 1.125 ml Solution I (565 µl per tube, rejoin the samples) and transferring in 2 ml eppendorf tubes (*see Note 12*).
11. Add 20 µl RNase A.
12. Incubate for 30 min at 37 °C.
13. Add 75 µl Proteinase K (20 mg/ml) and incubate at 50 °C for 60 min (*see Note 13*).
14. Spin down for 5 min at 4500 × *g* and transfer the supernatant to a new tube.
15. Add 800 µl Chloroform-Isoamyl alcohol.
16. Spin down for 5 min at 4500 × *g*.
17. Carefully collect the upper phase with 1 ml cut tips and transfer into a 2 ml tube.

18. Add 800  $\mu$ l isopropanol to the aqueous phase and invert several times until genomic DNA precipitates. Spool DNA on a glass rod (*see Note 5*).
19. Wash DNA with 70% Ethanol by allowing the EtOH run down glass rod.
20. Allow to air dry, break off the tip with DNA, and put it in a 1.5 ml tube with 150  $\mu$ l TE 1 $\times$ . Do not resuspend the DNA by over-pipetting/vortexing. Mix several times by gentle shaking (*see Note 6*).
21. Digest the DNA overnight using cocktail restriction enzymes according to supplier's instructions (add 1 $\times$  BSA, final volume is 250  $\mu$ l) (*see Note 7*).
22. Add 3  $\mu$ l StrataClean Resin, pipet with cut tips and spin down at full speed in a microcentrifuge for 1 min (*see Note 8*).
23. Purify over Sephadex g-50 column (in 50T5E buffer) (*see Note 9*).
24. Split the flow-through containing DNA into two new tubes (125  $\mu$ l).
25. Treat half of the DNA (one tube) with 8  $\mu$ l RNase H (NEB) overnight at 37  $^{\circ}$ C. In the meantime, keep the other half at 4  $^{\circ}$ C.
26. Repeat **step 12**.
27. Take 20  $\mu$ l of digested DNA (treated and untreated with RNase H) (INPUT) and process immediately as described in **step 35**.
28. Dilute digested DNA (treated and untreated with RNase H) in 450  $\mu$ l TE1 $\times$ .
29. Add 51  $\mu$ l of 10 $\times$  Binding Buffer.
30. Add 10  $\mu$ l of anti-DNA-RNA hybrid antibody (Stock 1 mg/ml).
31. Incubate overnight at 4  $^{\circ}$ C on a rotator mixer.
32. Wash 40  $\mu$ l of Protein A magnetic beads per DRIP sample, twice with 1 ml of 1 $\times$  Binding Buffer (10 $\times$  diluted in TE1 $\times$ ). Resuspend in the original volume (40  $\mu$ l).
33. Add the prepared magnetic beads to the extracts (40  $\mu$ l per reaction) and incubate for 2 h at 4  $^{\circ}$ C on a rotator mixer.
34. Place the precooled magnetic rack on ice, insert the tubes, aspirate all liquid off, and resuspend the beads in 1 ml of 1 $\times$  Binding Buffer. Repeat the same procedure to complete three washes with 1 $\times$  Binding Buffer.
35. Add 120  $\mu$ l Elution Buffer and 7  $\mu$ l Proteinase K and incubate at 55  $^{\circ}$ C for 45 min in a shaker. Add 27  $\mu$ l TE1 $\times$  to INPUT

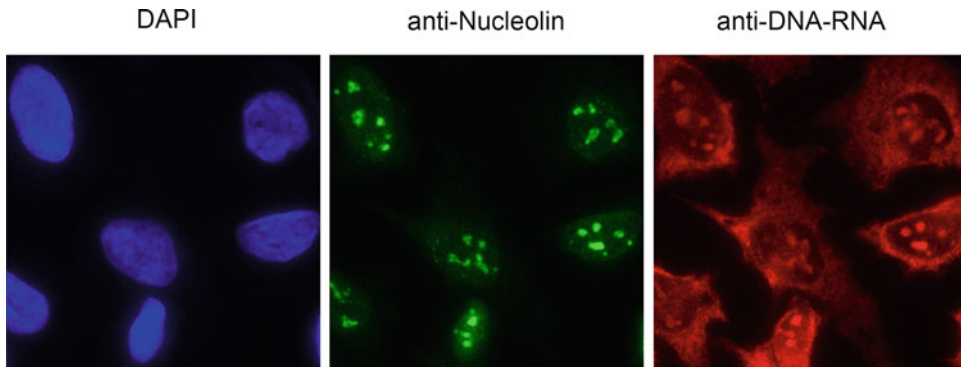


samples (**step 27**), 3  $\mu$ l Proteinase K and incubate at 55 °C for 45 min.

36. Purify INPUT and PRECIPITATE over PCR purification columns according to manufacturer's protocol and elute in 150  $\mu$ l TE1 $\times$ . Store samples at -20 °C.
37. Quantify the enriched DNA fragments in the PRECIPITATE by real-time quantitative PCR (qPCR). We perform qPCR using Fast SYBR Green dye and the absolute quantification protocol in the 7500 Fast Real-Time PCR system (Applied Biosystems). Standard curves for all pairs of primers are performed for each analysis. All the PCR reactions are performed in triplicated. The enrichment for each PCR of interest is normalized with respect to the corresponding ratios of the INPUT.

### **3.3 S9.6** **Immunofluorescence** **in Mammalian Cells**

1. HeLa cells are cultured on coverslips (*see Note 14*) at a concentration of  $2 \times 10^5$  cells/well in a 6-well plate.
2. After 24 h, coverslips are transferred to a 24-well plate well with cold PBS (*see Note 15*).
3. Aspirate the PBS with the vacuum system and add 1 ml of 100% ice-cold methanol (*see Note 16*). Incubate for 7 min at -20 °C.
4. Remove methanol (*see Note 16*) and wash twice with 1 ml/well of PBS.
5. Incubate the cells with Blocking Solution (1 ml/well) and incubate overnight at 4 °C.
6. Remove Blocking Solution and add anti-DNA-RNA hybrid (1:500 in Blocking Solution) and Nucleolin (1:1000 in Blocking Solution) antibodies in a total volume of 250  $\mu$ l/well. Incubate overnight at 4 °C.
7. Wash three times in PBS (5 min each) and incubate with Alexa Fluor 488 goat anti-rabbit and Alexa Fluor 594 goat anti-mouse secondary antibodies (1:1000 in Blocking Solution) 1 h at room temperature.
8. Wash twice for 5 min each in PBS. Incubate with DAPI solution for 5 min at room temperature and wash twice more for 5 min each in PBS. Wash once in distilled H<sub>2</sub>O.
9. Coverslips are removed using tweezers and placed on a microscope slide with Immumount mounting medium (*see Note 17*).
10. The slides are placed flat at room temperature for 24 h and then stored at 4 °C (*see Fig. 2*).



**Fig. 2** Representative image of HeLa cells stained with DAPI (DNA), S9.6 (DNA-RNA hybrids), and anti-nucleolin (nucleoli) antibodies. DNA-RNA hybrids can be observed at mitochondria, nucleoli, and nuclear DNA

### 3.4 S9.6 Immunofluorescence in Yeast

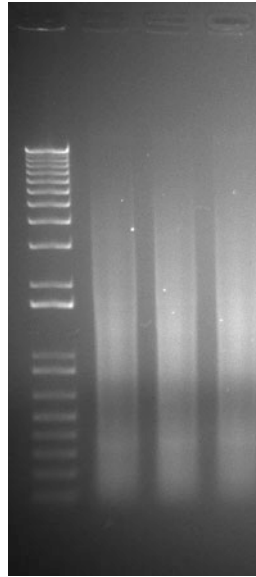
1. Collect at least 1 ml of mitotic culture.
2. Add 100  $\mu$ l of 37% formaldehyde to the media and vortex.
3. Spin the culture for 1 min at  $900 \times g$  and resuspend in 3.7% formaldehyde (*see Note 18*). Fix 10 min at room temperature.
4. Wash twice in 1 ml of 0.1 M Phosphate Buffer pH 6.4.
5. Resuspend in 1 ml of 1.2 M sorbitol-citrate (*see Note 19*).
6. Spin down and resuspend in digestion mix. Rotate on rack at  $30^\circ\text{C}$  for 5 min (*see Note 20*).
7. Spin down at  $900 \times g$  for 2 min. Remove the supernatant and resuspend in 1 ml of sorbitol by inversion (*see Note 21*).
8. Spin down at  $900 \times g$  for 3 min. Remove the supernatant and resuspend in sorbitol by inversion (*see Note 21*).
9. Clean the slide with a scrubber under distilled water (*see Note 22*). Let it dry completely.
10. Place 5  $\mu$ l of 0.1% Poly-L-lysine on each well and let it sit for 5 min (*see Note 23*).
11. Wash the slide under distilled water and let it dry completely.
12. Prepare a humid chamber by wrapping a petri dish in aluminum foil and put a wet paper towel inside.
13. In order to make the cells burst, drop 5  $\mu$ l of the cells onto the slide from a height of 30–40 cm (*see Note 24*). Let cells sit for 10 min.
14. Remove the supernatant by placing the vacuum tip to the side of the well (*see Note 23*).
15. Put slide in  $4^\circ\text{C}$  methanol (*see Note 16*) for 3 min followed by 10 s in  $4^\circ\text{C}$  acetone. Let it dry completely.
16. Add 5  $\mu$ l of primary antibody (anti-DNA-RNA hybrid 1:200 in PBS/BSA). Incubate in a wet chamber for 2 h at room temperature.

17. Remove the supernatant and wash five times in PBS/BSA (*see Note 25*).
18. Add 5  $\mu\text{l}$  of secondary antibody (anti-mouse Alexa Fluor 546). Incubate for 2 h at room temperature.
19. Remove the supernatant and wash five times in PBS/BSA.
20. Add 3  $\mu\text{l}$  of mounting medium with DAPI to each sample. Put on a coverslip. Paint with nail polish around the edges of the coverslip and let it dry for 10 min.

---

## 4 Notes

1. Always pipet DNA with cut tips to preserve DNA-RNA hybrids. Do not attempt to resuspend DNA by over-pipetting/vortexing.
2. The solution becomes viscous.
3. This step is extremely important. Phase lock gel heavy 2 ml tubes (5 PRIME GmbH, Hamburg, Deutschland) can be used to facilitate sample phenol treatment.
4. Wear gloves and a lab coat while working with phenol, because it is a very dangerous compound. Do it in the hood. Normally, the aqueous phase forms the upper phase. You should always dispose of phenol waste in a specially sealed container and ensure that it is eliminated according to your institution's policies for dangerous waste.
5. Use a glass Pasteur pipette to spool the DNA. Do not spin to precipitate the DNA. Centrifugation results in high loads of RNA contamination.
6. Do not attempt to resuspend the DNA by vortexing, eventually the DNA could be resuspended at 30 °C (1–2 h). Invert the tube occasionally.
7. Use Buffer 2.1 (NEB) for the restriction. You do not need to remove the cut glass rod. We often run 4  $\mu\text{l}$  on gel the next day to ensure the digestion is complete and the samples are free of RNA (*see Fig. 3*)
8. Better processing groups of two or three samples.
9. To prepare Sephadex G-50 columns, place empty micro Biospin chromatography columns in 2 ml eppendorf tubes, add 800  $\mu\text{l}$  of Sephadex G-50 slurry and centrifuge at  $2000 \times g$  for 2 min in a microcentrifuge. The Sephadex G-50 columns are then placed in new 1.5 ml tubes, which serve to collect the purified DNA after centrifugation at  $2000 \times g$  for 2 min in a microcentrifuge.



**Fig. 3** Genomic DNA digested with a mix of restriction enzymes for DRIP analysis. 4  $\mu$ l of each digestion reaction were run on agarose gel to confirm that they are free of RNA

10. We use the Primer Express application software (Applied Biosystems) to design the oligonucleotides.
11. The RNase A should be carefully boiled to eliminate residual DNase activity.
12. Spheroplasts are very difficult to resuspend at this step. Use 1 ml cut tips.
13. At this point if there are not many clumps in the samples you can proceed further with the extraction. If the preps are “dirty” you can incubate overnight at 30 °C.
14. Use tweezers to place two to three round coverslips on a well before seeding the cells.
15. Fill the 24-well plate with 1 ml of PBS per well and place one coverslip per well with the cells on the upper side.
16. Methanol is a hazardous chemical with significant toxic, flammable, and reactive properties that can produce deleterious impacts on human health and the environment when not properly handled. Wear lab coat and gloves when working with methanol and manipulate it always in a chemical fume hood.
17. Place a drop of about 30  $\mu$ l of mounting medium on a slide and place the coverslip with the cells facing the mounting medium.
18. Formaldehyde is a sensitizing agent and a cancer hazard. Wear gloves and lab coat and work always in a chemical fume hood.

19. At this step, samples can be stored at  $-20^{\circ}\text{C}$  overnight.
20. Do not vortex.
21. Check digestion by mixing equal amounts of cells and 1% SDS. You should see crystals at the microscope.
22. It is very important for cells to burst properly that the slides are really clean.
23. Do not touch the well with the tip.
24. You can stand on a chair to get a good height to drop the cells.
25. Lay down a drop and remove it with a vacuum line.

---

## Acknowledgments

A.A.'s lab is funded by Spanish Ministry of Economy and Competitiveness, Junta de Andalucía, European Research Council, the Worldwide Cancer Research and the European Union (FEDER).

## References

1. Aguilera A, Garcia-Muse T (2012) R loops: from transcription byproducts to threats to genome stability. *Mol Cell* 46(2):115–124
2. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* 4(5):442–451
3. Huertas P, Aguilera A (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell* 12(3):711–721
4. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–821
5. Gavalda S, Santos-Pereira JM, Garcia-Rubio ML, Luna R, Aguilera A (2016) Excess of Yra1 RNA-binding factor causes transcription-dependent genome instability, replication impairment and telomere shortening. *PLoS Genet* 12(4):e1005966
6. Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, Carrico RJ (1986) Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J Immunol Methods* 89(1):123–130
7. El Hage A, French SL, Beyer AL, Tollervey D (2010) Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* 24(14):1546–1558
8. Pohjoismäki JLO, Holmes JB, Wood SR, Yang M-Y, Yasukawa T, Reyes A, Bailey LJ, Cluett TJ, Goffart S, Willcox S, Rigby RE, Jackson AP, Spelbrink JN, Griffith JD, Crouch RJ, Jacobs HT, Holt IJ (2010) Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of RNA/DNA hybrid. *J Mol Biol* 397(5):1144–1155
9. Mischo HE, Gómez-González B, Grzechnik P, Rondón AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ (2011) Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell* 41(1):21–32
10. Wahba L, Amon JD, Koshland D, Vuica-Ross M (2011) RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol Cell* 44(6):978–988
11. Stirling PC, Chan YA, Minaker SW, Aristizabal MJ, Barrett I, Sipahimalani P, Kobor MS, Hieter P (2012) R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev* 26(2):163–175
12. Gómez-González B, Aguilera A (2007) Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proc Natl Acad Sci U S A* 104(20):8409–8414

13. Wahba L, Costantino L, Tan FJ, Zimmer A, Koshland D (2016) S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev* 30(11):1327–1338
14. Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* 45(6):814–825
15. Sollier J, Stork CT, García-Rubio ML, Paulsen RD, Aguilera A, Cimprich KA (2014) Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell* 56(6):777–785
16. Garcia-Rubio ML, Perez-Calero C, Barroso SI, Tumini E, Herrera-Moyano E, Rosado IV, Aguilera A (2015) The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet* 11(11):e1005674

## Analysis of De Novo Telomere Addition by Southern Blot

Diego Bonetti and Maria Pia Longhese

### Abstract

Telomere length is maintained in most eukaryotes by the action of a specialized enzyme, the telomerase. However, the complexity of mechanisms regulating telomeric DNA length as well as the heterogeneity in length of each telomere in a population of cells has made it very difficult to understand how telomerase is regulated in vivo. Here, we describe a method developed in *Saccharomyces cerevisiae* to monitor the addition of telomeric sequences to a single newly generated telomere in vivo. The primary strain consists of a HO endonuclease cleavage site that is placed directly adjacent to an 81-base-pair stretch of telomeric DNA inserted into the *ADH4* locus of chromosome VII. Upon cleavage by HO, the de novo DNA end is rapidly healed by the telomerase enzyme and the analysis of this process allows to gain a mechanistic understanding of how telomerase action is regulated in the cell.

**Key words** De novo telomere, HO endonuclease, *S. cerevisiae*, Southern blot, Telomerase

---

### 1 Introduction

Telomeres are specialized nucleoprotein complexes that maintain the integrity of eukaryotic chromosomes. They are believed to provide a “cap” to the end of the chromosome, thus preventing fusions between chromosomes and concomitantly inhibiting activities that could either degrade chromosomes or inappropriately use them as substrates for recombination events [1, 2].

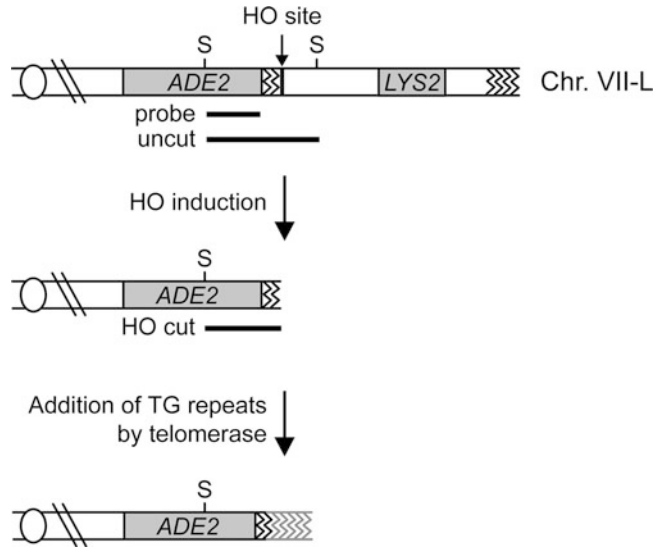
Telomeric DNA consists in most species of a simple tandemly repeated sequence that is guanine-rich in the strand running 5'-3' from the centromere towards the chromosome end (e.g., T<sub>2</sub>AG<sub>3</sub> in humans; T<sub>2</sub>G<sub>4</sub> in *Tetrahymena*; TG<sub>1-3</sub> in *Saccharomyces cerevisiae*). In most eukaryotes, a specialized enzyme called telomerase adds telomeric repeats to the 3' end of a telomeric G-rich strand. This process represents the main pathway for maintaining telomeric DNA repeats at the end of the chromosome [1, 2]. Importantly, work in *S. cerevisiae* has shown that the telomerase does not act on every telomere in each cell cycle, but preferentially extends the shortest telomeres [3–5].

Telomeric DNA repeat length has generally served as a readout for telomerase activity *in vivo*, because this enzyme is required to maintain a particular steady-state length of telomeric DNA at the end of the chromosome. For example, in most yeast *S. cerevisiae* strains the length of TG<sub>1-3</sub> repeats at the end of the chromosomes varies between 225 and 375 base-pair (bp) [1, 2]. However, the length heterogeneity at a given telomere in a population of cells and the complexity of telomeric DNA length regulation have made it difficult to understand how telomerase is regulated *in vivo*.

To circumvent this problem, Gottschling and colleagues have developed a *Saccharomyces cerevisiae* system that allows to monitor *in vivo* the addition of telomeric DNA sequences by telomerase onto a single *de novo* created telomere [6, 7]. This system was based on a previous finding that a double-strand break generated by the HO endonuclease can be healed by the telomerase enzyme at a low frequency when stretches of the T<sub>2</sub>G<sub>4</sub> repeat are present near the break site [8]. Gottschling and colleagues made the healing more efficient by constructing a haploid strain carrying the HO recognition sequence immediately adjacent to an 81 bp telomeric “seed” sequence of C<sub>1-3</sub>A/TG<sub>1-3</sub> repeats. Upon HO cleavage, the new DNA end is rapidly elongated in a telomerase-dependent manner to reach a steady-state length and this elongation is dependent upon most of the telomere regulating factors identified so far [6, 7]. In more detail, the *ADH4* locus on chromosome VII in strain UCC5913 is replaced with a 6 kb fragment consisting of the *S. cerevisiae* *ADE2* gene, 81 bp of TG<sub>1-3</sub> telomeric sequences and the recognition site for the HO endonuclease (Fig. 1). Since a recognition site for this endonuclease is normally present at the mating type locus (*MAT*), the UCC5913 strain carries a mutation that makes this site uncleavable (*MAT-inc*). Moreover, HO expression is under the control of the inducible *GALI* promoter, and the *GALI-HO* construct is integrated at the *LEU2* locus. As transcription from *GAL* promoter is repressed by glucose, cells are grown in raffinose-containing media followed by galactose addition to induce HO expression. After *GALI-HO* induction, telomeric DNA addition to the newly generated telomere end can be monitored by Southern blot analysis with an *ADE2* specific probe. It is important to mention that the distal part of chromosome VII containing the native telomere (ca. 20 kb) is lost at the first cell divisions after the HO cut (Fig. 1), but no essential genes are present in this region. Figure 2 is a representative Southern blot showing how this system allows to monitor telomere addition in wild type cells as well as in *sae2Δ* mutant cells that are defective in this process.

Here, we describe a protocol to monitor *de novo* telomeric DNA addition by telomerase at a single telomere in haploid yeast cells carrying the HO system described above. We divide it in three main steps: (i) generation of the *de novo* telomere; (ii) genomic DNA extraction; and (iii) DNA electrophoresis and Southern blot analysis.





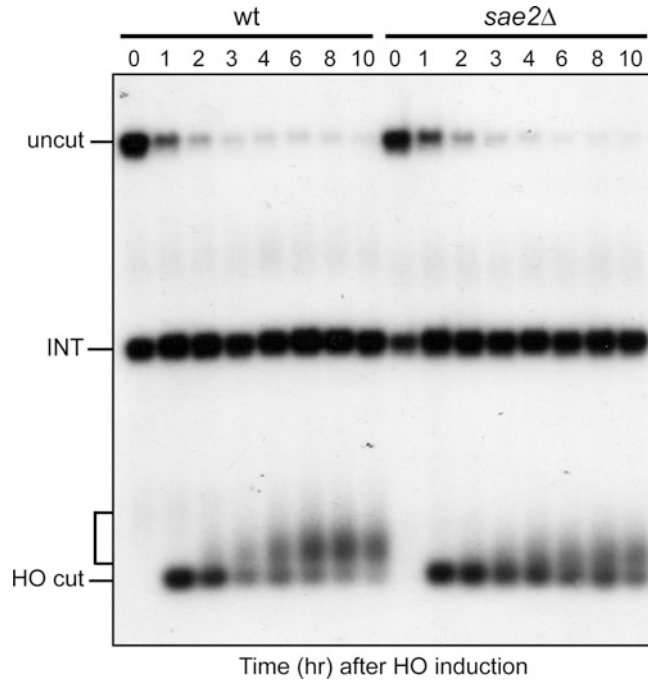
**Fig. 1** System to detect de novo telomere addition. In UCC5913 strain [*MATa-inc ade2-101 lys2-801 his3-Δ200 trp1-Δ63 ura3-52 leu2-Δ1::GAL1-HO-LEU2 VII-L::ADE2-TG<sub>(1-3)</sub> -HO site-LYS2*], the *ADH4* locus on chromosome VII is replaced with a 6 kb fragment consisting of the *S. cerevisiae ADE2* gene and 81 bp of TG<sub>1-3</sub> telomeric sequences (zigzag lines) flanking the recognition site for the HO endonuclease. SpeI cuts 700 bp upstream of the *ADE2* ORF starting codon (ATG to TAA is from TG repeats to centromere direction). A wild-type copy of the *LYS2* gene is present on the distal side of the HO cut site. This region of about 20 kb does not contain essential genes and is lost upon galactose induction. The probe used for Southern blot analysis is a 560 bp fragment of the *ADE2* gene located upstream of the ATG (–720 bp to –1280 bp). Telomere addition is indicated as *grey zigzag lines*

## 2 Materials

All solutions are prepared using filtered and deionized ultrapure water (ddH<sub>2</sub>O; resistivity 18.2 MΩ cm at 25 °C) and analytical grade reagents. Prepare and store all solutions at room temperature (unless otherwise indicated).

### 2.1 HO Endonuclease Induction in Yeast Cells

1. YEP medium: 2% Bacto peptone, 1% yeast extract, 0.0025% adenine hemisulfate salt. Dissolve in ddH<sub>2</sub>O. Autoclave.
2. 30% raffinose: Dissolve D-(+)-raffinose pentahydrate in ddH<sub>2</sub>O. Autoclave or sterilize by filtration. Add sterilized raffinose to YEP medium to 2% final concentration.
3. 30% galactose: Dissolve D-(+)-galactose ≥99.0% in ddH<sub>2</sub>O. Sterilize by filtration only (*see Note 1*).



**Fig. 2** Southern blot analysis to detect de novo telomere addition. Wild type and *sae2Δ* cells, exponentially growing in raffinose, were shifted to galactose containing medium at time zero to induce HO expression. Genomic DNA prepared at the indicated times after galactose addition was digested with *SpeI* and separated on an 0.8% agarose gel by about 16 h run at 2 V/cm. DNA fragments were transferred onto a nylon membrane and hybridized with the *ADE2* specific probe described in Fig. 1. A band of about 3 kb (uncut) can be detected in the absence of galactose. This band is converted by HO cleavage into a 0.7 kb fragment (HO-cut). A *bracket* points out new telomere repeats added to the exposed TG<sub>1-3</sub> telomeric sequences. The band of about 1.6 kb (INT) represents the endogenous *ade2-101* gene that can be used as a loading control

4. SD -Lys plates: Dissolve 6.7 g/L of yeast nitrogen base without amino acids, 20 mg/L uracil, adenine, histidine, tryptophan, leucine, arginine and methionine, 50 mg/L threonine, 60 mg/L phenylalanine and isoleucine, 2% D-(+)-glucose monohydrate. Dissolve in ddH<sub>2</sub>O. Add 2% Bacto agar. Autoclave, cool down and pour plates.

## 2.2 Genomic DNA Extraction

1. Spheroplasting solution: 0.9 M sorbitol, 0.1 M ethylenediaminetetraacetic acid (EDTA), pH 7.5.
2. Zymolyase solution: Dissolve 2 mg/mL Zymolyase 20T<sup>®</sup> (see **Note 2**) from *Arthrobacter luteus* (Nacalai Tesque) in spheroplasting solution +14 mM β-mercaptoethanol.
3. 1× TE: 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, pH 7.4. Autoclave.

4. Lysis solution: 2.2% sodium dodecyl sulfate (SDS), 278 mM EDTA, pH 8.5, 445 mM Tris base. Prepare the lysis solution just before use by mixing the appropriate amounts of 10% SDS, 0.5 M EDTA, pH 8.5, 2 M Tris base stock solutions.
5. 5 M potassium acetate: Dissolve 5 M potassium acetate in ddH<sub>2</sub>O. Autoclave.
6. Ice-cold ethanol: 96% and 70%. Prepare a 70% ethanol solution by diluting 96% ethanol in ddH<sub>2</sub>O. Store at -20 °C aliquots of both 70% ethanol and 95% ethanol in glass bottles.
7. RNase A solution: Dissolve 10 mg/mL RNase A, DNase-free, in 10 mL 10 mM Tris-HCl, pH 7.5, 15 mM NaCl. Heat to 100 °C for 5 min and cool down at room temperature. Prepare small aliquots and store them at -20 °C.
8. 2-propanol anhydrous, 99.5%.
9. 6× DNA loading buffer: 30% glycerol, 0.25% bromophenol blue.
10. 1× TAE buffer: 40 mM Tris-acetate, 10 mM EDTA. For 50× TAE buffer stock: for 1 L dissolve 242 g of Tris base in approximately 600 mL ddH<sub>2</sub>O. Add 57.1 mL glacial acetic acid and 100 mL 0.5 M EDTA, and bring final volume to 1 L with ddH<sub>2</sub>O. Before use, dilute in ddH<sub>2</sub>O to a final concentration of 1×.
11. Agarose gel: melt 0.8% agarose in 1× TAE buffer. Cool at approximately 60 °C and add 10 mg/mL ethidium bromide solution to a final concentration of 1 µg/mL. Pour the gel into a gel tank and insert a comb.
12. Ethidium bromide solution: prepare a stock of 10 mg/mL ethidium bromide in ddH<sub>2</sub>O and store in light-protected containers (*see Note 3*).
13. UV lamp with camera.

## **2.3 Native Gel Electrophoresis and Southern Blot**

### *2.3.1 DNA Digestion*

### *2.3.2 Native Gel Electrophoresis and Transfer*

1. SpeI restriction enzyme (20,000 U/mL; New England Biolabs) and 10× buffer supplied from the distributor.
1. Horizontal electrophoresis system with a large gel running chamber (gel size 25 × 20 cm) and standard 32-tooth comb (thickness 1.0 mm and width of teeth 4.0 mm).
2. 1× TAE buffer.
3. Denaturing solution: 0.2 N NaOH, 0.6 M NaCl. Dissolve in ddH<sub>2</sub>O just before use.
4. Neutralizing solution: 1 M Trizma base, 1.5 M NaCl. Dissolve in ddH<sub>2</sub>O just before use. Adjust pH to 7.4 with HCl.

5. Nylon hybridization transfer membrane (e.g., GeneScreen<sup>®</sup> Plus from Perkin Elmer).
6. 20× SSC buffer: 3 M NaCl, 300 mM sodium citrate. Adjust the pH to 7.0 with HCl. Autoclave.
7. Whatman<sup>™</sup> 3 MM paper.
8. Paper towels.
9. UV cross-linker (e.g., Stratalinker<sup>®</sup> from Stratagene).

### 2.3.3 Probe Labeling

1. DNA template. dsDNA for random priming labeling is obtained by PCR using yeast genomic DNA as a template (e.g., UCC5913 DNA prepared for Southern blot analysis) and oligos *ADE2-3'* 5'-ATTTACAGTTTTGATATCCTTGGC-3' and *ADE2-5'* 5'-TTCTAATGTAGATTCTTGTGTTGTCG-3' to amplify a 560 bp region at the *ADE2* locus.
2. Gel extraction kit (e.g., QIAgen).
3. Random priming labeling kit (e.g., Decaprime<sup>™</sup> II kit from Ambion<sup>®</sup>)
4. dATP- $\alpha^{32}\text{P}$  or dCTP- $\alpha^{32}\text{P}$ . Specific activity: 3000 Ci/mmol.

### 2.3.4 Filter Hybridization

1. Hybridization oven and hybridization tubes.
2. Hybridization solution: 0.5 M NaPO<sub>4</sub> pH 7.2, 1 mM EDTA pH 7.5, 7% SDS, 1% BSA. Prepare fresh in sterile ddH<sub>2</sub>O.
3. Washing solution: 0.2 M NaPO<sub>4</sub> pH 7.2, 1% SDS. Prepare fresh in sterile ddH<sub>2</sub>O.
4. Autoradiography cassette with intensifying screens.
5. Autoradiography films or imaging plates.

---

## 3 Methods

### 3.1 De Novo Telomere Generation

1. Streak out the strains to be analyzed on SD -Lys + 2% glucose plates and grow them for 2 days at 25–30 °C (*see Note 4*).
2. Inoculate cells in an appropriate volume of YEP + raffinose (2%) medium (*see Note 5*).
3. Grow the cell culture overnight at 25–30 °C.
4. Next day, when the cell culture has grown up to  $8 \times 10^6$  cells/mL (OD<sub>600nm</sub> 0.5) in an appropriate volume (*see Note 5*), take first 50 mL of the cell culture for genomic DNA extraction (uncut control) and then add galactose to 2% final concentration. If telomere addition has to be analyzed in a specific cell cycle phase (e.g., G<sub>2</sub>/M), first of all synchronize the cell culture (*see Note 6*).
5. Grow the cell culture and take 50 mL samples at the desired time points after galactose addition.

### 3.2 Genomic DNA Extraction

1. Pellet the cells by spinning for 3 min at RT at  $1600 \times g$  in 50 mL tubes.
2. Wash the cells in 1 mL spheroplasting solution and transfer the samples to 1.5 mL microcentrifuge tubes.
3. Spin for 3 min at  $1600 \times g$  and completely remove supernatant with a tip.
4. Freeze and store the pellets at  $-20\text{ }^{\circ}\text{C}$ .
5. Thaw the samples at room temperature and resuspend the cell pellets in 400  $\mu\text{L}$  spheroplasting solution, 14 mM  $\beta$ -mercaptoethanol.
6. Add 100  $\mu\text{L}$  of Zymolyase solution to each sample and invert the tube 4–6 times. Incubate the samples at  $37\text{ }^{\circ}\text{C}$ . After 30 min check the formation of spheroplasts under a light microscope (*see Note 7*).
7. When  $>95\%$  cells become spheroplasts, spin for 1 min at  $15,000 \times g$  and carefully remove supernatant with a tip.
8. Gently resuspend spheroplasts in 400  $\mu\text{L}$   $1\times$  TE (do not vortex).
9. Add 90  $\mu\text{L}$  lysis solution (prepared just before use). Immediately mix by inverting the tube several times and incubate the samples for 30 min at  $65\text{ }^{\circ}\text{C}$ . This causes spheroplasts lysis.
10. Add 80  $\mu\text{L}$  5 M potassium acetate and mix by inverting the tube several times. Place the tubes on ice for at least 1 h.
11. Spin for 15 min at  $15,000 \times g$  at  $4\text{ }^{\circ}\text{C}$ . Transfer the supernatant to new 1.5 tubes. Discard the pellets.
12. Add 1 mL ice-cold 96% ethanol and mix by inverting several times the tube. A white cloudy precipitate with the nucleic acids should form.
13. Spin for 5–10 min at  $15,000 \times g$  at  $4\text{ }^{\circ}\text{C}$  and remove the supernatant.
14. Wash the pellet with 1 mL ice-cold 70% ethanol and immediately discard the ethanol.
15. Air-dry the pellet.
16. Add 500  $\mu\text{L}$   $1\times$  TE. Let tubes sit for 15 min at room temperature (or overnight at  $4\text{ }^{\circ}\text{C}$ ), then gently dissolve the pellet (do not vortex).
17. Once pellets are completely dissolved, add 2.5  $\mu\text{L}$  RNase A solution to each sample and incubate for 1 h at  $37\text{ }^{\circ}\text{C}$ .
18. Add 500  $\mu\text{L}$  2-propanol and invert several times the tube. The solution should become cloudy.
19. Spin for 15–30 min at  $15,000 \times g$  at  $4\text{ }^{\circ}\text{C}$  and remove the supernatant.

20. Wash the pellet with 1 mL ice-cold 70% ethanol, invert and immediately discard the ethanol.
21. Air-dry the pellet.
22. Add 50  $\mu$ L 1 $\times$  TE. Let tubes sit for 30 min at room temperature or overnight at 4  $^{\circ}$ C, then gently dissolve the DNA pellet. Avoid pipetting or vortexing to prevent DNA shearing (*see Note 8*).
23. Load 1  $\mu$ L of each genomic DNA sample (added to 9  $\mu$ L of 1 $\times$  DNA loading buffer) on a 0.8% agarose gel with ethidium bromide and run in 1 $\times$  TAE buffer. Check under the UV lamp the quality of the extracted DNA (*see Note 9*).

### **3.3 DNA Native Gel Electrophoresis and Transfer**

1. Digest at least 5  $\mu$ g (up to 10  $\mu$ g) of DNA for 5–6 h at 37  $^{\circ}$ C with 20 U of SpeI restriction enzyme (New England Biolabs) with 1 $\times$  enzyme buffer in a total volume of 25  $\mu$ L.
2. In the meanwhile melt 0.8% agarose in 450 mL TAE 1 $\times$  buffer and ethidium bromide, and pour into a gel tray. Prepare also a smaller 0.8% agarose gel to check for the digestion reaction efficiency.
3. Once the gels are solid, put them in a gel running apparatus and fill in with TAE 1 $\times$ .
4. Before loading the Southern blot gel, test 1  $\mu$ L of each digestion reaction (added to 9  $\mu$ L of 1 $\times$  DNA loading buffer) on a 0.8% agarose gel with ethidium bromide and run in 1 $\times$  TAE buffer (*see Note 9*).
5. Load now on the 450 mL gel the whole digestion reaction mixed with 5  $\mu$ L of 6 $\times$  bromophenol blue loading dye, as well as a DNA ladder.
6. Run the gel overnight at 2 V/cm (*see Note 10*).
7. Once the run is complete remove the gel from the tank and take a picture under UV light.
8. Soak the gel 30–60 min with gentle agitation in Denaturing solution.
9. Soak the gel 30–60 min with gentle agitation in Neutralizing solution.
10. Blot overnight the DNA from the gel onto a nylon neutral membrane by capillary transfer with 10 $\times$  SSC buffer as for standard Southern blot procedure (*see Note 11*).
11. After overnight transfer quickly soak the membrane in SSC 4 $\times$  or water (optional).
12. Let the filter air-dry on 3MM paper for at least 30 min.
13. Cross-link the DNA on the membrane with an UV cross-linker by following the instructions of the manufacturer.

### 3.4 Probe Labeling and Filter Hybridization

1. Insert the filter in a hybridization tube and quickly rinse it with ddH<sub>2</sub>O.
2. Block the filter by incubating 5 h in a hybridization oven at 55 °C in 25 mL hybridization buffer (prehybridization step).
3. In the meanwhile, proceed with probe labeling according to the manufacturer protocol (*see Note 12*). We generally use 25–50 ng of gel-extracted DNA template.
4. As soon as the prehybridization step is complete, discard the prehybridization solution and add 25 mL of fresh hybridization buffer.
5. Denature the labeled probe for 5–10 min at 98 °C and immediately add it to the hybridization tube (eventually keep it on ice).
6. Incubate overnight at 55 °C by gently rotating in the hybridization oven.
7. Next day remove the hybridization solution (*see Note 13*).
8. Wash the filter 60 min at 55 °C with washing solution.
9. Wash the filter 30 min at 55 °C with washing solution.
10. Air-dry the filter on Whatman™ 3MM paper.
11. Expose the filter to an imaging plate or to an autoradiography film in an autoradiography cassette with intensifying screens (*see Note 14*).
12. Develop the filter. *See Fig. 2* for a representative gel image.

---

## 4 Notes

1. As galactose is reported to isomerize at elevated temperatures, its sterilization should be obtained by filtration. Autoclaving should be avoided.
2. There are two commercially available preparations of Zymolyase from Nacalai Tesque: Zymolyase® 20T and Zymolyase® 100T (lytic activity of 20 U/g and 100 U/g, respectively). Zymolyase–100T can also be used but the concentration in the spheroplasting solution and/or the time of incubation should be carefully adjusted.
3. Ethidium bromide can be substituted by less-toxic dyes such as GelRed, or SYBR Safe.
4. The UCC5913 strain has a *LYS2* wild type gene on the distal side (compared to the TG sequences) of the HO cut site and it is worth streaking out the cells on SD -Lys + 2% glucose plates (*see Subheading 2*) before inoculation. This allows to select for cells with an intact construct at chromosome VII.
5. Be sure to inoculate a proper amount of cells in YEP + raffinose medium to reach a density of at least  $6 \times 10^6$  cells/mL the next

day. Cells grow much slower in presence of raffinose compared to glucose as a carbon source. Usually the volume of the culture inoculated overnight should be no less than half of the volume needed the following day. We usually collect 50 mL of culture for each time point to be analyzed.

6. To monitor de novo telomere addition in G2/M phase, cells can be arrested in metaphase by treatment with the drug nocodazole that interferes with microtubule polymerization. To this end, prepare a 100× nocodazole stock by dissolving nocodazole at 1.5 mg/mL in 100% dimethyl sulfoxide (DMSO). Add the nocodazole stock to the exponentially growing cells in YEP + raffinose medium to reach a final concentration of 15 µg/mL. After 2.5–3 h at 25–30 °C, about 90% cells should be arrested as large-budded cells (score at the microscope). Nocodazole-mediated arrest can be maintained for several hours.
7. In order to monitor the efficiency of zymolyase digestion, prepare two drops (5 µL) of the same sample on a microscope slide and add 1–2 µL of 10% SDS to one of them. Spheroplasts will lyse in SDS solution. We usually obtain >95% spheroplasts in 40–50 min.
8. At this step the solution does not often look clear. In this case ensure that the pellet is fully resuspended, then spin again for 10–15 min at 15,000 × *g* at RT and take the clear supernatant.
9. Yeast genomic DNA is mainly visible as a unique discrete band of high molecular weight on an agarose gel. After digestion, depending on the enzyme used and the frequency of cutting sites in the genome, it usually appears as a pattern of several bands of different size.
10. In an 0.8% agarose gel the bromophenol blue usually migrates as DNA fragments of about 700 bp in size. This corresponds to the cut product to be monitored for de novo telomere addition. Since the latter will only increase of few dozens base pairs upon telomerase action, a nice separation of fragment sizing between 1000 and 700 bp is required. We usually stop electrophoresis run when the blue front has migrated about 16–18 cm from wells. This usually requires at least 16 h run at 55 V for a 25 cm long gel.
11. Briefly, fill a tank with 10× SSC buffer (approximately 1 L). Place a support in the tank, wet it with some SSC 10× and immediately create a “bridge” with Whatman™ 3MM paper onto the support. Pour some more 10× SSC over the 3MM paper so that it is fully wet and carefully remove air bubbles. Place the gel on top. Wet a nylon membrane in ddH<sub>2</sub>O (it should fit with the size of the gel, e.g., 20 × 20 cm) and lay it on top of the gel removing air bubbles. Wet 2 Whatman™



3MM papers (e.g., 20 × 20 cm) in 10× SSC and lay them on the membrane. Place a stack of absorbent paper towels on top, place a glass plate and put a 0.5 kg weight. Allow upward capillary transfer at room temperature for at least 18 h.

12. Clean up of the probe with G50 columns is usually not required.
13. Probes can eventually be labeled in advance and stored at 4 °C. Probes recovered after hybridization (in hybridization buffer) can be reused multiple times generally within a month (<sup>32</sup>P half-life is about 14 days). In all cases, probes need to be boiled for 5–10 min before each use.
14. In case autoradiography films are used, the cassette can be placed at –80 °C to further intensify the signal. With freshly labeled dATP- $\alpha$ <sup>32</sup>P, we usually obtain a good signal after a 4 h exposure at –80 °C.

---

## Acknowledgments

We thank G. Lucchini for critical reading of the manuscript. D.B. has been supported by a fellowship from CancerTelSys (grant 01ZX1302) in the E:med program of the German Federal Ministry of Education and Research (BMBF). Research in Longhese's lab is supported by Associazione Italiana per la Ricerca sul Cancro (AIRC) (grant number 15210) and Cofinanziamento 2015 MIUR/Università di Milano-Bicocca.

## References

1. Bonetti D, Martina M, Falcettoni M, Longhese MP (2014) Telomere-end processing: mechanisms and regulation. *Chromosoma* 123:57–66
2. Wellinger RJ, Zakian VA (2012) Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: beginning to end. *Genetics* 191:1073–1105
3. Marcand S, Brevet V, Gilson E (1999) Progressive cis-inhibition of telomerase upon telomere elongation. *EMBO J* 18:3509–3519
4. Teixeira MT, Arneric M, Sperisen P, Lingner J (2004) Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. *Cell* 117:323–335
5. Bianchi A, Shore D (2007) Increased association of telomerase with short telomeres in yeast. *Genes Dev* 21:1726–1730
6. Diede SJ, Gottschling DE (1999) Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerase  $\alpha$  and  $\delta$ . *Cell* 99:723–733
7. Diede SJ, Gottschling DE (2001) Exonuclease activity is required for sequence addition and Cdc13p loading at a de novo telomere. *Curr Biol* 11:1336–1340
8. Kramer KM, Haber JE (1993) New telomeres in yeast are initiated with a highly selected subset of TGI-3 repeats. *Genes Dev* 7:2345–2356

## Assays to Study Repair of Inducible DNA Double-Strand Breaks at Telomeres

Roxanne Oshidari and Karim Mekhail

### Abstract

The ends of linear chromosomes are constituted of repetitive DNA sequences called telomeres. Telomeres, nearby regions called subtelomeres, and their associated factors prevent chromosome erosion over cycles of DNA replication and prevent chromosome ends from being recognized as DNA double-strand breaks (DSBs). This raises the question of how cells repair DSBs that actually occur near chromosome ends. One approach is to edit the genome and engineer cells harboring inducible DSB sites within the subtelomeric region of different chromosome ends. This provides a reductionist and tractable genetic model system in which mechanisms mediating repair can be dissected via genetics, molecular biology, and microscopy tools.

**Key words** DNA double strand break (DSB), Telomeres, Subtelomeres, Inducible, Survival, Endonuclease, Nuclear envelope, Chromatin silencing, Homologous recombination (HR), Nonhomologous end joining (NHEJ)

---

### 1 Introduction

During S phase of the cell cycle, the ends of linear eukaryotic chromosomes cannot be replicated by the replication machinery [1]. This is a long-appreciated limitation known as the end replication problem [1]. To prevent rapid erosion of chromosome ends over rounds of replication, the ends are constituted of highly repetitive DNA sequences known as telomeres, whose length is maintained by the enzyme telomerase [2, 3]. Telomeres also prevent the deleterious fusion of various chromosome ends to each other by preventing the classic DNA repair machinery from recognizing them as the ends of a DNA double-strand break (DSB) [4]. Telomeres can also help initiate silent chromatin structures that then spread to nearby chromosomal regions known as subtelomeres [5, 6]. In yeast, telomeric TG<sub>1-3</sub> repeats help recruit the Silent Information Regulator (SIR) protein complex composed of Sir3, Sir4, and the catalytic subunit Sir2 histone deacetylase [7]. Iterative cycles of Sir2-mediated deacetylation of histone tails at neighboring

nucleosomes followed by recruitment of additional SIR complexes allow silent chromatin to spread within subtelomeres [7]. This form of chromatin silencing compacts chromatin, making it less accessible to the DNA repair and recombination machineries and ultimately more stable [7, 8].

Importantly, this chromatin silencing is greatly promoted by the clustering of telomeres in a handful of foci along the nuclear envelope [6, 8]. Specifically, perinuclear telomere clustering increases the local concentration of SIR complexes [8, 9]. Proteins mediating perinuclear telomere tethering include the Cohibin complex (composed of Lrs4 and Csm1), Enhancer of silent chromatin 1 (Esc1), the inner nuclear membrane protein Monopolar spindle 3 (Mps3), Ku proteins, and even telomerase [8, 10–13]. Given that chromosome ends are confined to perinuclear domains that are highly repressive to DNA recombination and repair, what happens to subtelomeric DSBs? Interestingly, DSBs within the rDNA repeats of yeast transiently escape their usual Rad52/recombination-repressive perinuclear/nucleolar environment to undergo repair [14, 15]. In *Drosophila*, heterochromatic DSBs exit their Rad51/recombination-repressive subnuclear environment to undergo homologous recombination (HR)-dependent repair at the nuclear periphery [16]. Similarly, heterochromatic DSBs relocate outside their domain for repair in cultured human cells [17]. It was recently shown that telomeric DSBs are transiently relocated to nuclear pore complexes (NPCs) enriched in factors mediating an error-prone repair mechanism called break-induced replication (BIR) [18–22]. This mechanism, also implicated at nontelomeric BIR-repairable DSBs, allows cells to survive DSBs but compromises overall genome integrity [18]. Importantly, yeast genetic screens employing inducible subtelomeric DSBs have revealed that BIR-dependent repair is dependent on perinuclear telomere tethering and chromatin remodeling that then allow kinesin motor proteins to transiently relocate subtelomeric DSBs to specialized nuclear envelope regions such as nuclear pore subcomplexes critical to key repair steps [18].

To study subtelomeric DSB repair, one can generate a single DSB by targeting a restriction endonuclease to a desired chromosome end locus [18, 21, 23]. This strategy allows for a site-specific and controlled approach to DSB investigation. Researchers commonly use the I-SceI endonuclease to generate single DSBs [18, 21, 23]. Here, we show how to exploit this approach in concert with classic yeast genetics to determine DSB repair efficiency as well as repair pathway choice at chromosome ends [24]. Briefly, the *URA3* reporter gene flanked by two inverted I-SceI restriction sites is inserted 3.5 kb downstream of the telomere on the left arm of chromosome XI, or another subtelomeric region as needed. As an internal control, one can alternatively insert the above-mentioned DSB site outside of the subtelomeric region on the

same chromosome arm studied. Cells are then transformed with a plasmid harboring the I-SceI open reading frame (ORF) under a galactose-inducible promoter allowing for controlled DSB induction via plating on galactose-containing media. These cells are subsequently grown and plated on media supplemented with either glucose (GLU, no induction) or galactose (GAL, induction). Cells that repair the induced DSB grow and generate colonies on media supplemented with galactose, while unsuccessful repair is lethal [18]. The DSB induction and repair efficiency of varying mutants can subsequently be calculated and the repair pathway employed by cells can further be determined via PCR. This method can be adapted to study any telomere, which can also be visualized via microscopy for movement and dynamic character.

---

## 2 Materials

### 2.1 Media

Media is prepared as detailed in [25].

1. SC-LEU(GAL) agar plates and liquid media.
2. SC-LEU(GLU) agar plates and liquid media.
3. SC and SC-URA agar plates.
4. SC and SC-URA agar plates.
5. Autoclaved ddH<sub>2</sub>O.

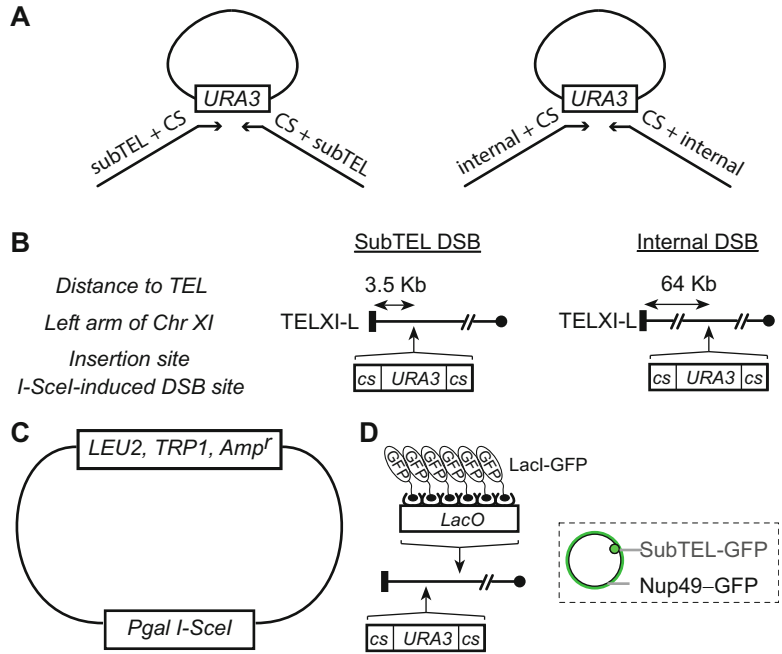
---

## 3 Methods

Carry out all procedures at room temperature unless otherwise indicated. All water used should be sterile ddH<sub>2</sub>O.

### 3.1 Preparation of Yeast Strains

1. Generate the *URA3* cassette flanked by two inverted I-SceI cut sites for integration into the preferred subtelomeric or internal locus. Briefly, PCR amplify *URA3* using primers harboring I-SceI cut sites and subtelomeric sequences that flank the insertion site (Fig. 1a). This allows for targeted insertion of the cassette via homologous recombination (Fig. 1b). Alternatively, insertion of the cassette can be carried out as described [21, 23]. For integration into any subtelomere, sequences with homology to all subtelomeres of *S. cerevisiae* are listed in Table 1. As an example in this paper, we will also focus on a DSB cassette inserted into the left subtelomere of chromosome XI within the ykl222c ORF 3.5 kb away from the telomere [18, 21]. As a control, the cassette can also be inserted ~64 kb away from the same telomere (Fig. 1b).
2. Transform strain of interest with the integrating PCR product via the lithium acetate method [18].



**Fig. 1** Production and integration of the I-SceI *URA3* cassette. (a) Amplify *URA3* using primers harboring I-SceI cut sites and sequences homologous to the desired locus of insertion, allowing for targeted insertion via homologous recombination. (b) Cassette is inserted 3.5 kb away from subTEL XI-L, and as an internal control is integrated ~64 kb away from the same telomere. (c) Major features of the I-SceI expression plasmid pKM97. (d) A *LacO* array can be inserted downstream of the I-SceI *URA3* cassette to allow for binding of ectopically expressed LacI-GFP. This allows for visualization of the DSB site by microscopy

3. Once integration of the cassette is confirmed by plating on SC-URA(GLU) media and PCR over the integration site, transform strains with the GAL-inducible I-SceI expression plasmid pKM97 (Fig. 1c) [18, 21, 23].
4. If microscopy-based tracking of the DSB site is desired, a *LacO* array can be integrated [26] adjacent to the DSB site allowing for binding of ectopically expressed GFP-tagged LacI (Fig. 1d).

**3.2 Inducible Subtelomeric DSB Repair Assay**

1. Grow 3 mL of *S. cerevisiae* culture in SC-LEU(GLU) medium overnight at 30 °C to saturation.
2. Inoculate 300 µL of saturated culture into 5 mL of SC-LEU (GLU) medium and grow at 30 °C for 3 h. or until cells reach log phase (see Note 1).
3. Pellet  $5.0 \times 10^7$  cells at  $18,000 \times g$ .
4. Wash pellets and resuspend in 1 mL of water.

**Table 1**  
**List of overhang sequences for insertion of I-SceI *URA3* cassette into any of the subtelomeres of *S. cerevisiae***

TEL	Forward <sup>a</sup>	Reverse <sup>a</sup>
I-L	GCTGGACATTATCCAGTAAAGGCTTGGTAGTAACCATAAT	ATGAGTCITTTCAAAGTTCCTAGCGTTTCGTACCTGGGTAAT
I-R	CCGGAGAGATTCCATATTCGGACGTCACATTCGCCAAT	GCCCTGACACAAATATAAACCAATGTACCACCTCATGTAA
II-L	GCTCTGTGTCGACTTTCITTCGCCAGGTAAACTTGCCTTG	GTTACTACACGTAGATGAGCTATCGATTTTTTCTGCAIAC
I-IR	GATGTTGATTTATCAATTTCCGGTATAGTTTCCGGTATGG	CTAGTCTGAAACAGCGGAACTTCCACTGGATCAAAGCTGG
III-L	GCTATATGCCCTACGGCCTCGTTTAGCCCCATCTCGCAT	ACTACTTCTCTGTGTATATACCTATGTTACGGTATTGC
III-R	GTCCATGACCTTAGGTCTTCAATATGAGCCTTCGAAAGTAC	TCCGTCCTCGCCTTATCACTTCTGGATCTTTGAAACGTTT
IV-L	ACGCTGATATCTTTATAAATAACCCCGGAATGGAAATTTGTG	TCACAGGTCTTAATACTCGAAACGCCAATTGATGACTGTG
IV-R	CATTGTTCAATCTAGTACTCGGAATGCGGAAGACGTAG	ATTTTTTGATCATGATCTGAACTCGCGCCTCACCTGCCA
V-L	CAGTGTACAGTTGGATGGCTTGTGTATCATCGACGTA	GCTTTCAGCTACTCGCTTTTGGCCGGTCTCTCAGCGCAACGTA
V-R	GACAAACAACGGGTTGGGATGTGGAATGGATCAAAAATTC	GAGTCCITTCAGCGGGGTTTCTCGGCTACTTTCGTGGTACA
V-R <sup>b</sup>	TCCAGGGCGGACCAATCCCTTCAAAACAACCCACGTGATGG	CTATTGAATTTCAAAGTATGCACITTGAAACAACCGTGTAGA
VI-L	TAAGCGACTGTCTGTCCCTTCAAGCACACATAGAAAAC	TTGCAGGGCCGAAATCAAAGCGATGAATGAGACATCCTTCT
VI-R	GAGATGCAGCAAAATAACCACTGGTGTTTAAGAAATAATA	AGCAGAGTGCCTACGGGCTCTTCGGTTACTGACGTAATCA
VII-L	CTTGACGGTCGCCATATCATGAGATTACTTTGAAGCTTCC	AAAGAAAAACGTATAAAAACCTGGTTACGGTATGAGCTTCAA
VII-R	GAAAGGATCGATCCATTACGAAATGAAGAACCATTATTC	CGCAACTTACATACTTTTCAAGCCATGCTCGTCCACAGAGG
VII-R <sup>b</sup>	ATATTACAATAATTTGAATAGATCATAAATGGTATGACCGAT	CTATACCAACCATGACGTCATTAACCTTAAATGTTCCCTTAAT
VIII-L	CGCGTTACTCTTGAAATTTTTTGTATCATGATCTGAACTCG	AGTACTCGGAATCGGAAAGACCGTAGTGGCAGGTGAGGGGG
VIII-R	GATGGATTTCTTGTCAAAAGCATAGCAATCAACATACTA	GAAACTTTCATTTTATTTTTGTAAAGTTTCGAAATTAACAA
IX-L	GTGCCGCTCAAAAAGATTGCTTTCTCAAAAAGCGTCAAAAT	GTCTACCGGCAGTCGCAATTTGGGGGCATAACTAACCTTGTG
IX-R	CTGTATTGAAGATCATCAAAATACATACTTTCATTAACAGAA	CCAGAATTACAGAAGACGATTATTGCAACTGAAGCTCAAT

(continued)

**Table 1**  
(continued)

TEL	Forward <sup>a</sup>	Reverse <sup>a</sup>
X-L	GGATTTCICCTTGATTAGATTAAACATCCGTTGGATAG	AGCTCAAACGGTCGGAAAGATCTCAGCAGAGGTCCTATCCAG
X-R	CGTACTTTGAGAAAACTATATGGTTGTACGAAAAAGAAAT	ATAGTAATGAGATTAATAAAGGCCCAAAGAATATTTGGCC
XI-L	GCAAAATTAGAAATACATAGGATATTATGATGTCACCTAAA	TAACGTTTGATATGTAAGTTCTAAGTATGTTTCTCTATAT
XI-R	TGACATAACAGGATAGGATAAACATGCTCGATAGACCCAC	TAAAAAGTACGGACACAGCTATATGCGTCGGCAAGTCAAAT
XII-L	CTGGAAATCTGGTTCGCTTCAACAAAACGTCAGGAACITTT	AAAAAATTTGCTGTTCCGTTCCGTCATGGTCGGGAAAAAATGTACA
XII-R	GATGCCGTGAAAGCTTTAATGGTGTGCTTGTGCTTGTGA	ACGCCGTTGGTGTGTCATCAAATATATCTAAATCCCTTGCAG
XII-R <sup>b</sup>	AGCGCAGATTCGCTTCTCTCAATAGAGTAGCTTAATTA	GATTGTCCAGTTTCCGTCCTTATCATCATCTAAGAATGTAA
XIII-L	GTCCAGTTTCCGCTTTATCATCATCTAAAGAAATGTAATAAT	GCAATAGCGCAGATTCGCTTCTTCTCAATAGAGTAGCTTA
XIII-R	CGCATAATGGAAATAGCTTCTTAAAGGATGGAGATAAGG	GAGAAAGCTCCCAAGTACTTCCCTTGTCTTTGTCATTTTTT
XIV-L	GTAAAGCCATCGCAACCTTCTTCAATAAAGTTTCTTTACAG	ATGATCAGGTTGAGCCGATGCGGTTGCTTGAATGTGGCCC
XIV-R	CGAGAGCTTTGACCAAAAAGTGTCCCGACTTTATAATGAG	CCAAAAGCATACCATATAAACCCTATGAATTAAGCGGCGGT
XV-L	CGTTCCAACATCAATACCTCTAGAATTTCCATTTTGGGTT	AAAACTAGGATCAACATCAGGATAGTTACTACCGTTAGCC
XV-R	TATTCTACATGTTTCTGAATACCGCGAAGGGCTGTCTGG	GGAAAGTTCCCTCATGTACTGCCGAAAAAGTTGCGTATTCAA
XVI-L	GGCAATATCAACCACCTGTAGGCATATAAAAAGCAGTCGTTT	TTGCATGCAAGGCTCAGTTTGATAACCTTTGGGAAGAAT
XVI-R	CATTCTCAGATGATGATAAGACGGAAAACCTGGACAATCTTT	ATTGTTATGCTTTTTTGACAAGAAATCCATCAATATAAACA

<sup>a</sup>These sequences should be added to the 5' end of the following *URA3* amplifying primers harboring cut sites (*underlined*): Forward: TAGGGATAACAGGGTAAATCACGCTTTTCAATTCAATTCATC, Reverse: TAGGGATAACAGGGTAAATTAGTTTGTGGCCGCAICTTC

Sequences are designed for integration 3–5 kb from the telomere, however in the event that this region is entirely occupied by an ORF, the <sup>b</sup> indicates an alternative set of sequences further from the telomere. They still reside within the subtelomeric region.

5. Prepare 1:10 serial dilution in sterile 1.5 mL microcentrifuge tubes by mixing 100  $\mu\text{L}$  of the  $5.0 \times 10^7$  cells/mL tube with 900  $\mu\text{L}$  of water to achieve an effective concentration of  $5.0 \times 10^6$  cells (*see Note 2*). Repeat this dilution procedure four more times to generate  $5.0 \times 10^5$  to  $5.0 \times 10^2$  cells/mL samples (Fig. 2a-i).
6. Plate 200  $\mu\text{L}$  of the  $5.0 \times 10^2$  cells/mL suspension on four SC-LEU(GLU) plates (~100 cells per plate) (*see Note 3*) (Fig. 2a-ii).
7. Plate 200  $\mu\text{L}$  of  $5.0 \times 10^3$  and/or  $5.0 \times 10^4$  cells/mL suspension on four SC-LEU(GAL) plates (~1000 and/or 10,000 cells/plate).
8. Incubate the plates at 30 °C for 7 days.
9. Calculate the ratio of the number of colonies growing over the number of cells plated and average the ratios for both glucose and galactose plates. Standard workflow and example calculations are shown (Fig. 2a).
10. For a final DSB survival rate, divide the average ratio of galactose plated cells over that of glucose (Fig. 2a-i).
11. Perform *t*-test analyses to evaluate statistical support where applicable.

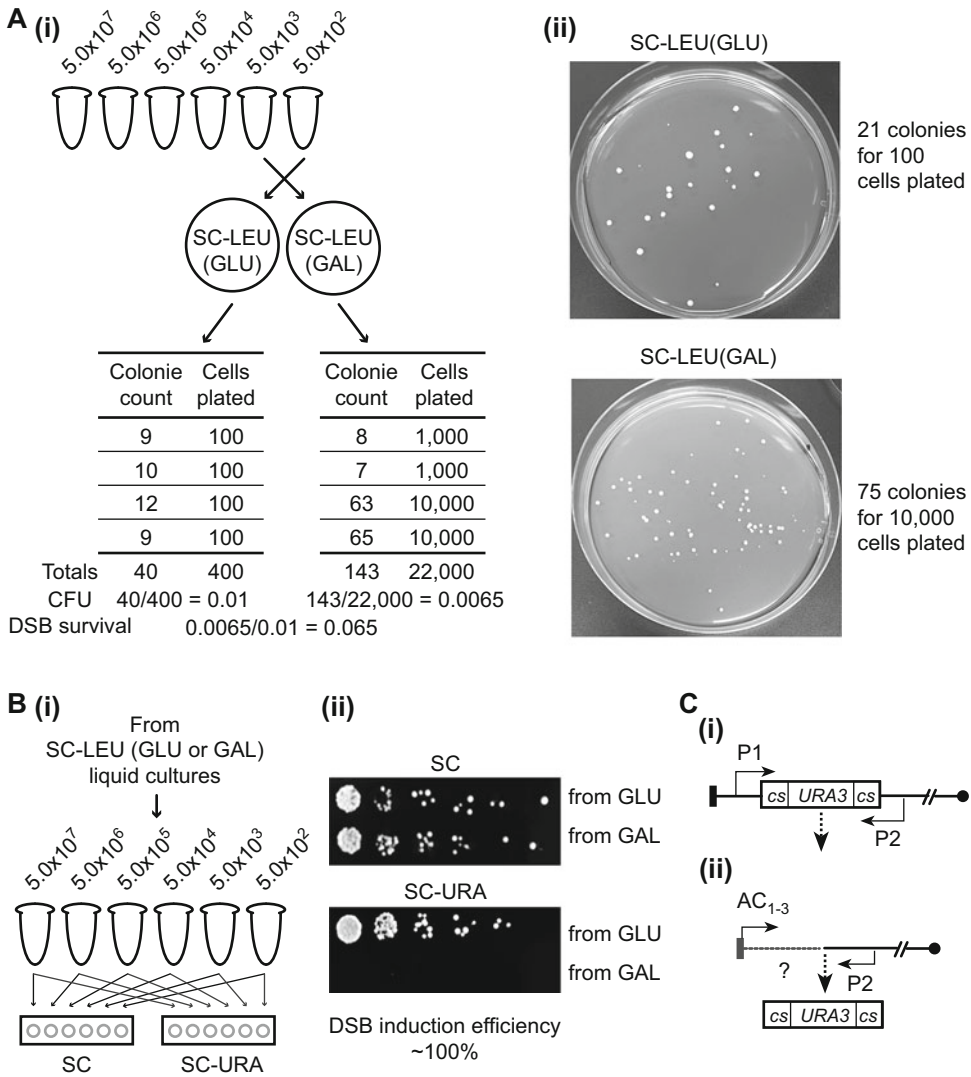
### 3.3 Assessing DSB Induction Efficiency

1. Grow 3 mL of *S. cerevisiae* culture in SC-LEU(GLU) medium overnight at 30 °C to saturation.
2. Inoculate 300  $\mu\text{L}$  of saturated culture into 5 mL of SC-LEU(GLU) medium and 5 mL of SC-LEU(GAL) and grow at 30 °C for 3 h. or until cells reach log phase.
3. Pellet  $5.0 \times 10^7$  cells at  $18,000 \times g$ .
4. Wash pellets and resuspend in 1 mL of water, then add the 225  $\mu\text{L}$  cell suspension to the first well of a 96-well plate.
5. Prepare 1:10 serial dilutions by mixing 25  $\mu\text{L}$  of the  $5.0 \times 10^7$  cells/mL well with 225  $\mu\text{L}$  of water in the next well to achieve an effective concentration of  $5.0 \times 10^6$  cells/mL. Repeat four times to achieve a final dilution of  $5.0 \times 10^2$  cells/mL (Fig. 2b-i).
6. Spot 3  $\mu\text{L}$  of each dilution on SC and SC-URA plates (Fig. 2b-i).
7. Incubate plates at 30 °C for 5–7 days.
8. Qualitatively and quantitatively determine DSB induction efficiency (Fig. 2b-ii).

### 3.4 DNA Repair Pathway Choice Determination

1. To determine the repair mechanism employed by cells to survive the induced DSB, perform colony PCR on colonies surviving on galactose-containing plates using primer pairs that bind just outside the DSB site (Fig. 2c-i; Table 2).





**Fig. 2** Standard workflow of the Inducible subtelomeric DSB repair and DSB induction efficiency assays. **(a)** (i) Begin with performing a 1:10 serial dilution to a final concentration of  $5.0 \times 10^2$  cells/mL. Plate on SC-LEU (GLU) and SC-LEU(GAL) media and count colonies after 7 days of growth. Calculate colony forming units (CFUs) and a final survival rate. (ii) An example of plated cells on SC-LEU(GLU) and SC-LEU(GAL). **(b)** (i) After culturing in liquid SC-LEU(GLU) and SC-LEU(GAL) media, perform serial dilutions as described previous and spot 3  $\mu$ L of each dilution on SC(GLU) and SC-URA(GLU) plates. (ii) Example of DSB induction efficiency plating and the subsequent determination of induction efficiency. **(c)** PCR primers for DNA repair pathway choice determination. The sequences of the primers shown here are included in Table 2 and can be used for subTEL XI-L. (i) For a subtelomeric DSB site, design two primers just outside the I-SceI *URA3* cassette. PCR product with the expected size indicates an incomplete DSB while a smaller amplicon suggests NHEJ and no amplicon indicates BIR or a similar pathway. (ii) BIR survivors can be further analyzed using a generic  $AC_{1-3}$  primer and the more internal P2 primer. Reactions would not yield an amplicon if the BIR event engaged a very long chromosome arm

**Table 2**  
**Primers used to screen survivors for type of DNA repair employed**

Primer	Sequence
P1	CTGAGTCTGCACTAGACAAT
P2	ATCTTGATCTCAAAAGCACC
AC1-3	ACCACACACCCACCAC

- (a) To perform colony PCR, pick up a small portion of the colony with a pipette tip and smear along the side of a PCR tube.
  - (b) Microwave PCR tubes on high for 1 min 30 s, allow cooling for 1 min, and repeat.
  - (c) Carry out preferred PCR protocol with the primers described in **step 1**.
2. Electrophorese PCR products in a 1% TAE-agarose gel.
  3. Calculate the expected amplicon size, taking into account the inserted cassette. A PCR product of that size would indicate an incomplete DSB. A visibly smaller amplicon indicates NHEJ. No amplicon indicates the BIR or similar pathway.
  4. For the system focusing on subTELXI-L, perform colony PCR with the primers shown in Fig. 2c-i and Table 2. No PCR product indicates the BIR or similar pathway, while an amplicon of ~0.7 kb indicates NHEJ. The production of a ~2 kb product indicates an incomplete DSB.
  5. BIR survivors can be further analyzed using a generic AC<sub>1-3</sub> forward primer and the more internal reverse primer used in **step 1** (Fig. 2c-ii; Table 2). The majority of reactions should produce no amplicon. For subTEL-XI-L, perform the PCR with the generic AC<sub>1-3</sub> primer and P2 (Table 2; *see Notes 4 and 5*).

---

## 4 Notes

1. The volume used for inoculation may require adjustment to compensate for varying rates of growth among strains. Overnight cultures may be at different levels of saturation. Alternatively, some mutants may reach log phase much slower than wild-type cells. All strains should be in log phase at the same time.
2. Between dilutions, mix the cell suspension by pipetting up and down to ensure that cells have not settled to the bottom of tubes. This should also be done before plating.

3. The  $5 \times 10^2$  dilution is generally appropriate for wild-type and mutants with normal growth. However, more concentrated dilutions may be required for slower growing mutants. Different dilutions can be plated during this assay to establish ideal conditions. This should also be done when plating on galactose-containing media.
4. When designing primers, ensure that only the AC<sub>1-3</sub> primer falls within repetitive elements/regions.
5. PCR-products can also be sequenced to confirm the DNA repair pathway used by surviving cells.

---

## Acknowledgments

We thank Mekhail lab members for comments. K.M. is supported by grants from the Canadian Institutes of Health Research (CIHR), the Ontario Ministry of Research and Innovation Early Researcher Award (MRI-ERA), and the Canada Research Chair (CRC) in Spatial Genome Organization. The authors have no conflict of interest.

## References

1. Pfeiffer V, Lingner J (2013) Replication of telomeres and the regulation of telomerase. *Cold Spring Harb Perspect Biol* 5(5):a010405. doi:cshperspect.a010405 [pii]. doi:10.1101/cshperspect.a010405
2. Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43(2 Pt 1):405–413. doi:0092-8674(85)90170-9 [pii]
3. Shampay J, Szostak JW, Blackburn EH (1984) DNA sequences of telomeres maintained in yeast. *Nature* 310(5973):154–157
4. Chan SW, Blackburn EH (2003) Telomerase and ATM/Tel1p protect telomeres from non-homologous end joining. *Mol Cell* 11(5):1379–1387
5. Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* 63(4):751–762. doi:0092-8674(90)90141-Z [pii]
6. Mekhail K, Moazed D (2010) The nuclear envelope in genome organization, expression and stability. *Nat Rev Mol Cell Biol* 11(5):317–328. doi:nrm2894 [pii]. doi:10.1038/nrm2894
7. Moazed D (2001) Common themes in mechanisms of gene silencing. *Mol Cell* 8(3):489–498
8. Taddei A, Gasser SM (2012) Structure and function in the budding yeast nucleus. *Genetics* 192(1):107–129. doi:192/1/107 [pii]. doi:10.1534/genetics.112.140608
9. Chan JN, Poon BP, Salvi J, Olsen JB, Emili A, Mekhail K (2011) Perinuclear cohibin complexes maintain replicative life span via roles at distinct silent chromatin domains. *Dev Cell* 20(6):867–879. doi:S1534-5807(11)00206-1 [pii]. doi:10.1016/j.devcel.2011.05.014
10. Taddei A, Hediger F, Neumann FR, Bauer C, Gasser SM (2004) Separation of silencing from perinuclear anchoring functions in yeast Ku80, Sir4 and Esc1 proteins. *EMBO J* 23(6):1301–1312. doi:10.1038/sj.emboj.7600144. 7600144 [pii]
11. Bupp JM, Martin AE, Stensrud ES, Jaspersen SL (2007) Telomere anchoring at the nuclear periphery requires the budding yeast Sad1-UNC-84 domain protein Mps3. *J Cell Biol* 179(5):845–854. doi:jcb.200706040 [pii]. doi:10.1083/jcb.200706040
12. Schober H, Ferreira H, Kalck V, Gehlen LR, Gasser SM (2009) Yeast telomerase and the SUN domain protein Mps3 anchor telomeres and repress subtelomeric recombination. *Genes Dev* 23(8):928–938. doi:23/8/928 [pii]. doi:10.1101/gad.1787509

13. Hediger F, Neumann FR, Van Houwe G, Dubrana K, Gasser SM (2002) Live imaging of telomeres: yKu and Sir proteins define redundant telomere-anchoring pathways in yeast. *Curr Biol* 12(24):2076–2089. doi: [10.1038/S0960982202013386](https://doi.org/10.1038/S0960982202013386) [pii]
14. Torres-Rosell J, Sunjevaric I, De Piccoli G, Sacher M, Eckert-Boulet N, Reid R, Jentsch S, Rothstein R, Aragon L, Lisby M (2007) The Smc5-Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. *Nat Cell Biol* 9(8):923–931. doi: [10.1038/ncb1619](https://doi.org/10.1038/ncb1619)
15. Mekhail K, Seebacher J, Gygi SP, Moazed D (2008) Role for perinuclear chromosome tethering in maintenance of genome stability. *Nature* 456(7222):667–670. doi: [10.1038/nature07460](https://doi.org/10.1038/nature07460) [pii]. doi: [10.1038/nature07460](https://doi.org/10.1038/nature07460)
16. Chiolo I, Minoda A, Colmenares SU, Polyzos A, Costes SV, Karpen GH (2011) Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell* 144(5):732–744. doi: [10.1016/j.cell.2011.02.012](https://doi.org/10.1016/j.cell.2011.02.012)
17. Jakob B, Splinter J, Conrad S, Voss KO, Zink D, Durante M, Lobrich M, Taucher-Scholz G (2011) DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Res* 39(15):6489–6499. doi: [10.1093/nar/gkr230](https://doi.org/10.1093/nar/gkr230)
18. Chung DK, Chan JN, Strecker J, Zhang W, Ebrahimi-Ardebili S, Lu T, Abraham KJ, Duracher D, Mekhail K (2015) Perinuclear tethers license telomeric DSBs for a broad kinesin- and NPC-dependent DNA repair process. *Nat Commun* 6:7742. doi: [10.1038/ncomms8742](https://doi.org/10.1038/ncomms8742)
19. Chung DK, Mekhail K (2015) Repair by a molecular DNA ambulance. *Oncotarget* 6(23):19358–19359. doi: [10.18632/oncotarget.5140](https://doi.org/10.18632/oncotarget.5140)
20. Lydeard JR, Jain S, Yamaguchi M, Haber JE (2007) Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature* 448(7155):820–823. doi: [10.1038/nature06047](https://doi.org/10.1038/nature06047) [pii]. doi: [10.1038/nature06047](https://doi.org/10.1038/nature06047)
21. Therizols P, Fairhead C, Cabal GG, Genovesio A, Olivo-Marin JC, Dujon B, Fabre E (2006) Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. *J Cell Biol* 172(2):189–199. doi: [10.1083/jcb.200505159](https://doi.org/10.1083/jcb.200505159) [pii]. doi: [10.1083/jcb.200505159](https://doi.org/10.1083/jcb.200505159)
22. Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ (2008) Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* 322(5901):597–602. doi: [10.1126/science.1162790](https://doi.org/10.1126/science.1162790)
23. Ricchetti M, Dujon B, Fairhead C (2003) Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. *J Mol Biol* 328(4):847–862. doi: [10.1006/jmb.2003.3003](https://doi.org/10.1006/jmb.2003.3003) [pii]
24. Chapman JR, Taylor MR, Boulton SJ (2012) Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* 47(4):497–510. doi: [10.1016/j.molcel.2012.07.029](https://doi.org/10.1016/j.molcel.2012.07.029) [pii]. doi: [10.1016/j.molcel.2012.07.029](https://doi.org/10.1016/j.molcel.2012.07.029)
25. Dunham MJ, Gartenberg MR, Brown GW (2015) Methods in yeast genetics and genomics: a CSHL course manual. CSHL, New York
26. Lettier G, Feng Q, de Mayolo AA, Erdeniz N, Reid RJ, Lisby M, Mortensen UH, Rothstein R (2006) The role of DNA double-strand breaks in spontaneous homologous recombination in *S. cerevisiae*. *PLoS Genet* 2(11):e194. doi: [10.1371/journal.pgen.0020194](https://doi.org/10.1371/journal.pgen.0020194)

## Telomerase RNA Imaging in Budding Yeast and Human Cells by Fluorescent In Situ Hybridization

David Guérit, Maxime Lalonde, and Pascal Chartrand

### Abstract

Telomerase, the enzyme that elongates telomeres in most eukaryotes, is a ribonucleoprotein complex composed of a reverse transcriptase catalytic subunit (TERT in human, Est2 in the budding yeast *S. cerevisiae*), regulatory factors and a noncoding RNA called hTERC (in human) or *TLC1* (in budding yeast). Telomerase trafficking is a major process in the biogenesis and regulation of telomerase action at telomeres. Due to its higher signal-to-noise ratio, imaging of the telomerase RNA moiety is frequently used to determine telomerase intracellular localization. Here we describe how to image telomerase RNA in human and yeast cells using fluorescence in situ hybridization.

**Key words** Telomerase, Fluorescent in situ hybridization, RNA localization, *TLC1* RNA, hTERC

---

### 1 Introduction

Telomerase biogenesis and action at telomeres is linked to its cellular trafficking and localization [1]. The link between telomerase RNA maturation and trafficking has been well studied in the budding yeast *S. cerevisiae*. The budding yeast telomerase RNA (*TLC1*) life cycle includes many events that occur at various subcellular localizations. As part of its biogenesis, this RNA goes to the nucleolus for its 5' end maturation, and is exported to the cytoplasm for assembly of the holoenzyme [2–4]. Telomerase is then reimported in the nucleus, and its nuclear accumulation depends on the yKu complex, which interacts with the *TLC1* RNA [2, 5, 6].

In human cells, telomerase RNA (hTERC or hTR) localization depends on its maturation status. After transcription, hTERC 3' end processing possibly occurs in the nucleolus, since 3' end extended hTERC precursors can be detected in this subnuclear compartment [7]. hTERC is a small Cajal body RNA (scaRNA) which accumulates in Cajal bodies in human cancer cells [8, 9]. This localization depends on the interaction between a CAB box element in hTERC with the Cajal body protein TCAB1 [8, 10].

Despite the fact that other techniques exist, such as live cell imaging [11], RNA fluorescence in situ hybridization (RNA-FISH) remains the technique of reference to study telomerase RNA localization. As a ribonucleoprotein complex, telomerase localization can be studied using RNA-FISH by imaging its RNA component. Properly designed FISH probes can push the sensitivity to image single telomerase RNA molecules, which is not always possible by immunofluorescence on proteins [12]. It is also important to note that RNA-FISH is noninvasive and does not require genetic modification of its target, which allows the study of endogenous telomerase RNA. In budding yeast, a mix of oligonucleotide probes labeled with an organic dye can be used to detect the endogenous *TLC1* RNA [2]. A standard inverted epifluorescence microscope equipped with a CCD camera is enough to detect single molecules of *TLC1* RNA. Regarding human cells, the small size of hTERC RNA (450 nt) and its high GC content, limit the number of probes that can be designed to detect this RNA, which results in low signal-to-noise ratio and therefore prevents the detection of single molecules. Therefore, only the accumulation of hTERC RNA in large foci or in Cajal bodies can be detected [8, 9].

RNA-FISH on telomerase RNA can easily be combined with immunofluorescence (IF) to investigate its colocalization with different proteins or telomeres. Indeed, several studies successfully described the colocalization between telomerase RNA and Cajal bodies in human cancer cells, or between telomerase RNA and telomeres in both yeast and human cells using simultaneous RNA-FISH and immunofluorescence [2, 10, 13–16].

---

## 2 Materials

### 2.1 Preparation of Poly-L-Lysine-Coated Coverslips

1. Glass coverslips (22 × 22 mm #1.5).
2. Six 35 mm well plastic plates.
3. Poly-L-lysine, diluted to 0.01% in DEPC water.

### 2.2 Probe Design and Preparation

The probes contain amino-allyl deoxythymidine (N-allyl-dT) modified nucleotides, allowing covalent dye coupling to the probe. Oligonucleotides may vary in length and N-allyl-dT modification is incorporated every ten base pair for >20 bases oligonucleotide, and only at the 5' or 3' end for oligonucleotide of about 20 bases. Tables 1 and 2 include the sequences of the probes used for *TLC1* and hTERC FISH in yeast and human cells, respectively.

1. G25 Mini quick spin columns (Roche Life Science).
2. Cy3 or Cy5 monoreactive dye pack (GE Healthcare).
3. Sodium bicarbonate buffer 0.1 M pH 8.8: prepare 0.1 M of sodium carbonate ( $\text{Na}_2\text{CO}_3$ ) and sodium bicarbonate

**Table 1**  
**Sequences of the modified oligonucleotides for TLC1 FISH**

TLC1-1	t*gcgcacacacaagcat*ctacactgacaccagcat*actcgaattctt*tg
TLC1-2	ct*aataacaatt*agctgtaacatt*tgtgtgtgggggt*gtggtgatggg*aggc
TLC1-3	*ccagagttaacgat*aagatagacat*aaagtgacagcgct*tagcaccgt*
TLC1-4	ttacgt*tcttgatctt*gtgtcattggt*cagttactgat*cgccccaacct*
TLC1-5	tgcat*cgaaggcat*taggagaagt*agctgtgaat*acaaccaagaat*tca

\*Amino allyl modified-T

**Table 2**  
**Sequences of the modified oligonucleotides for hTERC FISH**

hTERC-1	t*gcgcgcggggagcaaaagcacggcgct*acgcccttctcagtt*agggttagaca
hTERC-2	gct*gacatttt*tgtttgctct*agaatgaacggt*ggaaggcggcaggccgaggct*t

\*Amino allyl modified-T

(NaHCO<sub>3</sub>) solutions (106 mg Na<sub>2</sub>CO<sub>3</sub> and 84 mg NaHCO<sub>3</sub> in 10 mL of DEPC treated water). Add Na<sub>2</sub>CO<sub>3</sub> to NaHCO<sub>3</sub> until a pH of 8.8 is reached.

### 2.3 Preparation of Mounting Media

1. 4,6-diamidino-2-phenylindole (DAPI).
2. p-phenylenediamine.
3. Glycerol.

### 2.4 Fluorescent In Situ Hybridization on Yeast Cells

1. 32% paraformaldehyde, electronic microscopy grade, single-use sealed ampules (Electron Microscope Sciences) (*see Note 1*).
2. Phenylmethylsulfonyl fluoride (PMSF). Prepare a 17.5 mg/mL stock solution in ethanol, keep at -20 °C.
3. Vanadyl Ribonucleoside Complex (VRC) 200 mM (New England Biolabs #S1402S).
4. β-mercaptoethanol.
5. 70% ethanol solution prepared with DEPC water.
6. Lyticase from *Arthrobacter luteus* (Sigma-Aldrich #L2524). Reconstituted in 1× PBS at 25 U/μL. Store 30 μL (600 U) aliquot at -20 °C.
7. Protease inhibitor cocktail tablets (Roche #04 693 132 001). Dissolve 1 tablet in 1 mL DEPC water for a 50× concentrated solution.
8. 1× Buffer B: 1.2 M Sorbitol, 0.1 M potassium phosphate pH 7.5 (200 mL needed per sample, can be kept for a few month at 4 °C).

9. 28 mM  $\beta$ -mercaptoethanol, 0.06 mg/mL PMSF, protease inhibitor cocktail 1 $\times$ . Prepare fresh.
10. Formamide.
11. *Escherichia coli* tRNA (e.g., Roche #10 109 541 001).
12. Single-stranded DNA from salmon testes (ssDNA) 10 mg/mL (e.g., Sigma #D9156).
13. Solution of *E. coli* tRNA and salmon sperm single-stranded DNA in a 1:1 ratio (5 mg/mL each).
14. Sodium phosphate buffer 1 M, pH 7. Dissolve 11.9 g of sodium phosphate in 50 mL of DEPC water, adjust pH to 7 and make up to 100 mL with DEPC water, can be kept at room temperature.
15. Buffer F: 80% formamide, 10 mM sodium phosphate pH 7, freshly made, 15  $\mu$ L per sample.
16. Labeled probes, *see* Subheading 3.2.
17. Parafilm.
18. RNase free BSA 20 mg/mL.
19. Coplin Jars for 22  $\times$  22 mm coverslips.
20. Glass plate (16  $\times$  20 cm).
21. Forceps.
22. 20 $\times$  SSC. Dissolve 175.3 g NaCl and 88.2 g of sodium citrate in 1 L of water. Adjust pH to 7 with HCl. Treat with 0.1% DEPC overnight and autoclave. Keep at room temperature.
23. 2 $\times$  SSC, 40% formamide in DEPC water, freshly made. Can be stored at 4  $^{\circ}$ C for 2 days.
24. Buffer H: 4 $\times$  SSC, 20 mM VRC, 4  $\mu$ g/ $\mu$ L BSA, 20 U RNase OUT, freshly made.
25. 2 $\times$  SSC, from 20 $\times$  SSC diluted in DEPC water.
26. 1 $\times$  SSC, from 20 $\times$  SSC diluted in DEPC water.
27. Triton X-100 (Sigma-Aldrich).
28. Frosted microscope slides (76  $\times$  26  $\times$  1 mm).
29. Transparent nail polish.
30. 20 $\times$  PBS. Dissolve in 1 L water 2.88 g of sodium phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>), 160 g sodium chloride (NaCl), 4 g potassium phosphate dibasic (KH<sub>2</sub>PO<sub>4</sub>) and 4 g potassium chloride (KCl), adjust pH to 7.2 with sodium hydroxide (NaOH). Treat with 0.1% DEPC overnight and autoclave. Store at room temperature.
31. 2 $\times$  SSC, 0.1% Triton X-100, freshly made. Can be stored at 4  $^{\circ}$ C for 2 days.



## 2.5 Fluorescence In Situ Hybridization on Human Cells

1. Coplin jars.
2. 32% Paraformaldehyde, electronic microscopy grade, single-use sealed ampules (Electron Microscope Sciences) (*see Note 1*).
3. 1× phosphate buffered saline (PBS), diluted from 20× PBS in DEPC water, *see* Subheading 2.4, item 4.
4. Triton X-100 (Sigma-Aldrich).
5. Fixative solution: 3.2% paraformaldehyde in 1× PBS.
6. Permeabilization solution: 0.5% Triton X-100 in 1× PBS.
7. Formamide.
8. Dextran sulfate, make a 50% W/V solution in DEPC water.
9. *Escherichia coli* tRNA (Roche #10109541001). Dissolved at 10 mg/mL in DEPC water, store at  $-20^{\circ}\text{C}$ .
10. Single-stranded DNA from salmon testes (ssDNA) 10 mg/mL (e.g., Sigma-Aldrich #D9156).
11. Bovine serum albumin RNase free (e.g., Roche #10711454001).
12. Vanadyl Ribonucleoside Complex (VRC) 200 mM (New England Biolabs #S1402S).
13. 20× SSC, *see* Subheading 2.4, item 3.
14. 2× SSC in 40% formamide, store at  $4^{\circ}\text{C}$  for maximum 2 days.
15. Solution A, for ten coverslips: mix 80  $\mu\text{L}$  of formamide with 20  $\mu\text{L}$  of 20× SSC.
16. Solution B, for 10 coverslips: mix 20  $\mu\text{L}$  of BSA with 40  $\mu\text{L}$  of 50% dextran sulfate and 2  $\mu\text{L}$  200 mM of VRC, complete with 38  $\mu\text{L}$  of DEPC water.
17. Large glass plate (16 × 20 cm).
18. Parafilm.
19. Formamide.
20. Probes as described Subheading 3.2.
21. 20× SSC, *see* Subheading 2.4, item 3.
22. Frosted microscope slides.
23. Mounting media, *see* Subheading 3.4.
24. Transparent nail polish.
25. Formamide.
26. 1× PBS, diluted from 20× PBS, *see* Subheading 2.4, item 4.
27. Tween 20.
28. 20× SSC, *see* Subheading 2.4, item 3.
29. Wash solution 1: 2× SSC, 40% formamide in DEPC water.
30. 1× PBS, 0.1% Tween 20.

### 3 Methods

#### 3.1 Preparation of Poly-L-Lysine-Coated Coverslips

Yeast cells do not adhere stably on glass or plastic. Coverslips must be pretreated with a coating agent to assure that cells will stay on the coverslips during the hybridization and washing steps. We use poly-L-lysine, but alternatively concanavalin A can be used.

At the opposite, human cells adhere easily to glass but repeated washes during the process have the tendency to remove cells from coverslips and to alter normal cell shape. This can be prevented by using, as for yeast, poly-L-lysine-coated coverslips.

1. Coverslips are stripped using hydrochloric acid as follow: in a 1 L glass beaker, immerse a large number of coverslips with about 250 mL of HCl 0.1 N and cover it with aluminum foil.
2. Boil 30 min in a fume hood.
3. Let the beaker cool down and discard the HCl.
4. Wash the coverslips ten times with DEPC water. Leave at least 100 mL of water in the beaker after the last wash.
5. Cover the beaker filled with the coverslips with aluminum foil and autoclave it. Once autoclaved, the coverslips can be stored at 4 °C for several months. Alternatively, coverslips can be stored in 50 mL tubes filled with 75% ethanol, this has the advantage of a faster drying.
6. Place one coverslip in each well of a six-well plastic plate and let them dry.
7. Dilute poly-L-lysine at 0.01% in water, filtrate through 0.22 µm filter and apply 200–500 µL of diluted poly-L-lysine on each coverslip. The solution should cover at least 75% of the coverslip surface.
8. Incubate at room temperature for 5 min to let the poly-L-lysine adhere.
9. Aspirate the excess of poly-L-lysine and let dry completely. This will take around 3–4 h. Coverslips may stick to the bottom of the wells. During the washes, they will eventually get free.
10. Once dried, wash each coverslips with 3 mL sterile water three times. These washes are important to avoid excess adherence of yeast cells on the coverslips that will lead to aggregates formation instead of a single layer of individual cells.
11. After the last wash, rest each coverslip on the wall of the well to be sure they do not get stuck at the bottom while drying. Aspirate the excess and let dry completely.
12. The coverslips can be kept at room temperature for several months inside the six well plates.

### 3.2 Probe Design and Preparation

1. Using a Speed Vac, lyophilize 10  $\mu\text{g}$  of each probe in RNase-free Eppendorf tubes (do not heat).
2. Resuspend the probes in 35  $\mu\text{L}$  of sodium carbonate buffer.
3. Prepare the cyanine dye: dissolve the content of one vial in 30  $\mu\text{L}$  DEPC water. One vial can be used to label 2 probes. See manufacturer instruction/ protocol for more details.
4. Add 15  $\mu\text{L}$  of the reconstituted cyanine dye to each probe.
5. Incubate at room temperature for 16–24 h in the dark with occasional vortexing.
6. Purify the probe using a G25 oligo column.
  - (a) Prepare the column: open the top of the column and break the bottom seal. Remove the buffer by spinning the column in a 2 mL tube at  $2500 \times g$  for 1 min, discard the flow through.
  - (b) Place the column in a new 1.5 mL RNase free Eppendorf. Apply the 50  $\mu\text{L}$  labeled probe in the center of the column and centrifuge at  $2500 \times g$  for 4 min. Discard the column and determine the labeling efficiency of the probes (*see Note 2*).
7. Keep labeled probes at  $-80^\circ\text{C}$  in opaque RNase-free tubes.

### 3.3 Calculation of Fluorophore Incorporation

1. Measure the optical density at 260 nm to determine DNA concentration and at 552 nm (Cy3) or 650 nm (Cy5) for cyanine concentration. Usually a 1/50 dilution is needed and sufficient to avoid signal saturation.
2. Calculate the incorporation efficiency using the molecular extinction coefficient (MEC) of the oligonucleotide and cyanine dye. Given below is the calculation for Cy3 dye.
3. Cy3 labeling of probes by this method gives usually 70–90% incorporation efficiency, meaning that at least 3 modified nucleotides out of 5 have been labeled with Cy3. A reaction having a labeling efficiency below 60% is insufficient and must be redone.

$$[\text{Cy3}] = \frac{A_{552\text{nm}}}{\text{Cy3 MEC}}$$

$$[\text{Oligo}] = \frac{(A_{260\text{nm}} - 0.08(A_{552\text{nm}}))}{\text{Oligo MEC}}$$

$$\text{Incorporation efficiency} = \frac{[\text{Cy3}]}{[\text{Oligo}]}$$

\*The MEC of oligonucleotides is the sum of the MEC of each nucleotide.

	MEC ( $M^{-1}.cm^{-1}$ )
Cy3	150,000 (552 nm)
Cy5	250,000 (650 nm)
Adenosine	15,200 (260 nm)
Cytidine	7050 (260 nm)
Guanine	12,010 (260 nm)
Thymidine	8400 (260 nm)

### 3.4 Mounting Media Preparation

Mounting media can be easily made; it contains p-phenylenediamine as antifading agent. It is also convenient to add a DNA staining dye such as DAPI or Hoechst.

1. Dissolve 10 mg of p-phenylenediamine in 1 mL of  $10\times$  SSC.
2. Add DAPI at the final concentration of 1  $\mu$ g/mL.
3. Once completely dissolved, add 9 mL of glycerol, store at 20 °C in opaque tubes. Discard when it becomes dark brown.

### 3.5 Fluorescent In Situ Hybridization on Yeast Cells

#### 3.5.1 Fixation and Spheroplasting of Yeast Cells

Fixation of the cells is performed by cross-linking with paraformaldehyde (PFA). PFA is a small chemical compound that diffuses easily and rapidly inside the cells, and which cross-links protein–protein, protein–nucleic acid, and nucleic acid–nucleic acid interactions. Once the cells are fixed, keep them on ice to avoid proteolytic and RNase activities. Yeast cells have a cell wall outside their cell membrane that needs to be degraded to allow the probes to diffuse inside the cell. Several enzymes can be used to degrade this cell wall, such as lyticase, chitinase, zymolase, and gluculase. We typically use the lyticase from *Arthrobacter luteus* that hydrolyzes poly- $\beta$ -(1  $\rightarrow$  3)-glucose in the cell wall glucan.

1. Grow yeasts in 50 mL of media until mid-log phase growth is reached ( $OD_{600}$  between 0.2 and 0.4, starting from 0.05 to 0.1) (*see Note 3*).
2. Add 6.25 mL of 32% paraformaldehyde directly in the culture flask. Incubate at room temperature for 45 min with gentle shaking (*see Note 4*).
3. Transfer the fixed cells in a 50 mL tube and centrifuge at 4 °C at  $2500 \times g$  for 4 min.
4. Wash cell pellet three times with ice-cold buffer B to remove the culture media and the PFA. Do not resuspend the cell pellet when washing.
5. Resuspend gently the cell pellet in 1 mL of spheroplasting buffer by pipetting (no vortex).
6. Add 600 U of lyticase and mix by pipetting.

7. Incubate for 8 min in a 30 °C water bath. Invert the tube each minute to avoid sedimentation of the cells (*see Note 5*).
8. Pellet the cells at 4 °C at 2500 × *g* for 4 min.
9. Wash the cell pellet with 1 mL of ice-cold buffer B.
10. Resuspend gently the cell pellet in 750 μL of ice-cold buffer B.
11. Gently pipet 100 μL of the spheroplasting reaction on each prepared coverslips in a six well plate.
12. Let cells adhere for 30 min at 4 °C.
13. Very gently, wash the coverslips with 3 mL of ice-cold buffer B (*see Note 6*).
14. Dehydrate the spheroplasts by adding 5 mL of ice-cold 70% ethanol prepared with DEPC water. Incubate for at least 30 min at −20 °C before performing the hybridization. Spheroplasts can be kept at −20 °C for a maximum of 2 months (*see Note 6*).

### 3.5.2 Probe Preparation

1. Prepare two coverslips per experiment in case one of the two breaks during handling. The following protocol stands for one coverslip and can be scaled up.
2. In a 1.5 mL tube, add 1 ng of each labeled probe and complete up to 10 μL with DEPC water.
3. Add 4 μL of a 5 mg/mL of a solution 1:1 of *E. coli* tRNA and salmon sperm single-stranded DNA.
4. Lyophilize in a speed vac.
5. Resuspend the lyophilized probes in 12 μL of buffer F.
6. Keep at room temperature in the dark.

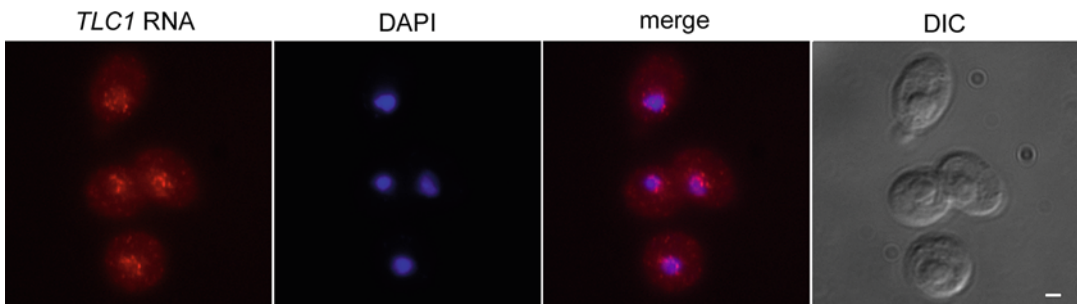
### 3.5.3 Hybridization

1. Cover a glass plate with a layer of Parafilm. Use the back part of a forceps to make the Parafilm adhere to the glass. At this point, all manipulation should be done in dim light.
2. Place the coverslips in a Coplin jar. Make sure the yeast-coated face of each coverslips face the same side (*see Note 7*).
3. Wash two times with 8 mL of 2× SSC for 5 min at room temperature with gentle shaking to rehydrate the cells.
4. Incubate in 2× SSC, 40% formamide for 5 min at room temperature (*see Note 8*).
5. Heat the probes at 95 °C for 3 min.
6. Dilute the probes with 12 μL of buffer H.
7. Drop the probes (24 μL) on the layer of Parafilm on the glass plate.
8. Lay the coverslip on the probes, with the spheroplasts side facing the probes (*see Note 9*).

- Cover the coverslips with another layer of Parafilm. Use the back of the forceps to seal the two layer of Parafilm together, making a humidified hybridization chamber. Wrap the glass plate with aluminum foil and incubate overnight at 37 °C.

#### 3.5.4 Washes and Coverslips Mounting

- All manipulation should be done in dim light. Cover the Coplin jar in aluminum foil during incubations.
- Gently put back the coverslips from the glass plate in a Coplin jar with the spheroplasts side facing the same direction.
- Wash two times with 8 mL of preheated 2× SSC, 40% formamide at 37 °C for 15 min.
- Wash with 2× SSC, 0.1% Triton X-100 for 15 min at room temperature with gentle shaking.
- Wash two times with 1× SSC 15 min at room temperature with gentle shaking (*see Note 10*).
- Incubate in 8 mL of 1× PBS, 1 ng/mL DAPI at room temperature for 2 min.
- Drop 7 µL of mounting media on a microscope slide.
- Gently take the coverslip and remove the excess of liquid at the bottom of the coverslip with an absorbent tissue.
- Lay the coverslip on the mounting media, with the spheroplast side facing the mounting media.
- Let stand for 2–3 min and then seal the coverslip by applying nail polish on the sides. Let dry. Microscope slides can be stored in the dark at –20 °C for up to 3 months, but image acquisition should be performed as soon as possible to assure good image quality (Fig. 1).



**Fig. 1** FISH on budding yeast *TLC1* RNA. Fluorescent in situ hybridization on budding yeast cells with *TLC1* specific probes. Endogenous *TLC1* RNA foci are visible in the nucleus of yeast cells. Acquisition was performed with a Nikon upright epifluorescence microscope and a CoolSnap Photometrics CCD camera. DAPI: nuclear staining, DIC: differential interference contrast. Scale bar = 1 µm

### 3.6 Fluorescent In Situ Hybridization on Human Cells

To ensure a good cell adherence and a normal shape, cell plating should be made the day before the FISH. Also it is important to reach around 70–80% of cell confluency the day of the fixation, thus the number of cells to plate must be determined empirically for each cell type by plating several cell densities.

#### 3.6.1 Cell Fixation/Permeabilization

All manipulations are done at room temperature and under a fume hood since paraformaldehyde is a toxic chemical.

1. Remove coverslips from the 6 well tissue culture plates containing the cells and place them in a Coplin jar (*see Note 11*).
2. Fill the Coplin jar with 1× PBS for 5 min, gently rock in your hands few seconds to remove any trace of serum.
3. Remove the PBS and place 7–8 mL of fixative solution, let stand for 15 min.
4. Pour off the fixative solution and wash two times with 7–8 mL of 1× PBS to ensure a complete removal of paraformaldehyde.
5. Remove the last wash solution, then permeabilize the cells with 7–8 mL of permeabilization solution for 5 min.
6. Wash one time with 1× PBS.
7. At this point, the cells can be stored in 1× PBS at 4 °C for a few days. For long-term storage, remove the PBS and add 2 mL of 70% ethanol. Keep at 4 °C, which allows the cells to remain usable for a few weeks.

#### 3.6.2 Probe Preparation

1. All the following steps should be performed in dim light in order to prevent dye photobleaching a thus a loss of signal.
2. For ten coverslips, mix 10 µL of 10 mg/mL ssDNA and 10 µL of 10 mg/mL *E. coli* tRNA in a 1.5 mL Eppendorf tube. Add 200 ng of probe per target RNA. If several probes are used, a total of 200 ng of the pool of probes should be used.
3. Lyophilize the probes using a speed vac. It should take around 20 min.
4. Add 100 µL of solution A, ensure a proper solubilization by vortexing the solution, then place the tube in a heater block at 95 °C for 3 min.
5. Chill the tube on ice for about 5 min.
6. Add 100 µL of solution B, mix well by vortexing.
7. Probes can be stored on ice for few hours.

#### 3.6.3 Hybridization

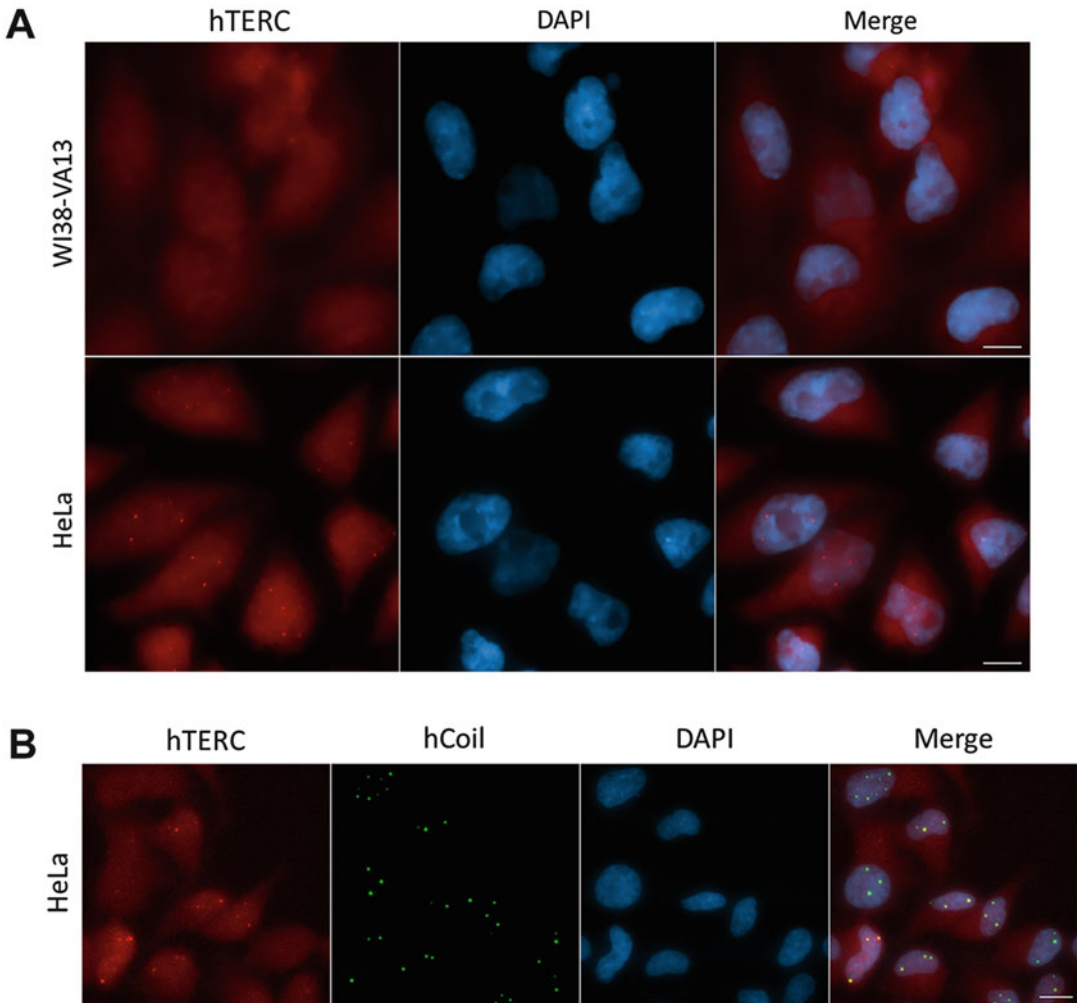
Hybridization is performed in a buffer containing formamide, in a moist chamber made of Parafilm. If detection of endogenous hTERC is required, hybridization should be performed overnight. If hTERC is overexpressed in the cells, incubation time can be reduced to 2 h.

1. Take coverslips prepared Subheading 3.6.1, they are already in a Coplin jar.
2. Remove the 1× PBS or the 70% ethanol and fill the Coplin jar with 7–8 mL of wash solution, incubate at room temperature for 5 min.
3. Meanwhile clean the glass plate and lay on a piece of Parafilm. Make the Parafilm adhere to the glass plate by scrubbing it with a forceps or a pencil.
4. Using a marker, criss-cross the Parafilm to define zones for coverslips and note the conditions.
5. Drop 20 µL of each probe onto the Parafilm
6. One by one, remove coverslips from the Coplin jar, drain excess of buffer on a paper and carefully deposit on the Parafilm, cells facing the Parafilm. Avoid contact between the coverslips (*see Note 12*).
7. Lay down a new piece of Parafilm onto the first one, seal using forceps or pencil to form a perfect moist chamber.
8. Wrap the glass plate in aluminum foil to protect from light and place at 37 °C overnight.

#### 3.6.4 Washes and Coverslip Mounting

1. Prewarm Coplin jars and 50 mL of wash solution at 37 °C for at least 30 min.
2. Take the glass plate containing coverslips out of the incubator and move one by one the coverslips into the prewarmed Coplin jars. Incubate for 30 min at 37 °C.
3. Remove wash solution and add a new volume of prewarmed wash solution, incubate at 37 °C for another 30 min.
4. Remove wash solution and fill the Coplin jar with 1× PBS, 0.1% Tween 20 containing 1 µg/mL of DAPI. Incubate at room temperature for 5 min.
5. Remove wash solution and fill the Coplin jar with 1× PBS, 0.1% Tween 20. Incubate at room temperature for 5 min.
6. Remove wash solution and fill the Coplin jar with 1× PBS. Incubate at room temperature for 5 min (*see Note 10*).
7. Annotate microscope slides and drop 20 µL of mounting media for each coverslip.
8. Remove coverslips from the Coplin jar, rapidly drain the PBS on a paper towel and apply onto the microscope slide, cell facing the mounting media.
9. Seal coverslips with few drops of nail polish, let dry in the dark.
10. To remove any traces of salt, coverslips can be whipped with a microfiber paper soaked with a drop of lens cleaning solution.
11. Slides are now ready to image. They can be kept up to a week in the dark at 4 °C or a few month at –20 °C. Human TERC foci





**Fig. 2** FISH on human telomerase (hTERC) RNA. **(a)** HeLa cells (hTERC positive cells) and WI38-VA13 cells (hTERC negative cells) were plated and processed for hTERC FISH, as described in this protocol. Endogenous hTERC in HeLa cells appears as foci in the nucleus. **(b)** hTERC FISH was performed on HeLa cells followed by immunofluorescence with anti-hCoilin-Alexa 647 coupled antibody (hCoil) to detect Cajal bodies. Acquisition was performed with a Zeiss upright epifluorescence microscope and Ixon EM-CCD camera. Images are maximal projections. Scale bar = 5  $\mu$ m

can be imaged using an epifluorescence or a confocal microscope (Fig. 2a). Colocalization experiments can also be performed using antibodies (to detect Coilin, as in Fig. 2b).

## 4 Notes

1. Formaldehyde quality is very important to preserve small details and low contrasts by FISH. We strongly recommend using EM-grade formaldehyde from Electron Microscope Science.

2. During probe purification, dropping probe solution far from the center of the G25 Sephadex column will result in the contamination of the probes with free dye molecules, leading to a false incorporation calculation. In this case, a second purification can be done with a new column.
3. When using yeast strains having the *ade2* marker, adding 20 µg/mL of adenine in the media decreases autofluorescence coming from the accumulation of phosphoribosylaminoimidazole, the adenine intermediate giving the pink color to *ade2* yeast colonies. For the same reason, yeast grown in YEPD medium present more autofluorescence (YEPD is poor in adenine).
4. To image simultaneously a GFP-tagged protein and the *TLC1* RNA by FISH in fixed yeast cells, grow the cells at room temperature to get a stronger GFP fluorescence. Also, spin the cells at  $2500 \times g$  for 4 min and resuspend the cells in 50 mL of RNase-free PBS before adding the PFA. Keeping the pH of the yeast culture at 7.0 during fixation helps preserve the GFP fluorescence [17].
5. Spheroplasting is a step that needs to be optimized by each user. If the cell wall is not enough degraded, the probes will not diffuse inside the cells. On the other hand, if the cell wall is too much digested, the cell may lose its integrity and shape. Each reaction needs to be followed using a differential interference contrast (DIC) microscopy. Typically each 2–3 min, apply 3.5 µL of the spheroplasting reaction between a coverslip and a microscopy slide, and observe under the 20× objective. Cells that still have an intact cell wall appear brighter and have a halo around them. Cells that have a degraded cell wall lose the halo and turn darker. When more than 75% of the cells show a degradation of their cell wall, stop the spheroplasting reaction.
6. Drop the liquid carefully on the side of the well, avoiding putting it directly on the coverslip to make sure not to disperse the spheroplasts.
7. Mark the side of the Coplin jar with a marker pen to remember on which face are the spheroplasted yeasts.
8. The concentration of formamide can be adjusted to have better signal if needed. If the background is too high, the concentration of formamide can be adjusted up to 50%.
9. Avoid making bubbles to assure a uniform hybridization and make sure not to move the coverslips once laid down to avoid displacing the spheroplasts from the coverslips.
10. After this wash, a standard immunofluorescence protocol can be performed if a colocalization experiment with a protein is needed [18].

11. Mark the side of the Coplin jar with a marker pen to remember on which face are the cells.
12. Avoid making bubbles to assure a uniform hybridization and make sure not to move the coverslips once laid down to avoid displacing the cells from the coverslips.

---

## Acknowledgments

This work was funded by a grant from the Canadian Institutes for Health Research (CHIR) MOP-89768 to P.C. M.L is supported by a fellowship from the Fonds de Recherche du Québec-Santé (FRQS). P.C holds a Research Chair from FRQS.

## References

1. Schmidt JC, Cech TR (2015) Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes Dev* 29(11):1095–1105. doi:[10.1101/gad.263863.115](https://doi.org/10.1101/gad.263863.115)
2. Gallardo F, Olivier C, Dandjinou AT, Wellinger RJ, Chartrand P (2008) TLC1 RNA nucleo-cytoplasmic trafficking links telomerase biogenesis to its recruitment to telomeres. *EMBO J* 27:748–757
3. Wu H, Becker D, Krebber H (2014) Telomerase RNA TLC1 shuttling to the cytoplasm requires mRNA export factors and is important for telomere maintenance. *Cell Rep* 8(6):1630–1638. doi:[10.1016/j.celrep.2014.08.021](https://doi.org/10.1016/j.celrep.2014.08.021)
4. Teixeira TM, Förstemann K, Gasser SM, Lingner J (2002) Intracellular trafficking of yeast telomerase components. *EMBO Rep* 3(7):652–659
5. Ferrezuelo F, Steiner B, Aldea M, Fitcher B (2002) Biogenesis of yeast telomerase depends on the Importin Mtr10. *Mol Cell Biol* 22(17):6046–6055. doi:[10.1128/mcb.22.17.6046-6055.2002](https://doi.org/10.1128/mcb.22.17.6046-6055.2002)
6. Pflingsten JS, Goodrich KJ, Taabazuing C, Ouenzar F, Chartrand P, Cech TR (2012) Mutually exclusive binding of telomerase RNA and DNA by Ku alters telomerase recruitment model. *Cell* 148:922–932
7. Nguyen D, Grenier St-Sauveur V, Bergeron D, Dupuis-Sandoval F, Scott Michelle S, Bachand FA (2015) Polyadenylation-dependent 3' end maturation pathway is required for the synthesis of the human telomerase RNA. *Cell Rep* 13(10):2244–2257. doi:[10.1016/j.celrep.2015.11.003](https://doi.org/10.1016/j.celrep.2015.11.003)
8. Jady BE, Bertrand E, Kiss T (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J Cell Biol* 164(5):647–652. doi:[10.1083/jcb.200310138](https://doi.org/10.1083/jcb.200310138)
9. Zhu Y, Tomlinson RL, Lukowiak AA, Terns RM, Terns MP (2004) Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol Biol Cell* 15(1):81–90. doi:[10.1091/mbc.E03-07-0525](https://doi.org/10.1091/mbc.E03-07-0525)
10. Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, Terns MP, Artandi SE (2009) A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* 323(5914):644–648. doi:[10.1126/science.1165357](https://doi.org/10.1126/science.1165357)
11. Gallardo F, Laterreur N, Cusanelli E, Ouenzar F, Querido E, Wellinger RJ, Chartrand P (2011) Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres. *Mol Cell* 44(5):819–827
12. Bajon E, Laterreur N, Wellinger Raymond J (2016) A single Templating RNA in yeast telomerase. *Cell Rep* 12(3):441–448. doi:[10.1016/j.celrep.2015.06.045](https://doi.org/10.1016/j.celrep.2015.06.045)
13. Jady BE, Richard P, Bertrand E, Kiss T (2006) Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Mol Biol Cell* 17(2):944–954. doi:[10.1091/mbc.E05-09-0904](https://doi.org/10.1091/mbc.E05-09-0904)
14. Tomlinson RL, Abreu EB, Ziegler T, Ly H, Counter CM, Terns RM, Terns MP (2008) Telomerase reverse transcriptase is required for the localization of telomerase RNA to cajal bodies and telomeres in human cancer cells. *Mol Biol Cell* 19(9):3793–3800. doi:[10.1091/mbc.E08-02-0184](https://doi.org/10.1091/mbc.E08-02-0184)

15. Abreu E, Aritonovska E, Reichenbach P, Cristofari G, Culp B, Terns RM, Lingner J, Terns MP (2010) TIN2-tethered TPP1 recruits human telomerase to telomeres in vivo. *Mol Cell Biol* 30(12):2971–2982. doi:[10.1128/mcb.00240-10](https://doi.org/10.1128/mcb.00240-10)
16. Stern JL, Zyner KG, Pickett HA, Cohen SB, Bryan TM (2012) Telomerase recruitment requires both TCAB1 and Cajal bodies independently. *Mol Cell Biol* 32(13):2384–2395. doi:[10.1128/mcb.00379-12](https://doi.org/10.1128/mcb.00379-12)
17. Gallardo F, Chartrand P (2011) Visualizing mRNAs in fixed and living yeast cells. In: Gerst EJ (ed) *RNA detection and visualization: methods and protocols*. Humana Press, Totowa, NJ, pp 203–219. doi:[10.1007/978-1-61779-005-8\\_13](https://doi.org/10.1007/978-1-61779-005-8_13)
18. Chartrand P, Bertrand E, Singer RH, Long RM (2000) Sensitive and high-resolution detection of RNA in situ. *Methods Enzymol* 318:493–506

## Methods to Study Repeat Fragility and Instability in *Saccharomyces cerevisiae*

Erica J. Polleys and Catherine H. Freudenreich

### Abstract

Trinucleotide repeats are common in the human genome and can undergo changes in repeat length. Expanded CAG repeats have been linked to over 14 human diseases and are considered hotspots for breakage and genomic rearrangement. Here, we describe two *Saccharomyces cerevisiae* based assays that evaluate the rate of chromosome breakage that occurs within a repeat tract (fragility), and a PCR-based assay to evaluate tract length changes (instability). The first fragility assay utilizes end-loss and subsequent telomere addition as the main mode of repair of a yeast artificial chromosome (YAC). The second fragility assay relies on the fact that a chromosomal break stimulates recombination-mediated repair. In addition to understanding the role of fragility at repetitive DNA sequences, both assays can be modified to evaluate instability of a CAG repeat using a PCR-based assay. All three assays have been essential in understanding the genetic mechanisms that cause chromosome breaks and tract-length changes at unstable repeats.

**Key words** Chromosome break, Fragility, Stability, Yeast artificial chromosome (YAC), CAG repeat

---

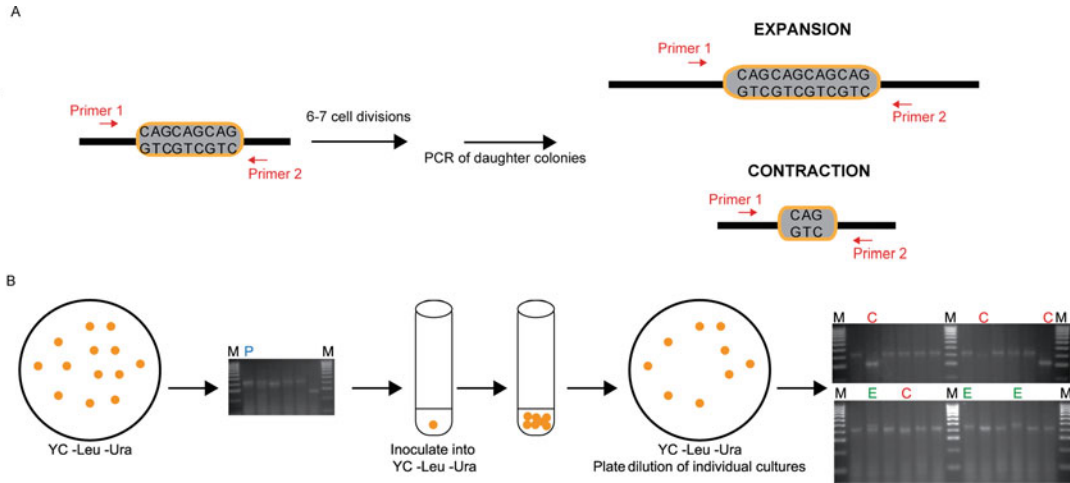
## 1 Introduction

Stretches of repetitive DNA sequence are common in eukaryotic genomes and are capable of forming alternative DNA structures, such as hairpin loops. These alternative DNA structures effectively function as a physical barrier, making it difficult for error free DNA replication and repair to occur, and can result in changes in repeat length, chromosomal breakage, and genome instability. One type of repetitive DNA sequence that is common in the human genome are CNG repeats, which when expanded, have been causally associated with over 16 human diseases including Huntington's disease (HD), myotonic dystrophy (DM), spinocerebellar ataxias (SCAs), and fragile X syndrome (FXS) [1, 2].

Long regions of CAG/CTG and CGG/GCC trinucleotide repeats break and undergo aberrant repair at a higher frequency than nonrepetitive DNA, which can result in mutation and changes in repeat copy number. Breakage of the DNA strands (referred to

herein as “fragility”) and changes in repeat copy number (referred to herein as “instability”) are two sides of the same coin. For example, if normal replication or repair is impaired when encountering a repeat tract, a stalled fork, nick, gap, or double-strand break (DSB) can occur. If the fork is restarted or if there is healing within the nick, gap or break, repeat units can be lost or gained. Alternatively, the DNA lesion may be repaired by a less conservative form of repair that leads to deletion of flanking sequence, loss of a chromosome arm, or other chromosome rearrangements. This dramatic loss of genetic material can be easily monitored in *S. cerevisiae* by loss of a neighboring genetic marker, and is indicative that a break in the chromosome occurred. Such a marker loss assay is an underestimate of the “true” rate of breakage, as it only measures the incidents that are healed after loss of flanking DNA sequence, and not those that heal within the repeat, or those that fail to repair resulting in cell death. Nonetheless, such genetic fragility assays are very useful for comparative studies and can be used to determine if the presence of a gene or condition is involved in replication or repair in the context of repetitive DNA. In contrast, a physical assay for trinucleotide repeat instability captures all the events that result in loss or gain of the number of repeats (at least those of a size that can be detected by the sensitivity of the assay), but will usually miss events that result in loss of flanking DNA sequence that includes one of the PCR primer binding sites.

Because of the clonal nature of yeast growth, repeat instability can easily be monitored over multiple generations. The PCR-based instability assay (Subheading 3.1) ideally begins with a parent colony with a known starting repeat tract length, determined by PCR using a subset of cells in the colony as a template (colony PCR; *see* Subheading 3.2). Subsequently, the cells in the parent colony are grown for a set number of cell divisions in liquid media to allow repeat tract length changes, plated and allowed to grow into daughter colonies. Colony PCR is then used to determine the repeat tract length in multiple individual daughter colonies in order to obtain a frequency of instability (Fig. 1). The goal is to use the tract length in the daughter colony as an output of the length present in the cell at the time of plating (which will contain one or two genomes, depending on cell cycle phase). This method is set up to ignore any changes that likely occur during the growth of the daughter colony, and only count the primary tract length changes that occurred in the cells during logarithmic growth in culture. It should be noted that there are other available genetic assay systems that can be used to select for repeat expansion or contraction events which are covered in this issue and previous papers [3, 4]. Such selection systems differ from the PCR-based instability assay as they capture only the events that yield a phenotype in the selective system. However they allow for the study of rare repeat length changes, such as those that occur in short (premutation) length CAG or



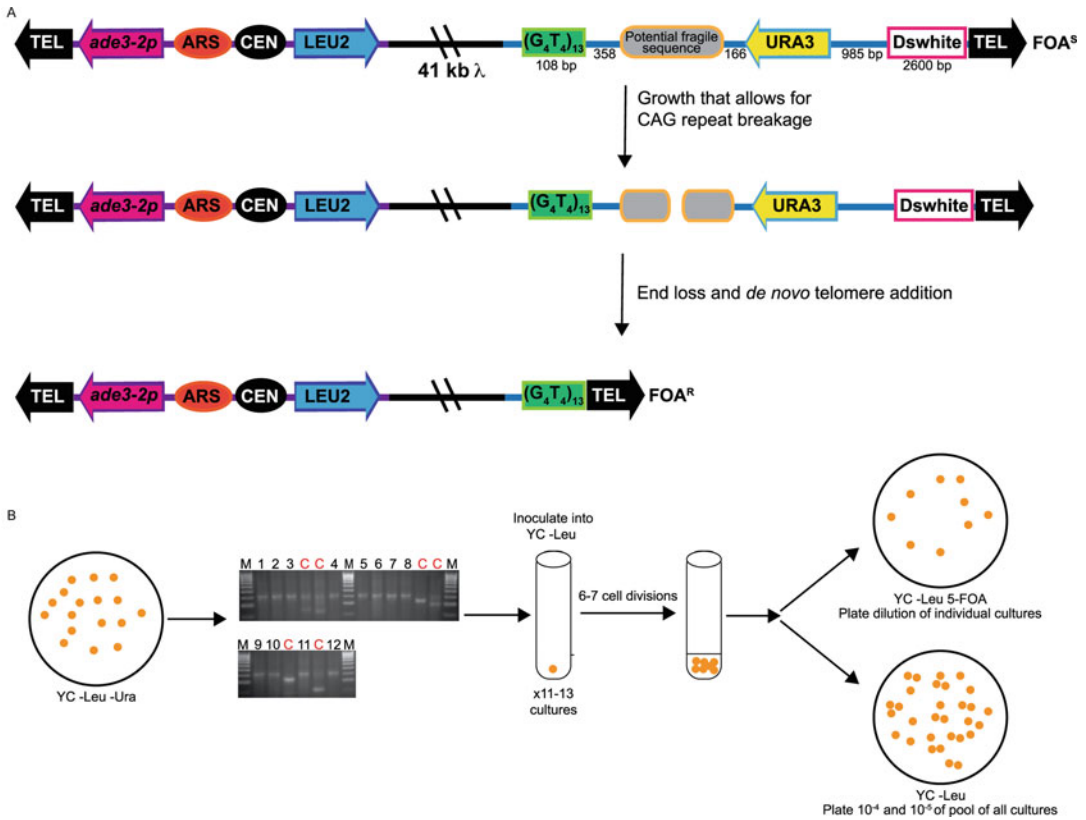
**Fig. 1** Schematic of CAG repeat instability assay (a) In this system, CAG repeats have either been integrated into a (CAG) $n$ -URA3 YAC (YAC CF1) or another chromosomal location. Colony PCR across the repeat is utilized to monitor the frequency of expansions and contractions in daughter cells. (b) In the (CAG) $n$ -URA3 YAC instability assay, strains are plated for single colonies on YC -Leu -Ura plates and individual parent colonies are tested for CAG tract length. Parent colonies of the correct tract length are then grown in YC -Leu -Ura for 6–7 cell divisions before plating each individual culture for single colonies on YC -Leu -Ura plates. Finally PCR is performed on approximately 25 daughter colonies for each parent colony to assess the frequency of changes in tract length sizes compared to the starting repeat size. At least four parent colonies should be tested for a total of at least 100 daughter colonies. *P* parental, *C* contraction, *E* expansion, *M* marker

CGG repeats (e.g., the Lahue lab system) or for large-scale repeat expansions (e.g., the Mirkin lab system).

To determine the rate at which breakage occurs within a repeat tract, two unique yeast systems have been developed. The first is a yeast artificial chromosome (YAC) system wherein the YAC VS5 is modified such that the potential fragile sequence is integrated between a telomere seed sequence ( $G_4T_4$ ) and a *URA3* marker gene [5, 6] (Fig. 2). The version of this YAC with an integrated CAG repeat tract has been termed the (CAG) $n$ -URA3 YAC, also referred to in the literature as YAC CF1. One feature of this YAC system is that other modifications are easily made by replacing the right end of the YAC (after  $G_4T_4$ ) with the desired sequence. We have used this technique to add additional genetic markers, promoters, terminators, etc. in addition to modifying the type of potential fragile element. For example, other derivatives of this YAC have been made that contain the CAG repeat in the opposite orientation (CTG) $n$  [7], AT repeats that are present in the FRA16D common fragile site [8], the expanded ATTCT repeat present in SCA10 [9], and inverted repeats [10].

Another important feature of the YAC is that it contains little homology to any of the natural yeast chromosomes. This feature, together with the telomere seed sequence placed proximal to the potential fragile sequence, ensures that the primary mode of





**Fig. 2** Schematic of the YAC breakage assay (a) In this system, a potential fragile sequence (such as an expanded CAG repeat tract) is integrated onto a yeast artificial chromosome between a telomere seed sequence ( $G_4T_4$ ) and the *URA3* gene. Breaks that occur within the fragile sequence are subject to resection and telomere addition, which results in loss of the *URA3* gene and renders cells 5- $FOA^R$ . The YAC additionally contains a *LEU2* marker gene, which allows for maintenance of the YAC, a centromere (*CEN4*) and an origin of replication (*ARS1*). The blue line indicates the pYIP5 plasmid backbone, the black line indicates lambda DNA, and the purple line corresponds to a pUC18 plasmid backbone. (b) In the YAC end loss assay, strains are plated for single colonies on YC -Leu -Ura plates and 11–13 individual colonies of the correct tract length (if testing an unstable repeat sequence) are grown in YC -Leu for 6–7 cell divisions. A portion of each individual culture is plated onto YC -Leu 5- $FOA$  plates, and a portion of the culture is pooled and serially diluted to obtain single colonies on YC -Leu, which serves as the total viable cell count

healing is by telomere addition rather than recombination. Healing via telomere addition is ideal as relatively few genes affect the efficiency of telomere addition. Known examples are the subunits of the enzyme telomerase, which are needed for telomere addition, and the Pif1 helicase, which inhibits promiscuous telomere addition [5]. Thus, in most cases an increase in YAC end loss is caused by increased breakage rather than decreased healing. The position of the potential fragile sequence on the YAC is such that on the proximal side of the sequence there is a 108 bp backup ( $G_4T_4$ )<sub>13</sub> telomeric seed sequence from *Oxytricha*, which serves as an efficient substrate for *S. cerevisiae* telomerase, but does not recombine with

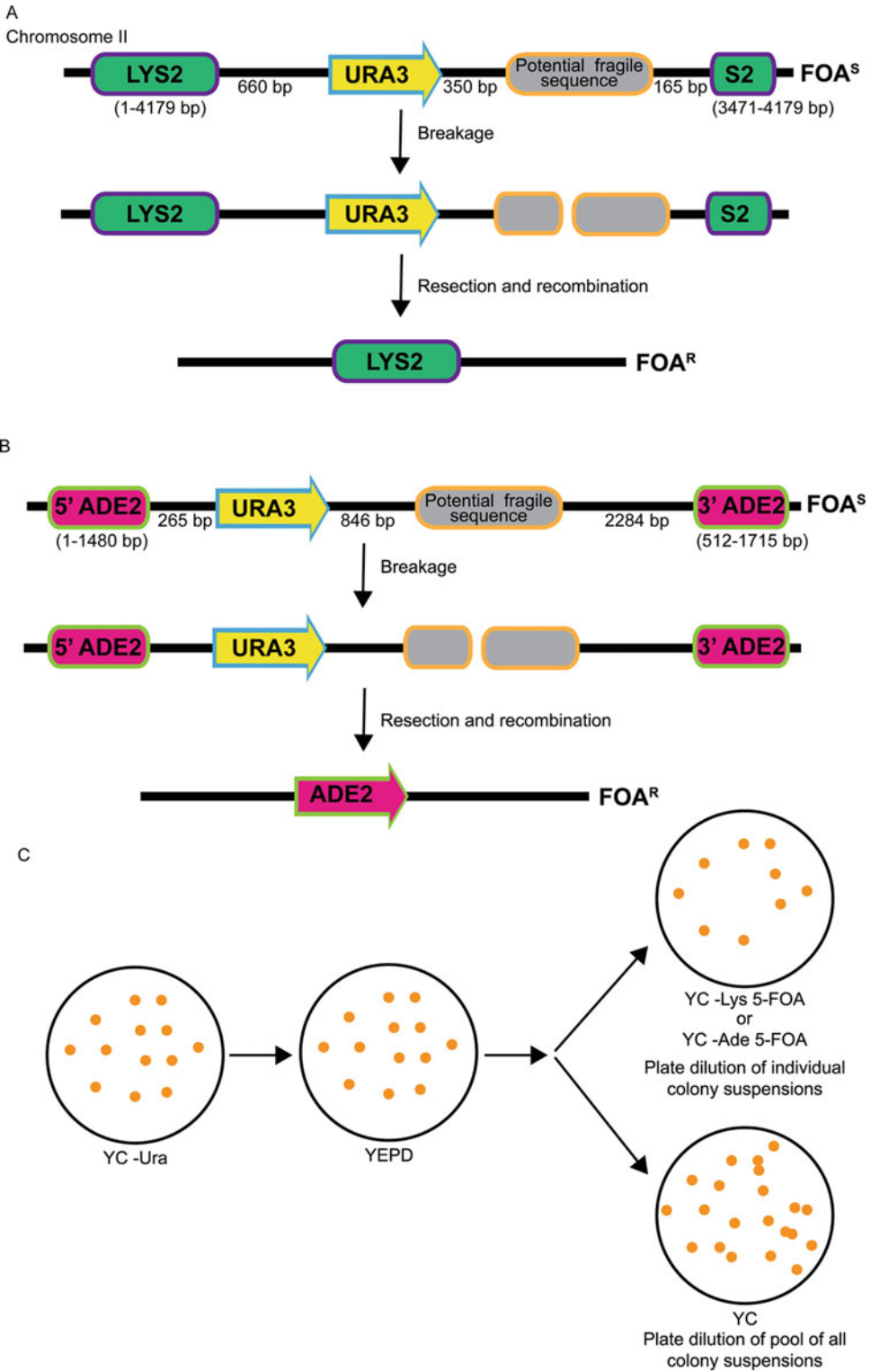


the natural yeast telomere sequence [11]. Therefore, if breakage within the fragile sequence occurs, resection and degradation of the broken end exposes the G<sub>4</sub>T<sub>4</sub> sequence as a template for telomerase. The G<sub>4</sub>T<sub>4</sub> seed provides an efficient pathway for healing of breakage events, which is different from another commonly used assay for end loss of chromosome V [12], where healing can occur by a variety of mechanisms including telomere addition onto short naturally occurring GT sequences or recombination with other chromosomes.

In the YAC fragility assay, breakage that results in loss of the right arm of the YAC renders cells *ura3*- and resistant to 5-flourorotic acid (5-FOA). The number of cells that are 5-FOA resistant serves as a measure of the rate of breakage, and the effect of a potential fragile sequence can be compared to a control sequence of the same length and base composition, or a YAC that does not contain the fragile sequence. One additional design consideration of the YAC is the distance between the *URA3* gene and the telomere, such that the placement of the *URA3* gene promoter is far enough away from the telomere not to be subject to telomere position effect. Elsewhere on the YAC, on the left arm there is a *LEU2* marker gene, which allows for maintenance of the YAC, a yeast origin of replication (*ARS1*), and a centromere (*CEN4*). Finally, an *ade3-2p* allele can be used to screen for cells that have obtained two YACs (which occurs occasionally by nondisjunction) in strains with an *ade3* background, but it is not necessary to use this feature for most applications.

For the original (CAG)<sub>n</sub>-*URA3* YAC, the CAG repeat tract is integrated such that the CAG repeat is on the lagging strand template, which has been shown to be the orientation less prone to repeat tract contractions [13]. For the CAG repeat, it has been established that there is a length dependent increase in breakage as measured by 5-FOA resistance [6]. This YAC system can also be used to study CAG repeat stability using similar conditions used for measuring breakage. These instability assays revealed a length dependent increase in contractions and expansions, though contractions occur with greater frequency than expansions in a replicating yeast population [6]. The CAG instability frequency measured on this YAC is similar to that measured at the chromosome II *LYS2* locus for the same orientation with respect to replication [13, 14], though it is known that factors such as distance from an origin, transcription level, and chromatin structure can alter instability frequencies (*see* ref. 1 for review).

The second system for measuring repeat fragility is a direct duplication recombination assay (DDRA), which is a chromosomal assay where a potential fragile sequence has been integrated adjacent to a *URA3* reporter and has been placed between a full length *LYS2* gene and a duplication of the 3' end of *LYS2* on chromosome II [15, 16] (Fig. 3). If breaks occur within the integrated sequence,



**Fig. 3** Schematic of the direct duplication recombination assay (DDRA). (a) In this chromosomal system, a potential fragile sequence and the *URA3* gene are integrated between a direct duplication of 708 bp of the 3'

resection occurs and repair results from single strand annealing of the duplicated regions of the *LYS2* gene. This results in the loss of the *URA3* gene and resistance to the drug 5-FOA. In this system, it was noted that there is a length dependent increase in 5-FOA resistance which was due to breaks within or very close to an integrated CAG repeat tract [15], and expanded CGG/CCG repeats also induced length-dependent fragility [16]. This system is ideal for understanding the role that recombination plays in the repair of breaks that occur within a repeat tract. Also, as it is located in the middle of a natural yeast chromosome, the replication program is known, and the location is far from elements such as the centromere and telomere that could influence chromatin structure or nuclear positioning. A modification of the DDRA utilizes the same principle, but with a duplication of the *ADE2* gene, designed such that the starting strain is Ade<sup>-</sup> and recombination products will become Ade<sup>+</sup> + 5-FOA<sup>R</sup> ([17]; Fig. 2b). This modification allows for selection of recombinants, eliminating other events that can generate FOA<sup>R</sup>, and the cassette can be integrated anywhere in the genome. It is important to note that many genes that are involved in replication or repair are also required for efficient recombination, which makes it more difficult to determine whether a gene is required to prevent fragility or facilitate healing in this system, as compared to the YAC end loss system.

These assay systems are unique as they can be utilized to monitor both fragility and instability of long stretches of repetitive DNAs under the same conditions (genomic location, strain background, transcriptional status, etc.). Taken together, both assays, the YAC end loss system and the chromosomal direct duplication recombination assay, have provided novel insight into the mechanisms of fragility and instability of repetitive DNA sequences and can be utilized to understand the breakage and repair of any sequence of interest.



**Fig. 3** (continued) end of the *LYS2* gene, rendering cells LYS<sup>+</sup> and FOA<sup>S</sup>. Breaks that occur within the fragile sequence are subject to resection and recombination, which renders cells Lys<sup>+</sup> and FOA<sup>R</sup>. **(b)** In the *ADE2* DDRA, a potential fragile sequence and the *URA3* gene are integrated between 5' and 3' portions of the *ADE2* gene, rendering cells Ade<sup>-</sup> and FOA<sup>S</sup>. Breaks that occur within the repeat are subject to resection and repair via single strand annealing between the 968 bp duplicated region of *ADE2*, thus rendering cells Ade<sup>+</sup> and FOA<sup>R</sup>. **(c)** Schematic of DDRA protocol. Cells are initially plated on YC -Ura media and then plated for single colonies on YEPD (or other nonselective media) to allow loss of *URA3*. Ten individual colonies are selected and each resuspended in water. A portion of each colony suspension is then plated onto ten separate 5-FOA plates (-Lys or -Ade, depending on the assay) and 100  $\mu$ L of each colony suspension is pooled, diluted and plated for single colonies onto YC plates for a total cell count

## 2 Materials

### 2.1 Determination of Repeat Instability Frequency

1. Starting colonies that have been checked for tract length by colony PCR (*see* Subheading 3.2).
2. OD spectrophotometer.
3. Cuvettes.
4. 5 mL glass test tubes.
5. P20 or P200 pipette tips.
6. Roller drum at 30 °C.
7. Media (standard yeast media, *see* ref. 18 for recipes):
  - For YAC repeat instability assays:
    - (a) YC -Leu -Ura liquid media.
    - (b) YC -Leu -Ura plates.
  - For DDRA repeat instability assays:
    - (c) YC -Ura plates.

### 2.2 Colony PCR to Determine CAG Repeat Tract Length

1. Taq polymerase.
2. Taq polymerase buffer.
3. 20 mM magnesium sulfate (MgSO<sub>4</sub>) solution.
4. 5× CG buffer (Empire Genomics cat# IDL028).
5. 10 mM dNTPs.
6. Primers that span the CAG repeat (*see* Note 9).
  - (a) Forward primer: 5' CCC AGG CCT CCA GTT TGC 3' (starts 74 bp before CAG repeat on (CAG)<sub>n</sub>-URA3 YAC).
  - (b) Reverse primer: 5' TAA TAC GAC TCA CTA TAG GG 3' (starts 79 bp after CAG repeat on (CAG)<sub>n</sub>-URA3 YAC).
7. 2% Metaphor agarose or other high resolution gel system.
8. Refrigerator.
9. Power supply.
10. Electrophoretic gel box.
11. 10 mg/mL ethidium bromide.

### 2.3 Determination of Chromosome Fragility by a YAC end Loss Assay

1. Starting colonies that have been checked for repeat tract length by colony PCR (*see* Subheading 3.2 for CAG repeats).
2. OD spectrophotometer.
3. Cuvettes.
4. 5 mL glass test tubes.
5. P20 or P200 pipette tips.

6. Roller drum at 30 °C.
7. Media (standard yeast media, *see* ref. 18 for recipes):
  - (a) YC -Leu liquid media.
  - (b) YC -Leu -Ura plates.
  - (c) YC -Leu plates.
  - (d) YC -Leu 5-FOA plates.

**2.4 Direct Duplication Recombination Assay (DDRA)**

1. Media (standard yeast media, *see* ref. 18 for recipes):
  - (a) YC -Ura plates.
  - (b) YEPD plates.
  - (c) YC -Lys 5-FOA plates (*LYS2* assay) or YC -Ade 5-FOA plates (*ADE2* assay).
  - (d) YC plates.

---

## 3 Methods

### 3.1 Determination of Repeat Instability Frequency

1. Plate or streak for single colonies on YC -Leu -Ura, which will select against breakage events, from a master patch (*see* Note 1).
2. Perform colony PCR (*see* Subheading 3.2) on 3–5 colonies to identify two parent colonies with the desired CAG tract length that have no other visible repeat sizes. The colonies should ideally be less than 2 weeks old. Usually two separate colonies from each strain are done in parallel.
3. Using a pipet tip, suspend a small amount of the yeast colony (from the same region of the colony tested for colony PCR) into 1 mL of YC -Leu -Ura media. The ideal starting OD<sub>600</sub> is between 0.02 and 0.04 (*see* Note 2).
4. Grow the culture for 6–7 divisions at 30 °C with constant agitation in a roller drum (*see* Note 3).
5. Check 100 µL aliquots of your cultures (diluted in 900 µL diH<sub>2</sub>O) to determine the final OD<sub>600</sub>. Multiply this reading by 10. Grow the culture to whatever OD<sub>600</sub> will give 6–7 divisions based on your starting OD<sub>600</sub>, usually ~OD 1.2–5.0 (±1.0).
6. Take a 10 µL aliquot from the culture and add to a microfuge tube containing 90 µL diH<sub>2</sub>O. Dilute tenfold three more times (to 10<sup>-4</sup>), and plate 100 µL on a YC -Leu -Ura plate. Do this for each culture, making sure to keep cultures separate.
7. Incubate for 3 days at 30 °C.
8. Do colony PCR on ~25 colonies from each YC -Leu -Ura plate (*see* Notes 4 and 5).

9. Keep track of which PCR reactions came from which plate/parent colony and which daughter colony (#1–25). Mark the lanes with expansions (E) and contractions (C) on the picture (*see Note 6*).
10. Repeat the assay until ~100 daughter colonies from four parent colonies have been tested. More assays can be added if greater statistical significance is desired.
11. Stability data is calculated as frequency of change/total number of reactions and are either represented as: number of contractions/total number of reactions = frequency of contractions or number of expansions/total number of reactions = frequency of expansions. Statistical significance is determined by a Fisher's exact test using the raw numbers from the sum of reactions performed (for example 2 out of 100 versus 12 out of 100 would be compared) (*see Notes 7 and 8*).

### 3.2 Colony PCR to Determine CAG Repeat Tract Length

1. Select individual colonies and clearly mark which PCR reactions will come from which colony.
2. Combine a bulk PCR reaction mix on ice and aliquot 12.5  $\mu\text{L}$  per PCR tube. The 1 $\times$  PCR reaction is:

Sterile diH <sub>2</sub> O 5.8 $\mu\text{L}$	
10 $\times$ buffer (w/o MgCl <sub>2</sub> ):	1.25 $\mu\text{L}$
20 mM MgSO <sub>4</sub> :	1.25 $\mu\text{L}$
6.25 pmol primer 1:	0.625 $\mu\text{L}$ (of 10 pmol/ $\mu\text{L}$ stock)
6.25 pmol primer 2:	0.625 $\mu\text{L}$ (of 10 pmol/ $\mu\text{L}$ stock)
10 mM dNTPs:	0.25 $\mu\text{L}$
Mix thoroughly, then add,	
Taq polymerase:	0.2 $\mu\text{L}$
5 $\times$ CG-Rich buffer:	2.5 $\mu\text{L}$
Total reaction mix is	12.5 $\mu\text{L}$

Using a pipette tip, add a small amount of yeast cells from the center of the colony into the PCR reaction mix. When mixing the cells with reaction mix, either place a gloved finger on the wider end of pipette tip or use a pipetteman and push the plunger completely to ensure all liquid entering the tip by surface tension is expelled (*see Notes 9–11*).

## 3. PCR cycle profile:

95 °C	4'		
95 °C	30''		35 × times
54 °C	1'		
68 °C	3'		
68 °C	7'		
4 °C	Hold		

- Once the PCR protocol is complete, add an appropriate amount of loading dye to the 12.5  $\mu$ L reaction and run out the entire PCR reaction on a 2% Metaphor gel or other sizing gel system. Run Metaphor gels in 1 × TAE at 70 V for 90 min and load markers on both ends of the gel; a third marker in a middle lane is helpful for gel systems that do not run evenly across the gel box (*see Notes 12 and 13*).
- Stain the gel with ethidium bromide for 30 min. Transfer gel tray to diH<sub>2</sub>O and photograph the gel using a UV transilluminator (*see Note 14*).
- In Adobe Photoshop or similar program draw a line between the marker bands on each side, put a curve in if necessary, and decide if an expansion or contraction occurred based on whether the band falls above or below the line (*see Notes 15 and 16*).

### 3.3 Determination of Chromosomal Fragility by a YAC End Loss Assay

- Plate for single colonies on YC -Leu -Ura, which will select against breakage events, from a master patch.
- If the potential fragile sequence is an unstable repeat able to be sized via PCR, do colony PCR on ~20 colonies to check for the appropriate repeat tract length.
- Inoculate colonies of correct tract length for the fluctuation assay. Colonies should be as fresh as possible.

#### 3.3.1 Methods of Inoculation

- 10 colony protocol (preferred):  
For this methodology there will be 11–13 unique colonies that are inoculated into 11–13 separate cultures.
  - Pipet 1 mL of YC -Leu into 11–13 5 mL test tubes.
  - Using a pipet tip, inoculate 1–3 tubes with a small number of cells and take the OD<sub>600</sub> (using the entire 1 mL culture; these are referred to as the “OD<sub>600</sub> cultures”). The starting OD<sub>600</sub> should be ~0.02–0.06. Once the OD<sub>600</sub> has been taken, return the 1 mL culture to the initial test tube (*see Note 17*).

- (c) Inoculate the remaining tubes with approximately the same amount of cells by eye, so that they will also have a starting OD<sub>600</sub> of ~0.02–0.06. These are referred to as the “experimental cultures” (*see* **Note 18**).

2. 1 colony protocol:

For this methodology there will be one colony that is divided into ten experimental cultures and 1–3 OD<sub>600</sub> cultures.

- (a) Suspend one entire colony with the correct tract length in 1.3 mL of YC -Leu media.
- (b) Take 100 μL of the resuspended culture and inoculate into 11–13 5 mL test tubes that contains 0.9 mL of YC -Leu media.
- (c) Take the 1 mL OD<sub>600</sub> culture out of the test tube and measure the OD<sub>600</sub>. The OD<sub>600</sub> should be between ~0.02 and 0.06. Return the 1 mL culture to the initial test tube (*see* **Notes 19** and **20**).
- Allow the cultures to grow with constant agitation in a spinning roller drum for 6–7 doublings at 30 °C (*see* **Note 21**).
  - In order to determine the number of doublings, take 100 μL of the OD<sub>600</sub> cultures and resuspend into 900 μL of YC -Leu media and measure the final OD. Multiply this reading by 10. If the ideal starting OD<sub>600</sub> was ~0.02, the appropriate final OD<sub>600</sub> range is 1.25–2.5 (*see* **Note 3**). Once the correct OD<sub>600</sub> has been obtained, discard the 1–3 OD<sub>600</sub> tubes.
  - For the remaining ten cultures, spread 100 μL from each of the ten tubes onto ten YC -Leu 5-FOA plates using a consistent number of sterile glass beads or a sterile glass spreader. The amount plated can be adjusted up or down for strains with a higher or lower rate of FOA<sup>R</sup>; just remember to adjust the number of colonies counted to the number of mutants expected per 100 μL in the rate calculations.
  - To determine the viable total cell count, combine 100 μL from each of the ten undiluted cell cultures and mix together. Serially dilute the mixed cultures in sterile water (usually to 10<sup>-4</sup> and 10<sup>-5</sup>) and plate 100 μL on of each dilution onto YC -Leu plates. Total cell count plates should be plated in duplicate.
  - Incubate the plates at 30 °C for 4 days.
  - Count colonies on the YC -Leu 5-FOA plates.
  - Order the plate counts from low to high to determine the median number of colonies for each set (i.e., each



ten plate assay), and average the number of colonies on the two “total cell count” plates to determine the number of viable cells in 100  $\mu\text{L}$ . Determine the rate of 5-FOA<sup>R</sup> by using a fluctuation rate calculator such as FALCOR (<http://www.keshavsingh.org/protocols/FALCOR.html>) [19] or FluCalc (*see* this issue; <http://flucalc.ase.tufts.edu/>) using either the Maximum Likelihood Method (preferred, if the spread is not too great) or the Method of the Median. *See* ref. 20 for an excellent discussion of calculating fluctuation rates (*see* Notes 21–24).

- The assay should be repeated at least three times per strain, and standard error and confidence intervals calculated.

**3.4 Direct Duplication Recombination Assay: A Recombination Based Method to Assess Chromosome Breakage**

1. Patch strain to test onto a YC -Ura plate and allow to grow for 2 days at 30 °C.
2. Plate for single colonies on YEPD at 30 °C for 3 days such that there are  $50 \pm 10$  colonies per plate (this step allows recombination events to occur).
3. Randomly select ten colonies and resuspend half of each colony in 400  $\mu\text{L}$  diH<sub>2</sub>O (*see* Note 25).
4. Plate a portion of the suspension on a 5-FOA plate with the goal being to get ~50–200 colonies as the median after 4–5 days of growth at 30 °C. For the *LYS2* assay (Fig. 3a), this should be a YC -Lys 5-FOA plate; For the *ADE2* assay, (Fig. 3b), this should be a YC -Ade 5-FOA plate (*see* Note 26).
5. To determine the viable cell count combine 100  $\mu\text{L}$  of each of the ten colony suspensions. Plate a dilution that will yield ~30–300 colonies single colonies per YC plate. Total cell count plates should be plated in duplicate (*see* Note 27).
6. Count the number of colonies on the 5-FOA plates. Determine the number of mutants and the total number of cells per a certain volume, typically per 10  $\mu\text{L}$ , by multiplying the number of colonies growing on 5-FOA and the average of the two total cell count plates by the dilution factors used. To determine the rate of recombination, use the method of maximum likelihood or the method of the median and a fluctuation analysis calculator (e.g., FALCOR [19] or FluCalc; *see* Subheading 3.3).

---

## 4 Notes

### 4.1 Determination of Repeat Instability Frequency

1. This protocol is written for a CAG tract on the CFI YAC, but can be adapted to other locations or situations.

2. The low starting OD<sub>600</sub> ensures that cells have only been grown in log phase before being plated. The method is designed to test for instability occurring during log phase growth. The OD<sub>600</sub> range could be altered according to the purpose of the experiment.
3. For strains with poor division potential, the total number of divisions can be altered to fewer divisions, e.g., 3–5 divisions. Wildtype strains done as a comparison should be grown for the same number of divisions as mutant strains.
4. Pick colonies in as unbiased approach as possible: either pick a section of the plate or pick representative colonies that reflect the natural variability of colony size on the plate.
5. The PCR reactions do not have to be done all at once—choose the number by how proficient you are at colony PCR, but do all within 2 weeks. Keep track of which PCR reactions came from which plate.
6. For lanes where both a starting size band and new size band are visible, an expansion can be counted if it is clearly visible (intensity about 30% or more compared to starting size). A contraction can be counted if the intensity of the band is at least equal to the starting size band. The logic is that these events are likely have been present in the plated cell or occurred in the first cell division (expanded repeats PCR less efficiently than shorter repeats, thus the difference in cut-off criteria; (*see* ref. 13 for examples). The goal is to count events that occurred during the culture growth, and not those that occurred during the growth of the plated cell into a colony.
7. Typically, 100–200 PCR reactions in total are used. If a set of reactions from the same parent show a clear difference from the others (a jackpot), this assay should be dropped. For example, an undetected contraction in the starting parent colony could lead to an unusually high frequency of “contractions” in the daughters, which are not independent events that occurred during the culture growth. These often show up as many contractions of the same size.
8. The instability analysis can be combined with a fragility assay by analyzing the repeat tract length of the daughter colonies plated for the total cell count. Even though these plates do not select for the presence of the repeat (e.g., the YC -Leu plates for the YAC assay), usually the rate of breakage is low enough that most colonies tested will have retained a repeat tract. One disadvantage to this method is that the starting parent colony and the daughter colonies are no longer correlated, reducing the ability to detect jackpot events.

**4.2 Colony PCR to Determine CAG Repeat Tract Length**

9. In order to obtain robust tract lengths PCR, primers should be designed such that there is a region of nonrepetitive flanking sequence included on both sides of the repeat. Generally, we recommend between 40 and 80 bp of nonrepetitive sequence be included on each side of the repeat.
10. Taq polymerase and CG buffer should be added last to the master mix after everything else. CG rich buffer should not be vortexed.
11. The amount of cells added to the PCR reaction should be no larger than the tip of the p200 or p20 pipette tip. Too few cells will result in a very faint PCR product, while too many will result in a background smear in the lane.
12. The 2% Metaphor agarose gels should be made according to the manufacturer's instructions.
13. Instead of using 2% Metaphor gels for separation of fragments, PCR reactions can be run on a denaturing PAGE gel or other high resolution gel system for size analysis, with appropriate DNA detection.
14. Ethidium bromide can be added directly to Metaphor gel if preferred. The ideal concentration is 15  $\mu$ L 10 mg/mL stock in 300 mL 1 $\times$  TAE.
15. The ideal molecular weight marker used should have bands that change in size every 100 bp or less. Adding an additional marker in the middle of the gel helps with sizing the CAG repeat.
16. If the PCR is of poor quality, then expansions may not have amplified well and could have been underestimated; therefore, only use good-quality PCR reactions to determine expansion frequencies. A 2% Metaphor gel can resolve fragments that differ by 2%, i.e., 8 bp for a fragment size of 400 bp or about 3 (CAG) repeats added or subtracted for a (CAG)<sub>85</sub> repeat tract.

**4.3 Determination of Chromosomal Fragility Using a YAC End Loss Assay**

17. Usually the "OD<sub>600</sub> cultures" are only used for taking the OD<sub>600</sub> throughout the experiment and not plated to be counted, due to risk of contamination during the process of taking the OD<sub>600</sub>.
18. The method is designed to test for fragility occurring during log phase growth. The OD<sub>600</sub> range could be altered according to the purpose of the experiment.
19. The 1 colony method is less preferred, but can be used when finding 10 starting colonies of the same repeat tract length is technically difficult.
20. For wild type, 6–7 divisions will take approximately ~16–18 h.

21. Make sure you enter the number of mutants and total cells for the *same volume*—usually per 100  $\mu\text{L}$  (per 1 mL can be used for very low rates).
22. For experiments with zeros or a large spread between low and high numbers of mutants (happens with lower rate assays), use the Method of the Median to calculate the rate.
23. Certain mutants could alter *URA3* gene function or FOA<sup>R</sup> by increasing the mutation rate of *URA3*, altering *URA3* expression, or affecting the toxicity of 5-FOA, resulting in increased FOA<sup>R</sup> that is not due to chromosome breakage. To confirm that there is loss of the right arm of the YAC, PCR amplify the *URA3* gene from 1 colony off of each YC -Leu 5-FOA plate (ten colonies total) to confirm there is loss of the right arm of the YAC. This should be done for three assays (a total of 30 independent colonies checked). The expected end loss frequency for wild type is 90–100%. Make sure to also run your starting strain as a positive control.
24. Verify the structure of the YAC in 20 independent FOA<sup>R</sup> colonies (plated from independent cultures) by Southern blot, using a lambda probe to the YAC and digestion with BstEII (*see ref. 6*).
25. Establish a standard size of colony to use to allow equal generations of growth between colonies and strains.
26. Generally, for rates in the  $10^{-5}$  range, approximately 10  $\mu\text{L}$  of the colony suspension is plated onto 5-FOA. A pilot experiment can determine the best amount to plate. Keep in mind that the more fragile the sequence, the less should be plated.
27. The dilution plated for the viable cell count depends on the size of the colonies that were picked for the assay. For initial assays it is best to plate a series of dilutions, such as  $10^{-4}$  and  $10^{-5}$ .

**4.4 Direct  
Duplication  
Recombination Assay**

---

## Acknowledgments

Work in our laboratory is currently supported by grants from the National Institute of Health (GM105473) and National Science Foundation (MCB 1330743). We would like to thank past and present members of the C. H. Freudenreich and V.A. Zakian labs for their contributions to the development of these assays.

## References

1. Usdin K, House NC, Freudenreich CH (2015) Repeat instability during DNA repair: Insights from model systems. *Crit Rev Biochem Mol Biol* 50(2):142–167. doi:[10.3109/10409238.2014.999192](https://doi.org/10.3109/10409238.2014.999192). PubMed PMID: 25608779; PMCID: 4454471
2. McMurray CT (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* 11(11):786–799. doi:[10.1038/nrg2828](https://doi.org/10.1038/nrg2828). PubMed PMID: 20953213; PMCID: 3175376

3. Shishkin AA, Voineagu I, Matera R, Cherng N, Chernet BT, Krasilnikova MM, Narayanan V, Lobachev KS, Mirkin SM (2009) Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol Cell* 35(1):82–92. doi:[10.1016/j.molcel.2009.06.017](https://doi.org/10.1016/j.molcel.2009.06.017). PubMed PMID: 19595718; PMCID: 2722067
4. Dixon MJ, Bhattacharyya S, Lahue RS (2004) Genetic assays for triplet repeat instability in yeast. *Methods Mol Biol* 277:29–45. doi:[10.1385/1-59259-804-8.029](https://doi.org/10.1385/1-59259-804-8.029). PubMed
5. Schulz VP, Zakian VA (1994) The saccharomyces PIF1 DNA helicase inhibits telomere elongation and de novo telomere formation. *Cell* 76(1):145–155. PubMed
6. Callahan JL, Andrews KJ, Zakian VA, Freudenreich CH (2003) Mutations in yeast replication proteins that increase CAG/CTG expansions also increase repeat fragility. *Mol Cell Biol* 23(21):7849–7860. Epub 2003/10/16. PubMed PMID: 14560028; PMCID: 207578
7. Kerrest A, Anand RP, Sundararajan R, Bermejo R, Liberi G, Dujon B, Freudenreich CH, Richard GF (2009) SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. *Nat Struct Mol Biol* 16(2):159–167. doi:[10.1038/nsmb.1544](https://doi.org/10.1038/nsmb.1544). PubMed Epub 2009/01/13. doi: nsmb.1544 [pii]
8. Zhang H, Freudenreich CH. An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae*. *Mol Cell*. 2007;27(3):367–379. doi: [10.1016/j.molcel.2007.06.012](https://doi.org/10.1016/j.molcel.2007.06.012). PubMed PMID: 17679088; PMCID: 2144737
9. Cherng N, Shishkin AA, Schlager LI, Tuck RH, Sloan L, Matera R, Sarkar PS, Ashizawa T, Freudenreich CH, Mirkin SM (2011) Expansions, contractions, and fragility of the spinocerebellar ataxia type 10 pentanucleotide repeat in yeast. *Proc Natl Acad Sci U S A* 108(7):2843–2848. doi:[10.1073/pnas.1009409108](https://doi.org/10.1073/pnas.1009409108). PubMed PMID: 21282659; PMCID: 3041125
10. Lu S, Wang G, Bacolla A, Zhao J, Spitzer S, Vasquez KM (2015) Short inverted repeats are hotspots for genetic instability: relevance to cancer genomes. *Cell Rep*. doi:[10.1016/j.celrep.2015.02.039](https://doi.org/10.1016/j.celrep.2015.02.039). PubMed
11. Pluta AF, Zakian VA (1989) Recombination occurs during telomere formation in yeast. *Nature* 337(6206):429–433. doi:[10.1038/337429a0](https://doi.org/10.1038/337429a0). PubMed
12. Chen C, Kolodner RD (1999) Gross chromosomal rearrangements in *Saccharomyces cerevisiae* replication and recombination defective mutants. *Nat Genet* 23(1):81–85. doi:[10.1038/12687](https://doi.org/10.1038/12687). PubMed
13. Freudenreich CH, Stavenhagen JB, Zakian VA (1997) Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Mol Cell Biol* 17(4):2090–2098. PubMed PMID: 9121457; PMCID: 232056
14. Su XA, Dion V, Gasser SM, Freudenreich CH (2015) Regulation of recombination at yeast nuclear pores controls repair and triplet repeat stability. *Genes Dev* 29(10):1006–1017. doi:[10.1101/gad.256404.114](https://doi.org/10.1101/gad.256404.114). PubMed PMID: 25940904; PMCID: 4441049
15. Freudenreich CH, Kantrow SM, Zakian VA (1998) Expansion and length-dependent fragility of CTG repeats in yeast. *Science* 279(5352):853–856. PubMed
16. Balakumaran BS, Freudenreich CH, Zakian VA (2000) CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in *Saccharomyces cerevisiae*. *Hum Mol Genet* 9(1):93–100. PubMed
17. Paeschke K, Capra JA, Zakian VA (2011) DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* 145(5):678–691. doi:[10.1016/j.cell.2011.04.015](https://doi.org/10.1016/j.cell.2011.04.015). PubMed PMID: 21620135; PMCID: 3129610
18. Dunham MJ, Dunham MJ, Gartenberg MR, Brown GW (2015) Methods in yeast genetics and genomics : a Cold Spring Harbor Laboratory course manual/Maitreya J. Dunham, University of Washington, Marc R. Gartenberg, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, Grant W. Brown, University of Toronto. 2015 edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. xvii, 233 pages
19. Hall BM, Ma CX, Liang P, Singh KK (2009) Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 25(12):1564–1565. doi:[10.1093/bioinformatics/btp253](https://doi.org/10.1093/bioinformatics/btp253). PubMed PMID: 19369502; PMCID: 2687991
20. Rosche WA, Foster PL (2000) Determining mutation rates in bacterial populations. *Methods* 20(1):4–17. doi:[10.1006/meth.1999.0901](https://doi.org/10.1006/meth.1999.0901). PubMed PMID: 10610800; PMCID: 2932672

## Quantitative Analysis of the Rates for Repeat-Mediated Genome Instability in a Yeast Experimental System

Elina A. Radchenko, Ryan J. McGinty, Anna Y. Aksenova,  
Alexander J. Neil, and Sergei M. Mirkin

### Abstract

Instability of repetitive DNA sequences causes numerous hereditary disorders in humans, the majority of which are associated with trinucleotide repeat expansions. Here, we describe a unique system to study instability of triplet repeats in a yeast experimental setting. Using fluctuation assay and the novel program FluCalc we are able to accurately estimate the rates of large-scale expansions, as well as repeat-mediated mutagenesis and gross chromosomal rearrangements for different repeat sequences.

**Key words** Trinucleotide repeats, Repeat expansions, Repeat-induced mutagenesis, Fluctuation assay, Expansion rate, Mutation rate, MSS-MLE, FluCalc

---

### 1 Introduction

The fluctuation test has been a cornerstone of the field of microbial genetics since the Luria–Delbrück experiment, which determined that mutations arise spontaneously during growth of a culture prior to selection [1]. The mathematical methods underlying mutation rate calculations are complicated by the fact that mutations spontaneously arise by chance at different generations in the culture, and subsequently accumulate with each cell division [1, 2]. This produces fluctuations in the frequency of mutants in each independent experiment. The mathematical methods of accounting for this natural fluctuation have improved over time, with innovations to account for variables such as the relative fitness of the mutant population, partial plating of the culture, or variation in the total population size across the cultures in the experiment [3–5].

Laboratory methods have also evolved since the days of Luria and Delbrück. An experimental strain typically contains a forward-selection reporter, which allows the detection of mutant cells on

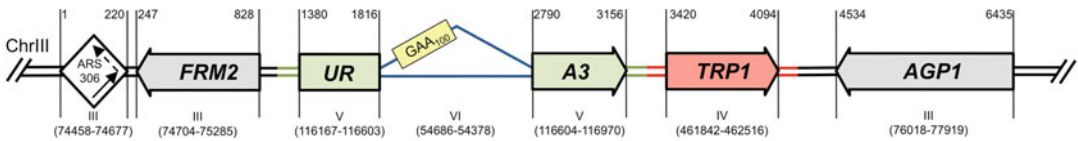
particular selective media. A fluctuation test usually begins with growth of several parallel cultures in nonselective conditions, where a small number of inoculated cells grow to saturation. In our experiments, we first spread cells on a Petri dish, where a single cell grows to form a clonal colony. Individual colonies are then resuspended in water and various dilutions are plated onto both selective and nonselective solid media. The number of colonies grown on each plate is used to determine the number of mutations per cell per generation, also known as  $\mu$  or the mutation rate.

Notably, however, different colonies on selective media do not necessarily arise from the same type of mutation. A reporter gene may be inactivated by a point mutation, a small insertion or deletion causing a frameshift, a larger deletion removing part or all of the gene, or by chromosomal events such as translocations or the loss of a nonessential arm of the chromosome. Furthermore, mutations can occur within the body of the gene, or in an important regulatory region. Finally, reporter gene inactivation can also occur indirectly, for instance via mutations or downregulation of genes responsible for the uptake of a selective drug. These different events do not necessarily have the same consequences in terms of colony formation, with a regulatory change producing a weaker selective pressure than a deletion of the entire marker. Variations on the fluctuation assay may focus on one or more of these types of mutations. A classic example is the frameshift reversion assay, which begins with a selective marker inactivated by a 1 bp insertion, and selects for those mutations that restore the frame of translation.

One unusual type of mutation is microsatellite instability. Tracts of short repetitive sequences (repetitions of 2–12 bases) can mutate in their length, with the number of repeats either expanding or contracting. Expansions of microsatellite sequences have been linked to a variety of human disorders [6]. The instability of microsatellite sequences can be attributed to the formation of various non-B-form DNA secondary structures, depending on the particular sequence. Repeats can be added a few at a time, or by a single large-scale event. Different microsatellite sequences may be more or less prone to large or small-scale events, as well as chromosomal fragility and repeat-induced mutations, all of which differ in their underlying mechanisms and rates [7]. Importantly, the rates of these various events appear to increase exponentially with the length of the starting repeat tract [8, 9]. Depending on the particular repeat, the rate of such events can be higher than the average point mutation rate of the genome by many orders of magnitude.

In the protocol below, we describe the use of our experimental system to select for large-scale expansions of H-DNA-forming (GAA)<sub>n</sub> repeats in *S. cerevisiae* [8, 10]. The yeast cells carry a selectable cassette consisting of a modified *URA3* gene, which is counter-selectable on media containing the drug 5-fluoroorotic





**Fig. 1** Experimental system to study GAA repeats expansions in yeast. The selectable cassette contains flanking sequences from chromosome III (depicted in *black* and *grey*), *URA3* flanking and coding sequences from chromosome V (*green*), intron sequence inside *URA3* gene derived from *ACT1* gene from chromosome VI (*blue*) and *TRP1* flanking and coding sequences from chromosome IV (*red*)

acid (5-FOA) [see Fig. 1]. This cassette is located ~1 kb downstream of the early-firing replication origin ARS306. The *URA3* gene has been modified to include 100 GAA repeats within a 974 bp artificial intron. Due to the limitations of the yeast spliceosome, an intron exceeding ~1000 bp is no longer spliced, leading to 5-FOA resistance in strains bearing our genetic construct. It is important to note that this threshold is not absolute, but rather splicing efficiency appears to diminish in an exponential manner as the intron length grows past 1000 bp. Thus, we designed our system to begin with the overall intron length close to this threshold, such that even relatively short expansions should be sufficient to induce a level of 5-FOA resistance distinguishable from the background.

Importantly, however, *URA3* can become inactivated not only due to a repeat expansion, but also by point mutations, deletions, translocations, downregulation, or even indirectly. In practice, each of these situations creates varying degrees of 5-FOA resistance, resulting in a range of colony sizes. The smallest 5-FOA-resistant colonies are typically formed owing to indirect resistance or slight *URA3* downregulation. The largest colonies typically result from the strongest inactivating events, such as nonsense mutations and large deletions. Expansions typically occupy a middle ground, as they severely but not totally reduce *URA3* expression, with the strength of this effect depending on the size of the expansion. Very large expansions may be the result of secondary expansions, *i.e.*, expansions that occurred following a primary expansion. To distinguish between these different events, we perform a PCR on various (ideally all) 5-FOA-resistant colonies. Using primers flanking the repetitive tract, we can measure the size of each expansion, or infer the presence of a deletion/translocation (if the PCR product is absent) or a point mutation (if there is no change in repeat length).

A major advantage of our system is that it can be adapted to favor the recovery of certain types of events, simply by altering the concentration of 5-FOA in the selective media. The weakest events, including small-scale expansions, can be observed by growing on a low concentration of 5-FOA. The concentration of 5-FOA can be optimized to reduce these background colonies, favoring the recovery of large-scale primary expansions. If the concentration of



5-FOA is increased further, this can eliminate most expansions, highlighting only the most damaging types of mutations, including deletions, translocations and point mutations.

Indeed, there are a number of possible modifications to the fluctuation test, and to our expansion system, that can be made. For instance, the starting repeat length can be increased or decreased [8], or the  $(GAA)_n$  repeats can be swapped out for other sequences such as SCA10 pentanucleotide  $(ATTCT)_n$  repeats, or telomeric repeats [11, 12]. The *URA3* promoter can be swapped for an inducible *GALI* promoter to study the effects of transcription [10]. The entire cassette can be moved to different locations, such as the nonessential arm of chromosome V, allowing for recovery of arm-loss events (double-strand breaks that were not repaired) [13]. In addition to forward selection for repeat expansions, reverse-selection to a strong  $Ura^+$  phenotype can be used to study contractions of the repetitive tracts [14]. Thus, our system is highly adaptable to the study of various molecular mechanisms.

As written, the protocol below describes the steps used for measuring the rate of primary expansions in our wild-type strain with the  $(GAA)_{100}$  selectable cassette located ~1 kb downstream of the early-firing origin ARS306. This is perhaps the most challenging use of the assay, as it requires optimization of 5-FOA concentration and growth times, as well as careful scrutiny of the distribution of repeat lengths. We have included a number of notes in the protocol to guide with this optimization, as well as suggestions for where the assay can be modified. Variations of this protocol can be found in our previous publications [8–12, 14], but this document represents our most up-to-date and refined version of the assay.

Finally, we introduce here a new tool, named FluCalc, for calculating mutation rates from these fluctuation tests (*see* Fig. 2). Although several mathematical tools such as SALVADOR, or web-tools such as FALCOR and bz-rates were designed for accurate estimation of mutation rate [15–18], they appeared either unsuitable or inconvenient for our experiments. FluCalc uses the Ma-Sandri-Sarkar maximum likelihood estimation (MSS-MLE) equations for calculation, regarded as one of the most accurate estimators available [19]. MSS-MLE is also applicable across most mutational spectra [20]. FluCalc calculates the values for mutant frequencies, mutations per culture, mutation rates and 95% confidence intervals of mutation rates. FluCalc also automatically incorporates plating efficiency calculations to account for dilutions. FluCalc displays a full description of all equations used in the calculations, and shows all values of the intermediate variables at each step. FluCalc is presented in a web format suitable for all modern browsers, and does not rely on plugins that become outdated and broken. FluCalc allows one to download and save the data as a text file, which can be subsequently reopened by the

# FluCalc

Version: 2016.7.2

Enter data below or select a file with data:  No file chosen

---

Volume of a culture ( $\mu\text{l}$ ),  $V_{\text{tot}}$

### Selective media

Observed numbers of clones,  $C_{\text{sel}}$

27

53

47

73

21

40

22

30

45

19

32

38

Dilution factor,  $D_{\text{sel}}$

Volume plated ( $\mu\text{l}$ ),  $V_{\text{sel}}$

### Complete media

Observed numbers of clones,  $C_{\text{com}}$

34

60

55

43

34

59

42

32

44

56

31

58

Dilution factor,  $D_{\text{com}}$

Volume plated ( $\mu\text{l}$ ),  $V_{\text{com}}$

## Results

### Raw values

m	11.0726
$\mu$	12.1233e-06
$CI_{\text{lower}}$	8.4580e-06
$CI_{\text{upper}}$	16.2564e-06
Frequency mean	8.1569e-05

### Corrected values

m	15.9743
$\mu$	1.7490e-05
$CI_{\text{lower}}$	1.2736e-05
$CI_{\text{upper}}$	2.2781e-05

**Fig. 2** Screenshot of the FluCalc interface. A user needs to enter the number of colonies on both selective ( $C_{\text{sel}}$ ) and complete media ( $C_{\text{com}}$ ), a dilution factor ( $D_{\text{sel}}$  and  $D_{\text{com}}$ ), and a volume of cultures ( $V_{\text{sel}}$  and  $V_{\text{com}}$ ) plated on both media. A volume of the initial culture ( $V_{\text{tot}}$ ) can be changed from its default value of 200  $\mu\text{L}$ . Raw values and corrected values for  $m$ —number of mutations;  $\mu$ —mutation rate;  $CI$ —confidence intervals for  $\mu$

program or imported into a spreadsheet program. This feature is especially handy when processing and comparing large amounts of data. We hope that FluCalc will be useful for many researchers wanting an accurate and efficient tool for mutation rate calculations. The tool is available at <http://ase.tufts.edu/biology/labs/mirkin/resources/>.

---

## 2 Materials

All experiments are described based on the *Saccharomyces cerevisiae* WT strain CH1585 (*MATa*, *leu2Δ1*, *trp1Δ63*, *ura3Δ52*, and *his3Δ200*) with (GAA)<sub>100</sub> repeats inserted in the intron-containing *URA3* gene on chromosome III (see **Note 1**).

### 2.1 Nonselective Growth Period

1. 15% glycerol stock solution: 300 μL of autoclaved 50% glycerol solution +700 μL of YPD.
2. YPD + Uracil plates (example makes 1 L): 1 L dH<sub>2</sub>O, 10 g yeast extract, 20 g peptone, 20 g dextrose, 20 g agar. Autoclave, allow to cool slightly, then add 24 mL uracil from a 2 mg/mL stock solution. Pour into Petri dishes, solidify overnight, seal and store at 4°C.

### 2.2 Selective Growth Period

1. 5-FOA media (example makes 1 L): Autoclaved portion: 500 mL dH<sub>2</sub>O, 20 g agar, 20 g dextrose, 2 g synthetic complete amino acid supplement minus uracil with added adenine, 6.7 g yeast nitrogen base without amino acids or carbohydrates and with ammonium sulfate. Nonautoclaved portion: 500 mL dH<sub>2</sub>O, 1 g 5-FOA, 50 mg uracil—shake at 37°C until 5-FOA is completely dissolved and use a 0.2 μm bottle-top vacuum filter to sterilize the solution. Allow autoclaved portion to cool slightly, then combine both portions, mix on stirrer and immediately pour into Petri dishes. Let the media solidify overnight and use the following day (see **Note 2**).
2. YPD media (example makes 1 L): 1 L dH<sub>2</sub>O, 10 g yeast extract, 20 g peptone, 20 g dextrose, and 20 g agar. Autoclave, allow to cool slightly, and pour into Petri dishes. Let the media solidify overnight, seal and store at 4°C.

### 2.3 Colony PCR

1. Lyticase solution: 2.5 mg/mL lyticase (L2524 Sigma ≥2000 units/mg protein) in 0.9 M sorbitol, 0.1 M EDTA (pH 7.4) (see **Note 3**). Zymolyase (e.g., 100 T Zymo Research) performs equally well.
2. PCR primers (see **Note 4**).
3. PCR mastermix (example makes 24 μL):

H <sub>2</sub> O	13.5 $\mu$ L
5 $\times$ Green GoTaq buffer (Promega)	5 $\mu$ L
10 mM dNTPs	0.25 $\mu$ L
10 $\mu$ M Forward Primer	2.5 $\mu$ L
10 $\mu$ M Reverse primer	2.5 $\mu$ L
Sibenzyme Taq	0.25 $\mu$ L

### 2.4 Gel Electrophoresis

1. 10 $\times$ TBE stock solution: 0.89 M Tris, 0.89 M borate, 0.02 M EDTA. Store at room temperature.
2. Ultrapure high-resolution agarose.
3. NEB Quick-Load 100 bp or 50 bp DNA ladder.
4. 1% ethidium bromide solution.

---

## 3 Methods

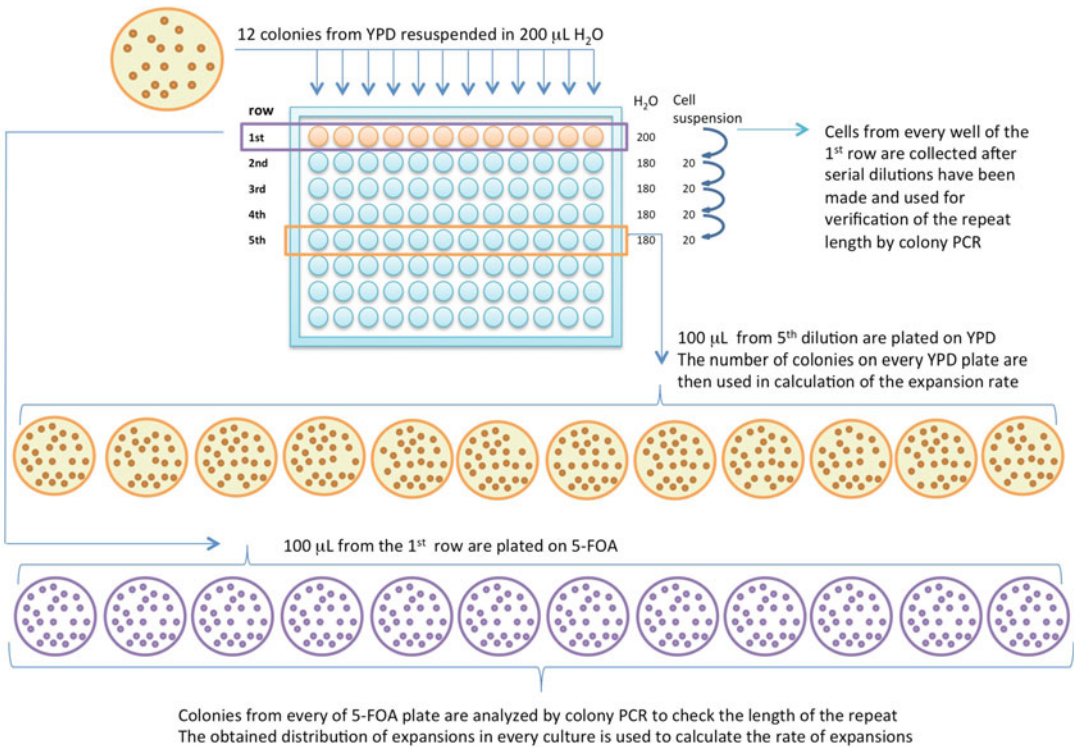
### 3.1 Nonselective Growth Stage (See Note 5)

1. Streak a small amount of yeast from 15% glycerol frozen stock onto a YPD plate supplemented with uracil such that distinct individual colonies form (*see Note 6*).
2. Incubate for 40 hours at 30°C (*see Note 7*).

### 3.2 Selective Growth and Total Cell Count (See Fig. 3)

1. Fill a sterile 96-well polypropylene, U-well microtiter plate with 200  $\mu$ L of sterile deionized water in the 1st row and 180  $\mu$ L in rows 2–5.
2. Use a small plastic inoculating loop to scoop up an average-sized colony and resuspend it in the 1st row of the plate. Repeat until 8–12 similarly sized colonies have been resuspended, one per well (*see Note 8*). Mix thoroughly with a multichannel pipette.
3. Perform 10 $\times$  serial dilutions, transferring 20  $\mu$ L from the 1st to 2nd rows, and then into subsequent rows, mixing thoroughly at each step (*see Note 9*).
4. Mark corresponding YPD and 5-FOA plates with the strain name and colony number.
5. Plate 100  $\mu$ L from the first row onto a 5-FOA plate and 100  $\mu$ L from the 5th row onto a YPD plate for each sample (*see Note 10*) and spread until dry with a large disposable plastic inoculating loop.
6. Spin down microtiter plate to collect yeast cells for PCR analysis, discard supernatant, add 1.5  $\mu$ L lyticase solution, mix thoroughly and go to Subheading 3.3.3 (*see Note 11*).

**Scheme of the fluctuation test**



**Fig. 3** Scheme of the fluctuation test

7. Incubate 5-FOA and YPD plates at 30°C for 3 days (*see Note 12*).
8. Count individual colonies on all plates (*see Note 13*).
9. (Optional) Using a flatbed scanner, scan images of all plates. Load images into OpenCFU colony counting software. Adjust threshold and radius values until all 5-FOA-resistant colonies are included and all background growth is excluded from the count (*see Note 14*).

**3.3 PCR Analysis of Repeats (See Note 15)**

**3.3.1 Option A (See Note 16)**

1. If necessary, randomly subdivide 5-FOA plates into 1/2, 1/4 or 1/8 size wedges containing no less than 4 and no more than 16 colonies within a wedge from each plate. Count total number of colonies within random wedge of each 5-FOA plate.
2. Pick all colonies from within the selected 5-FOA plate or wedge using a pipette tip and resuspend one colony in each well of a 96-well PCR plate containing 1.5 µL lyticase solution per well.

**3.3.2 Option B (See Note 17) and Option C (See Note 18)**

1. For each 5-FOA plate, choose four 5-FOA-resistant colonies for PCR analysis.
2. Pick colonies using a pipette tip and resuspend one colony in each well of a 96-well PCR plate containing 1.5  $\mu$ L lyticase solution per well.

**3.3.3 Continue from Option A, B or C**

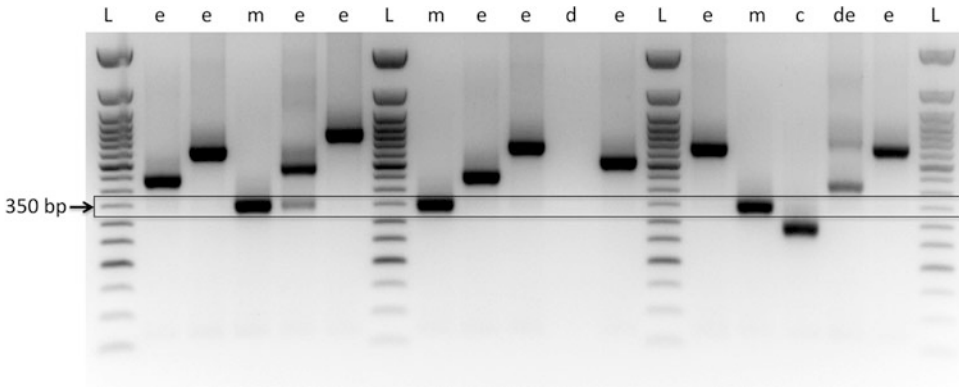
1. Incubate cells in lyticase solution for 8–15 min at 37°C
2. Add 50  $\mu$ L of PCR-grade water to all wells. Resuspend if necessary.
3. Incubate plate at 100°C for 5 min.
4. Spin plate at 2500–3000  $\times g$  for 2–5 minutes to pellet cell debris. The solution containing yeast genomic DNA can be stored for up to 12 months at –20°C and repeatedly analyzed.
5. Prepare PCR master mix for the appropriate number of samples.
6. Add 12–24  $\mu$ L of PCR master mix into each well of a new 96-well PCR plate.
7. Add 1  $\mu$ L of genomic DNA supernatant to each PCR reaction.
8. Run PCR program:
  - (a) 95°C—15 s.
  - (b) 94°C—15 s.
  - (c) 72°C—2 min.
  - (d) Go to Step 2  $\times$  30.
  - (e) 72°C—2 min.
  - (f) 12°C—hold.

**3.4 Gel Electrophoresis (See Fig. 4)**

1. Prepare a 1.5% agarose gel using ultrapure agarose and 0.5 $\times$  TBE (*see* Note 19). Add 2  $\mu$ L of ethidium bromide per 100 mL of TBE.
2. Load samples along with 100 bp DNA ladder, run gel, and image.
3. Calculate expansion sizes by subtracting the starting length of the unexpanded PCR product from expanded PCR product sizes and dividing by 3 (for triplet repeats) (*see* Note 20).
4. For each individual experimental strain, visualize the distribution of expansion sizes by graphing the percentage of expanded products that fall into 10-repeat bins (0–10 repeats added, 11–20 repeats added, etc.) (*see* Note 21).

**3.5 Expansion Rate Calculation (See Note 22)**

1. Insert the following information into FluCalc to obtain an accurate expansion rate with corresponding 95% confidence interval:



**Fig. 4** Representative gel electrophoresis of the PCR analysis of 5-FOA-resistant colonies, originated in a strain containing our selectable cassette with the  $(GAA)_{100}$  repeat. *L*—ladder, *e*—expansion, *d*—deletion, *m*—mutation, *c*—contraction with a simultaneous mutation in the *URA3* body, *de*—double expansion. The initial length of PCR product containing  $(GAA)_{100}$  repeat is ~350 bp that is marked by *arrow*

- (a)  $V_{\text{tot}}$ —the total volume of the culture following resuspension of colonies in **step 2** of Subheading 3.2 (typically 200  $\mu\text{L}$ ).
  - (b)  $C_{\text{sel}}$ —the observed number of colonies with expanded repeats from each plate or sector (adjusted to reflect the results of the gel electrophoresis) (*see Note 23*).
  - (c)  $C_{\text{com}}$ —the observed number of colonies on nonselective media.
  - (d)  $D_{\text{sel}}$ —the dilution factor used for plating on selective media (e.g., if liquid from the second row of a  $10\times$  serial dilution was plated on 5-FOA, then  $D_{\text{sel}} = 10$ ).
  - (e)  $D_{\text{com}}$ —the dilution factor used for plating on nonselective media (e.g., if liquid from the fifth row of a  $10\times$  serial dilution was plated on YPD then  $D_{\text{com}} = 10,000$ ).
  - (f)  $V_{\text{sel}}$ —the volume of culture plated per plate or sector on selective media (*see Note 24*).
  - (g)  $V_{\text{com}}$ —the volume of culture plated per plate on nonselective media.
2. Download and save the data in FluCalc.txt format so that it can be subsequently reopened by the program.
  3. Click on the “Calculate” button to get results for the expansion rate, presented as either raw values or corrected values (*see Note 25*). Frequencies for each plate will also be reported (*see Note 26*). 95% confidence intervals are also reported (*see Note 27*).



---

## 4 Notes

1. The strain used for the example in this protocol was described and thoroughly examined in Shah et al. 2012 [9]. A key feature of this strain is the “balanced” intron length just under the splicing threshold length, allowing a single expansion to inactivate the *URA3* gene. The rate of expansions for a  $GAA_{100}$  repeat was determined to be  $\sim 2 \times 10^{-5}$  expansions per cell per generation. Altering the starting repeat length or sequence can dramatically change this rate. Our protocol is ideal for measuring rates in the range of  $10^{-6}$  to  $10^{-3}$ . Measuring rates below this range is difficult due to the larger number of cells required for plating. Measuring rates above this range results in large natural fluctuations, making this type of assay inappropriate. We recommend adjusting the starting length of the repeat until the rate is close to  $10^{-5}$ . This will allow the investigation of various mutants or experimental conditions that raise or lower the rate within the measurable range.
2. 5-FOA will remain effective when stored at 4°C for a period of weeks. However, as water slowly evaporates from the plates, this will raise the effective concentration of 5-FOA, which can shift the selection toward larger expansions, point mutations, and deletions. Thus it is important, if determining the rate of expansions in a mutant strain, to concurrently retest the wild-type strain to help control for subtle differences in the media. Plastic wrap can be used to reduce evaporation, but it is best to use the plates within days. The above recipe is for 0.1% 5-FOA media, which is ideal for the selection of primary repeat expansions. To select for deletions/point mutations/translocations, a concentration of 0.15% works well. The strength of the 5-FOA media can also be increased by lowering the pH, though we find this method of adjustment to be less precise.
3. The pH of the lyticase solution should be kept around 7.4–7.5 and is critical for the activity of the enzyme. This solution should be stored in small aliquots at  $-20^{\circ}\text{C}$ .
4. Primers should be closely flanking the repeats, and should have  $T_m$  above  $72^{\circ}\text{C}$ . The  $T_m$  should be high to discourage secondary structure formation which can lead to PCR failures that create faint or smeared bands. The PCR product should be as small as possible to facilitate accurate length determination. PCR products with longer nonrepetitive flanks can sometimes help to reduce PCR failures, but these longer products will not be as accurate for measuring small changes in length.
5. If assessing any experimental variables (such as temperature sensitivity, chemical perturbation, or gene induction), it is important to introduce them during this stage of growth.



6. Extra uracil is added to the media to prevent cells from becoming dependent on the *URA3* gene, which will be inactivated by an expansion. If petites are particularly frequent following unfreezing, the strains can first be patched onto YPEG plate for 1 day. The YPEG patches can then be used to streak for single colonies on YPD + Uracil. For strains in which the *URA3* promoter is replaced by the galactose-inducible *GALI* promoter, start with a 5 mL overnight culture in YPD + Raffinose to remove glucose inhibition of the *GALI* promoter. Dilute the overnight culture in deionized water and plate on YPD or YPGal, supplemented with uracil.
7. This amount of time can be extended for any slow-growing strains or alternate growth conditions. However, it is important not to grow the colonies for too many generations, which can result in a higher probability of secondary expansions. This is because a primary expansion will increase the next generation's starting repeat length, which will in turn have an exponentially higher rate of expansion.
8. It is important to pick colonies that are of the same size. Most methods to calculate mutation rates (including MSS-MLE) rely on an assumption that culture sizes are the same (within 20% of the coefficient of variation) across the experiment [5, 20]. To decrease the inter-culture variation, do not choose the largest or smallest colonies on the plate—it is better to aim for average sized colonies. When testing our wild-type *GAA*<sub>100</sub> strain, we choose single colonies that are about 1 mm in diameter when uncrowded. The starting colony size can be adjusted in anticipation of higher or lower expansion rates. After colonies have been counted (**Step 8** in Subheading 3.2), if there is a YPD plate that deviates significantly from the median, both the YPD and corresponding 5-FOA plate should be excluded from further analysis.
9. Accurate pipetting is critical throughout the serial dilution, as small random errors can be magnified after each step. Accuracy can be improved through the use of low-retention tips, by discarding tips between each suspension, and by visually checking for air bubbles, droplets stuck in the tips, or other pipetting errors.
10. This recommendation is applicable to our wild-type *GAA*<sub>100</sub> strain, and should result in uncrowded single colonies. Both the volume and choice of dilution can be altered for different mutant backgrounds or growth conditions.
11. PCR analysis of the starting colonies should confirm that expansions did not take place prior to the nonselective growth period. All future analysis depends on these cells beginning from the same initial repeat length. Encountering expanded

or contracted clones at this stage should be rare, but strains with very high rates of instability may require careful monitoring.

12. The growth period may be altered if necessary. Alterations to the temperature or media components at this stage will not affect the rate at which mutations were generated during the nonselective growth phase, but may affect the recovery of those events for counting. Only change these conditions if you have reason to believe you are not recovering certain events. Also, if using a galactose-inducible strain, the 5-FOA media must also contain galactose in order for the *URA3* gene to be active and thus counter-selectable.
13. For YPD plates, all colonies must be counted. For 5-FOA plates, there will be a certain amount of small background colonies, which do not have expansions or other mutations of interest. A rule of thumb is that if the colony is too small to be easily picked up with a pipette tip, it cannot be analyzed by PCR and therefore should not be counted. It is possible to empirically determine for each strain the relationship between colony size and mutation status. In any case, it is important to be consistent between those colonies counted and those chosen for PCR analysis.
14. The OpenCFU program is freely available at <http://opencfu.sourceforge.net/>. An alternative is the Image J program with a cell-counter plugin (<https://imagej.nih.gov/ij/index.html>). Use a consistent image size for every plate (a 300 dpi scan will produce 1000 × 1000 pixels per plate). Image the underside of the plate with the top cover removed. This will produce the most consistent and clear images, allowing proper comparison of colony sizes across different plates.
15. The goal of the PCR analysis is to determine which 5-FOA resistant colonies contain expansions, as well as the size of individual expansions. An ideal solution would involve PCR analysis of every 5-FOA-resistant colony on a plate. To achieve a manageable number of colonies on each selective plate (e.g., ~10 per plate) you may need to dilute cells further or even plate several dilutions on 5-FOA plates in a given experiment. While performing PCR for each colony on every 5-FOA plate may become prohibitively costly and laborious, it eliminates sampling error from confidence intervals for the mutation rates. Because of this consideration, we recommend to PCR all 5-FOA resistant colonies; if it is not practically doable, three options to make PCR analysis feasible are presented below.
16. Option A is to plate a lower density of cells on each plate, or to count and PCR every colony only from a sector of each plate. This approach eliminates any kind of bias (especially by size)

that might occur while choosing colonies for PCR analysis. Of course, the subdivision of each plate must occur randomly in order to avoid unintentionally choosing a sector with larger or smaller colonies. One solution is to make a “Wheel-of-Fortune,” marking sectors on the top cover of a Petri plate, then place the “wheel” over each plate, spin, and mark each colony that lands within the prechosen sector. If there are a large number of colonies in the sector, a smaller sector may be used. Note, that the size of a sector should be the same for every 5-FOA plate within a batch. If very small sectors are needed, it may be better to perform the fluctuation test again with a lower concentration of cells. Option A also has the advantage of being mathematically simpler when performing mutation rate calculations. The caveat is that by plating fewer cells, the fluctuations in rate may be higher due to the greater number of dilutions and pipetting steps. Counting from sectors of a plate also assumes that the cells were spread evenly.

17. Option B is to PCR a set number of colonies on each plate, and to apply that percentage to the total number of 5-FOA-resistant colonies. Option B has several advantages. This method allows a higher plating density, provided that the colonies are not overcrowded to the point where they affect the growth of their neighbors. It also has the advantage of producing a predictable number of PCR reactions per fluctuation test, allowing more high-throughput processing. A caveat is that the user must choose a few colonies from each plate to represent the entire population, potentially introducing bias. Some of this bias can be eliminated, since we know roughly the relationship between colony size and mutation type. To best account for the various events that cause 5-FOA resistance, choose colonies by size in proportion to their approximate representation in the population, *i.e.*, if one out of four colonies are larger than average, then choose three average sized colonies and one larger colony for analysis. Although this tactic is not completely void of observer’s bias, it might provide a comprehensive analysis of the population. It is important not to choose any small background colonies if they were not included in the total colony count. Another caveat of Option B is that the additional step of the percentage calculation can affect the mathematics of the mutation rate calculation. Rate estimations can be skewed, particularly if expansions (or other events analyzed) do not represent the majority of the population. Because FluCalc applies plating efficiency adjustments, the calculated mutation rate in Option A is expected to be lower than in Option B. It is best, therefore, to be consistent in choosing either Option A, B or C for all experiments.

18. In Option C, as in Option B, a fixed number of colonies per plate are analyzed. However, this step is treated as the mathematical equivalent of a dilution, with a different dilution factor for each plate. Like Option A, the actual number of observed events can then be used for each calculation, though this requires applying an additional plating efficiency correction to each plate. The calculations accounting for this “variable dilution” are not currently integrated into FluCalc. *See Note 25* and refs. [12](#), [14](#) for more details.
19. In our current system, the repetitive tract starting length PCR is 350 bp, and expansions will typically be in the range of 375–1000 bp. A 1.5% agarose gel is appropriate for this size range. Switching to a 2–2.5% agarose gel in  $1\times$  TAE buffer can help provide better resolution when analyzing small-scale changes [[14](#)]. Voltage and running time should be adjusted for clear separation of bands within this size range, and size estimation to the nearest 25 bp. Bands can also be analyzed electronically, for example via the Bio-Rad QuantityOne program, for greater accuracy. We also recommend large format gels with multichannel-compatible wells, due to the high volume of PCR samples that must be analyzed per assay.
20. In some cases, a gel lane may contain two or more bands, indicating the presence of a secondary expansion or contraction that occurred following plating on 5-FOA (as opposed to secondary expansions that occurred during the nonselective growth stage, which will only produce a single band). We typically attribute the preplating event to the darkest band on the gel (representing the greatest number of cells within the colony), with the caveat that shorter repeat lengths typically result in a more successful PCR reaction. On the other hand, an extremely long expansion may push the limits of the PCR reaction, creating a smeared band. We typically use the darkest part of the smear as the length of the preplating expansion. If one of the two bands has the original size of the repeat, the expansion most likely occurred postplating and the colony should be counted as nonexpanded.
21. We have found that primary expansions occur in a normal distribution with a median of 50–60 repeats added [[9](#)]. Only experiments in which the expansion size histogram follows a unimodal distribution will produce an accurate estimate of mutation rate. A bimodal distribution may indicate the presence of secondary expansion events. Secondary expansions can be selected for if the 5-FOA concentration is too high. Secondary expansions become more and more prevalent as starting repeat tract length increases, in mutant strains with high expansion rates, or if the nonselective growth period is too long (because primary expansions will increase the likelihood of

further expansions in the next generation). It is also not uncommon for the distribution to reveal that smaller-scale expansions were not allowed to grow due to a higher 5-FOA concentration. Expansion rates calculated in nonnormally distributed assays will thus reflect an uncertain mixture of primary, secondary, and missed events.

22. This example protocol is used to determine the rate of repeat expansions. However, FluCalc can be used to determine rates for any type of observed mutation, depending on which set of  $C_{\text{sel}}$  values are entered. For example, the user can instead enter the number of colonies with an observed deletion, or the total number of 5-FOA-resistant colonies.
23. For Options A and C, enter the number of colonies with expansions after running gel electrophoresis. For Option B, the numbers entered for  $C_{\text{sel}}$  should be the number of observed 5-FOA-resistant colonies multiplied by the overall percent of expansions (*i.e.*, total number of colonies with expansions divided by total number of analyzed colonies for all cultures in the experiment).
24. If counting sectors (Option A), assume an even distribution of liquid sample during plating (e.g., if 100  $\mu\text{L}$  of solution was plated on 5-FOA and all colonies from a 1/8 wedge were analyzed for expansions then  $V_{\text{sel}} = 12.5 \mu\text{L}$ )
25. Corrected values are raw values corrected for plating efficiency, resulting in a more accurate representation of the mutation rate per cell per generation. If using Option C for PCR analysis, plating efficiency calculations must be performed manually. This is because a fixed amount of PCRs represent a variable amount of colonies, and thus the plating efficiency correction is not the same for each plate. We denote the following calculation method as VZ-MLE, or variable  $z$  maximum likelihood estimation. This method also relies on the assumption that the cultures were of roughly equal size (*see Note 8*). We hope to fully integrate VZ-MLE into FluCalc in the near future. At present, additional steps must be performed to employ FluCalc: Enter  $C_{\text{sel}}$  values as the actual number of observed expansions (e.g., 2 out of 4) obtained for each individual plate. Enter all other values ( $V_{\text{tot}}$ ,  $D_{\text{sel}}$ ,  $V_{\text{sel}}$ ,  $C_{\text{com}}$ ,  $D_{\text{com}}$  and  $V_{\text{com}}$ ) as defined in **step 1** of Subheading 3.5 and press “Calculate” to obtain the raw mutation rate and confidence intervals. Next, calculate the  $z$ -factor for each culture by dividing the number of colonies analyzed by PCR by total number of 5-FOA resistant colonies in the culture, which is  $C_{\text{sel}}D_{\text{sel}}V_{\text{tot}}/V_{\text{sel}}$ . From there, determine the plating efficiency coefficient ( $P = 1 - z / z \ln(z)$ ) for each plate, and then the average value of  $P$  (AVP). Then multiply the rate and confidential limits obtained with FluCalc by AVP.

26. **Step 5** of Subheading 3.2 describes resuspending and plating dilutions of each colony on corresponding 5-FOA and YPD plates. Thus the number of entries for  $C_{sel}$  and  $C_{com}$  should be the same, and FluCalc will determine individual frequencies for each corresponding set of plates, making the assumption that the first 5-FOA plate corresponds to the first YPD plate, and so forth. If the number of entries for  $D_{com}$  and  $C_{com}$  are not the same, FluCalc assumes that individual 5-FOA and YPD plates do not correspond, and instead uses the mean of  $C_{com}$  for each calculation.
27. The reported confidence intervals are accurate for any given experiment, but they may not fully reflect the inherent nature of the fluctuation assay. Due to the relatively high baseline rate of expansion events, the median expansion rate of our wild-type  $GAA_{100}$  strain does indeed fluctuate from trial to trial. This natural fluctuation is usually within  $\pm 50\%$  of the typical rate, however we do caution at making conclusions for mutant strains with any less than a threefold difference from the wild-type. Despite all of the complications and caveats, we find this assay to be a reliable and consistent method to determine expansion rates.

---

## Acknowledgments

We thank Alexander A. Shishkin and Kartik A. Shah for their invaluable contributions in developing cassettes to study repeat instability, and for developing experimental protocols for the selection and PCR procedures, Timofei S. Bondarev for developing FluCalc program, and Durwood Marshall for statistical consulting. This study was funded by NIH grants GM105473 and GM60987 to S.M.M and RFBR grant #15-04-08658 and research project in the Centre for Molecular and Cell Technologies (Research Park, Saint-Petersburg State University) for A.Y.A.

## References

1. Luria SE, Delbruck M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28(6):491–511
2. Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci U S A* 88:7160–7164. doi:10.1073/pnas.88.16.7160
3. Foster PL (2006) Methods for determining spontaneous mutation rates. *Methods Enzymol* 409:195–213. doi:10.1016/S0076-6879(05)09012-9
4. Zheng Q (2015) A new practical guide to the Luria–Delbrück protocol. *Mutat Res* 781:7–13. doi:10.1016/j.mrfmmm.2015.08.005
5. Zheng Q (2016) A second look at the final number of cells in a fluctuation experiment. *J Theor Biol* 401:54–63. doi:10.1016/j.jtbi.2016.04.027
6. Mirkin SM (2007) Expandable DNA repeats and human disease. *Nature* 447(7147):932–940. doi:10.1038/nature05977

7. Kim JC, Mirkin SM (2013) The balancing act of DNA repeat expansions. *Curr Opin Genet Dev* 23(3):280–288. doi:[10.1016/j.gde.2013.04.009](https://doi.org/10.1016/j.gde.2013.04.009)
8. Shishkin AA, Voineagu I, Matera R et al (2009) Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Mol Cell* 35(1):82–92. doi:[10.1016/j.molcel.2009.06.017](https://doi.org/10.1016/j.molcel.2009.06.017)
9. Shah KA, Shishkin AA, Voineagu I et al (2012) Role of DNA polymerases in repeat-mediated genome instability. *Cell Rep* 2(5):1088–1095. doi:[10.1016/j.celrep.2012.10.006](https://doi.org/10.1016/j.celrep.2012.10.006)
10. Shah KA, McGinty RJ, Egorova VI et al (2014) Coupling transcriptional state to large-scale repeat expansions in yeast. *Cell Rep* 9(5):1594–1602. doi:[10.1016/j.celrep.2014.10.048](https://doi.org/10.1016/j.celrep.2014.10.048)
11. Cherng N, Shishkin AA, Schlager LI et al (2011) Expansions, contractions, and fragility of the spinocerebellar ataxia type 10 pentanucleotide repeat in yeast. *Proc Natl Acad Sci U S A* 108(7):2843–2848. doi:[10.1073/pnas.1009409108](https://doi.org/10.1073/pnas.1009409108)
12. Aksenova AY, Greenwell PW, Dominska M et al (2013) Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proc Natl Acad Sci U S A* 110(49):19866–19871. doi:[10.1073/pnas.1319313110](https://doi.org/10.1073/pnas.1319313110)
13. Schmidt KH, Pennaneach V, Putnam CD et al (2006) Analysis of gross-chromosomal rearrangements in *Saccharomyces cerevisiae*. *Methods Enzymol* 409:462–476. doi:[10.1016/S0076-6879\(05\)09027-0](https://doi.org/10.1016/S0076-6879(05)09027-0)
14. Aksenova AY, Han G, Shishkin AA et al (2015) Expansion of interstitial telomeric sequences in yeast. *Cell Rep* 13(8):1545–1551. doi:[10.1016/j.celrep.2015.10.023](https://doi.org/10.1016/j.celrep.2015.10.023)
15. Zheng Q (2002) Statistical and algorithmic methods for fluctuation analysis with SALVADOR as an implementation. *Math Biosci* 176(2):237–252. doi:[10.1016/S0025-5564\(02\)00087-1](https://doi.org/10.1016/S0025-5564(02)00087-1)
16. Zheng Q (2016) rSalvador 1.5: an R tool for the Luria–Delbrück fluctuation assay. <http://eccceric.github.io/rSalvador>. Accessed 7 July 2016
17. Hall BM, Ma CX, Liang P et al (2009) Fluctuation AnaLysis CalculatOR: a web tool for the determination of mutation rate using Luria–Delbrück fluctuation analysis. *Bioinformatics* 25(12):1564–1565. doi:[10.1093/bioinformatics/btp253](https://doi.org/10.1093/bioinformatics/btp253)
18. Gillet-Markowska A, Louvel G, Fischer G (2015) bz-rates: a web tool to estimate mutation rates from fluctuation analysis. *G3 Bethesda* 5(11):2323–2327. doi:[10.1534/g3.115.019836](https://doi.org/10.1534/g3.115.019836)
19. Sarkar S, Ma WT, Sandri GH (1992) On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. *Genetica* 85(2):173–179. doi:[10.1007/BF00120324](https://doi.org/10.1007/BF00120324)
20. Rosche WA, Foster PL (2000) Determining mutation rates in bacterial populations. *Methods* 20(1):4–17. doi:[10.1006/meth.1999.0901](https://doi.org/10.1006/meth.1999.0901)

## Measuring Dynamic Behavior of Trinucleotide Repeat Tracts In Vivo in *Saccharomyces cerevisiae*

Gregory M. Williams and Jennifer A. Surtees

### Abstract

Trinucleotide repeat (TNR) tracts are inherently unstable during DNA replication, leading to repeat expansions and/or contractions. Expanded tracts are the cause of over 40 neurodegenerative and neuromuscular diseases. In this chapter, we focus on the (CNG)<sub>n</sub> repeat sequences that, when expanded, lead to Huntington's disease (HD), myotonic dystrophy type 1 (DM1), and a number of other neurodegenerative diseases. We describe a series of in vivo assays, using the model system *Saccharomyces cerevisiae*, to determine and characterize the dynamic behavior of TNR tracts that are in the early stages of expansion, i.e., the so-called threshold range. Through a series of time courses and PCR-based assays, dynamic changes in tract length can be observed as a function of time. These assays can ultimately be used to determine how genetic factors influence the process of tract expansion in these early stages.

**Key words** Trinucleotide repeat, Expansion, Contraction, Polymerase chain reaction, *Saccharomyces cerevisiae*, DNA replication, Repeat tract dynamics, Microsatellite instability

---

### 1 Introduction

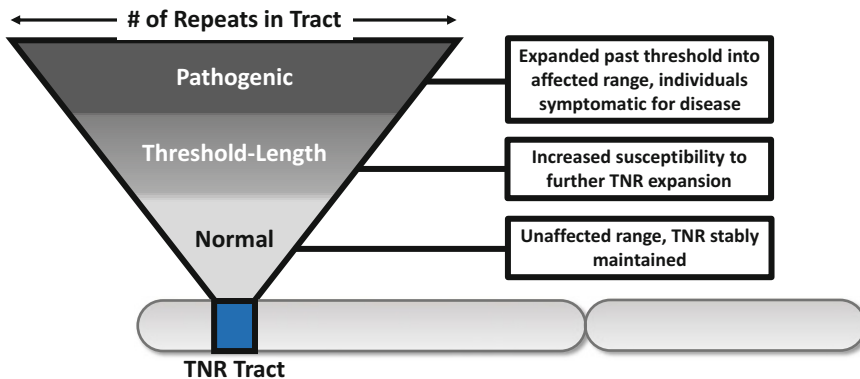
Trinucleotide repeat (TNR) sequences, like all repetitive sequences, pose a problem for DNA metabolism. Stretches of DNA containing repetitive nucleotides (mono-, di-, tri-, tetra-, etc.), known as regions of microsatellites are inherently difficult to replicate accurately. These repetitive sequences are hotspots for genetic mutation due to the increased frequency of replicative polymerase slippage events that can occur during replication through these regions [1–4]. This leads to an elevated rate of change in the number of repeats and is known as microsatellite instability. If left unrepaired, these slippage events are mutagenic and can be detrimental to the cell. (CNG)<sub>n</sub> TNR sequences present an additional hurdle to accurate DNA replication because of their propensity to form secondary structure (due to C–G base pairing) when the DNA is single-stranded, as can happen during replication, particularly during lagging strand synthesis [5, 6]. While the mismatch repair



(MMR) system, particularly Msh2-Msh3, typically directs repair of these slippage events [7–9], these structures are refractory to MMR [10]. In fact, Msh2 and Msh3 have been shown to promote TNR expansions in a variety of systems [6, 11–18].

The addition of TNRs leads to expansion of the TNR tract and can be within a coding region or a noncoding region of DNA. Expansions of TNR sequences are associated with several hereditary neurological diseases, including Huntington’s disease (HD), myotonic dystrophy (DM1), and Friedrich’s ataxia (FRDA) [18–21]. As the tract size increases, the expression and/or function of a protein is increasingly compromised, eventually leading to disease (Fig. 1). The transition from a stable repeat tract to an expanded, pathogenic tract is a critical step in the development of TNR-mediated disease, but is poorly understood. However, it is known that a small change in tract size, with the addition of only a few repeats, can mean the difference between stability and instability. Therefore, at this critical stage, a small reduction in tract length, or in the probability of expansion, has significant implications for the prevention and treatment of disease. Tract lengths substantially affect age of onset, penetrance of the disease, and severity of disease symptoms [22–24].

Initial expansions bring the tract into a “threshold-range” in which the tracts are not pathogenic but are increasingly unstable and are particularly prone to expansions [15, 17, 22, 25] (Fig. 1). Such intermediate alleles for HD (~27-35 CAG repeats) have recently been reported in ~6% of the general population [23], and these individuals are carriers for disease. Small increases in tract



**Fig. 1** TNR tracts within the normal range (which is tract-dependent) are stably maintained within that range. However, through mechanisms that remain unclear, a TNR tract can expand, increasing the number of repeats within the tract. Initially this brings the tract into a “threshold-length” range, in which these somewhat longer tracts are not pathogenic, but are increasingly susceptible to expansion; these individuals are carriers for disease. Once a tract has expanded sufficiently, it crosses a threshold; tracts above this threshold (which is disease-specific) are pathogenic and cause disease. As the size of the tract increases, it becomes increasingly unstable and prone to changes in length, particularly expansions

length make these individuals, and their offspring, at risk for disease. As a result, it is of interest to understand the dynamic behavior of TNR tracts within the threshold or intermediate range. Historically, the field has relied on end-point experiments, including single sperm typing [26–30] to assess changes in tract length. These studies led to modeling that predicted that threshold-length TNR tracts will preferentially increase in increments smaller than the repeats themselves [31–33]. Using the dynamic system described in this chapter, we recently demonstrated that threshold length tracts are highly dynamic, they increase (or decrease) in increments as small as a single repeat unit and the changes in tract length are biased toward expansion [17], consistent with our *in vitro* observations [6].

Assessing dynamic length changes in TNR tracts is challenging in higher eukaryotic systems [34]. Our laboratory has developed a set of genetic assays, in *Saccharomyces cerevisiae*, designed to assess repeat tract length with one repeat resolution under several different growth conditions, ranging from large heterogeneous cell populations in liquid culture down to repeat dynamics originating from a single cell. In this protocol, based on a system developed by the Lahue lab [37], we focus specifically on  $(CAG)_n$  and  $(CTG)_n$  repeat tracts within the threshold range, i.e., those poised to expand—or contract. We describe a series of related approaches that enable monitoring of the tract length as a function of time, using the model organism *Saccharomyces cerevisiae*. Using these approaches, we can detect dynamic tract length changes and start to understand the underlying mechanism(s) (including MMR) that lead to tract instability. These assays can also be used to determine the effect of a therapeutic drug on repeat tract dynamics as a function of time.

---

## 2 Materials

Make up all buffers with deionized water (dH<sub>2</sub>O). Store at room temperature unless otherwise noted.

### 2.1 General Reagents

1. 10× TBE: For 1 L: 108 g Tris base, 55.0 g boric acid, 7.44 g ethylenediaminetetraacetic acid (EDTA). The 0.5× TBE solution is 45 mM Tris–borate, 1 mM EDTA.
  - (a) Add each of the measured components into a beaker containing a stir bar and 500 mL of deionized water and place beaker onto a stir plate.
  - (b) Allow components to dissolve in solution (~1 h). Heat can be used to assist components into solution (not to exceed 60 °C).

- (c) Bring dissolved solution up to a final volume of 1 L with dH<sub>2</sub>O. Store at room temperature.
  - (d) To make polyacrylamide gel running buffer, dilute the 10× TBE buffer down to 0.5× TBE (950 mL dH<sub>2</sub>O + 50 mL 10× TBE).
2. Solutions for lithium acetate transformations:
- (a) 1 M lithium acetate (LiAc), pH 7.5 (adjusted with a few drops of diluted acetic acid). Autoclave.
  - (b) 10× TE (100 mM Tris-HCl pH 7.5, 10 mM EDTA, pH 7.5). Autoclave.
  - (c) 50% PEG (50% w/v). Filter-sterilize.
  - (d) LiAc-TE Mix: 1 mL of 1 M LiAc, 1 mL of 10× TE, and 8 mL of sterile dH<sub>2</sub>O. Make fresh.
  - (e) 40% PEG Mix: 1 mL of LiAc, 1 mL 1× TE, and 8 mL of 50% PEG. Make fresh.
3. Chromosome Prep Buffer:
- (a) 8.68 mL of H<sub>2</sub>O-Triton X-100 solution (8.48 mL of H<sub>2</sub>O and 200 μL Triton X-100).
  - (b) 1 mL of 10% SDS.
  - (c) 200 μL of 5 M NaCl.
  - (d) 100 μL of 1 M Tris pH 8.0.
  - (e) 20 μL of 0.5 M EDTA.

**2.2 Polymerase Chain Reaction (PCR)**

1. PCR Reaction Mixture:
- (a) 25 mM MgCl<sub>2</sub> (to 2 mM final concentration in reaction).
  - (b) 10× Taq Buffer (200 mM Tris pH 8.8, 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM KCl, and 1% Triton X-100).
  - (c) dNTP mix (2.5 mM each).
  - (d) *Taq* Polymerase (*see Note 1*).
  - (e) Primer SO295: 5' AA ACTCGGTTTGACGCCT CCCATG.
  - (f) Primer SO296: 5' AGCAACAGGACTAGGATGA GTAGC.

**2.3 Expansion Assays and Time Courses**

1. Synthetic Complete (SC) Plates (1 L): 7.0 g Yeast Nitrogen Base without amino acids (+Ammonium Sulfate) (Difco), 0.87 g Amino Acid drop-out mix (-histidine), 20.0 g agar (US Biological), 950 mL sterile dH<sub>2</sub>O.
- (a) Add stir bar and all measured contents into a 2 L flask, capping the top with aluminum foil.
  - (b) Autoclave at 121 °C for 40 min.

- (c) Remove from autoclave and add 50.0 mL of sterile 40% glucose to flask.
  - (d) Allow flask to cool before pouring into plates. Store at room temperature or at 4 °C.
2. Liquid Synthetic Complete Media (1 L): 7.0 g Yeast Nitrogen Base without amino acids (+Ammonium Sulfate) (Difco), 0.87 g of –Histidine Amino Acid drop-out mix, 950 mL sterile dH<sub>2</sub>O.
- (a) Add all contents into a 1 L screwcap bottle.
  - (b) Autoclave at 121 °C for 40 min.
  - (c) Remove from autoclave, add 50.0 mL of sterile 40% glucose to bottle and allow media to cool to room temperature before use.
  - (d) Liquid media can be stored for long periods of time without addition of 40% glucose.
3. Synthetic Complete (SC) Plates containing 5-FOA (1 L):
- (a) In a 2 L flask, add a stir bar, 700 mL of deionized water, 20.0 g of Agar (US Biological), and 20.0 g of Glucose (US Biological). Autoclave at 121 °C for 40 min.
  - (b) In a separate beaker containing a stir bar and 200 mL of deionized water, add 7.0 g Yeast Nitrogen Base without amino acids (+ammonium sulfate) (Difco), 0.87 g of –Histidine –Uracil Amino Acid drop-out mix, 50 mg of uracil, and 1.0 g 5-FOA (0.1%).
  - (c) Stir contents of beaker on a stir plate until dissolved (~1–2 h). Bring the total volume up to 300 mL using dH<sub>2</sub>O.
  - (d) Filter-sterilize the 5-FOA solution using a 0.2 µm filter.
  - (e) Once the flask of agar and glucose cools to approximately 80 °C, add the 5-FOA solution to the flask, allowing it to completely mix on the stir plate. Immediately pour into plates.
  - (f) Allow the plates to cool, then store at 4 °C. Note that 5-FOA is light sensitive.
4. 12% Polyacrylamide, 0.5 × TBE Gels
- (a) Acrylamide–bis solution (29:1).
  - (b) 10 × TBE.
  - (c) 10% ammonium persulfate (APS).
  - (d) TEMED.
  - (e) Horizontal gel apparatus, 16 cm × 14 cm glass plates, thin spacers, and a 20-well comb.
  - (f) 10 × polyacrylamide loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, and 30% glycerol).

- (g) Low molecular weight DNA Marker (e.g., NEBiolabs).
- (h) Plastic staining tray.
- (i) Ethidium bromide.

## 2.4 Southern Blotting

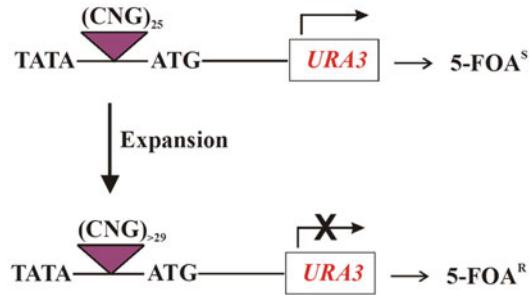
1. Denaturation buffer: 0.5 N NaOH and 1.5 M NaCl.
2. Neutralization buffer: 0.5 M Tris-HCl pH 7.0, 3.0 M NaCl. Start with ~700 mL dH<sub>2</sub>O, bring Tris and NaCl into solution. Adjust pH to 7.0 with concentrated HCl and then bring final volume (1 L) with dH<sub>2</sub>O.
3. 10× SSC: dissolve 87.7 g of NaCl and 44.1 g of sodium citrate in ~800 mL dH<sub>2</sub>O. Adjust pH to 7.0 with a few drops of 10 N NaOH and adjust to 1 L dH<sub>2</sub>O. This concentrated stock of NaOH is required to quickly adjust the pH.
4. 1.85 MBq (50 μCi) [ $\alpha$ -<sup>32</sup>P] dCTP (~111 TBq/mmol, 370 MQq/mL).
5. Random Priming DNA labeling kit (e.g., TaKaRa).
6. Stratalinker or equivalent UV crosslinker.
7. HYBAID hybridization oven, or similar.
8. Hybridization bottles.
9. STORM PhosphorImager, or equivalent.
10. Storage phosphor screen.

---

## 3 Methods

### 3.1 Construction of Strains Containing TNR Tracts

1. A *URA3* reporter system construct is used to select for expansion events in vivo [35–37]. TNR substrates are integrated into each strain of interest, using plasmids that contain the reporter construct, tract sequences and sequences to integrate the system at the *LYS2* locus [6, 17, 37]. Plasmids encoding (CAG)<sub>25</sub>, (CTG)<sub>25</sub>, and scrambled (C,A,G)<sub>25</sub> or (C,T,G)<sub>25</sub> are used [36–38]. Each repeat tract is cloned into the regulatory region controlling expression of the *URA3* reporter gene. When the distance between the TATA box and the initiator ATG for the *URA3* gene is increased beyond 29 repeats, *URA3* is no longer expressed, making the cells resistant to 5-FOA (*see Note 2*) (Fig. 2).
2. Each plasmid is digested with *Bsu36I* to linearize and is then transformed, by the lithium acetate method [39], to allow integration at the *LYS2* locus by homologous recombination.
3. For each transformation, set up a small culture in liquid media from a single colony. Incubate overnight in a 30 °C shaker.
4. From the saturated culture, dilute 1:30 to 1:50 to start a new 20–30 mL liquid culture in a 125 mL sterile flask. Incubate in a



**Fig. 2** TNR reporter system used to select expansions. The *URA3* gene is used as a reporter for expansions on the lagging strand [36]. The *Schizosaccharomyces pombe adh1* promoter is fused upstream of the *URA3* reporter gene. This promoter has a stringent requirement for the distance between TATA and ATG (55–125 bp) for function. Insertion of a trinucleotide repeat (TNR) makes *URA3* expression dependent on repeat tract length. Cells with an unexpanded TNR tract will express *URA3* and exhibit sensitivity to the drug 5-FOA, which is toxic to the cell in the presence of the Ura3 gene product. Expansions within the TNR tract of 4 or more repeats alters the transcriptional start site, effectively inactivating *URA3* expression and resulting in 5-FOA resistance. The TNR construct is integrated into the chromosomal *LYS2* locus by homologous recombination with *HIS3* as a selectable marker. Single integration is confirmed by Southern blotting

30 °C shaker until the culture reaches mid-logarithmic growth (OD<sub>600</sub> ~ 0.5–0.8).

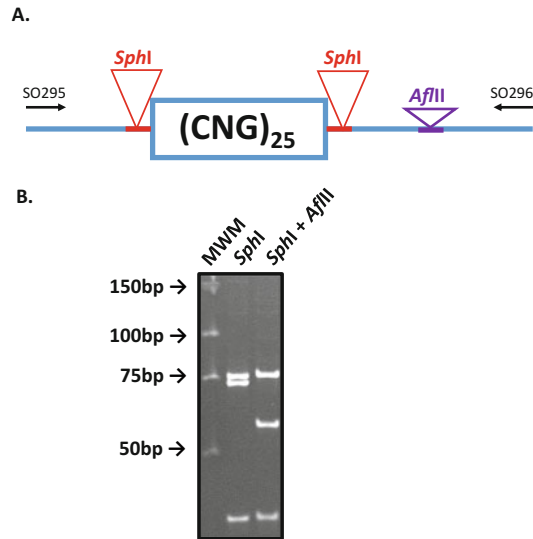
5. Collect the cells by centrifugation (3000 × *g*). Wash the pellet in 1 mL sterile dH<sub>2</sub>O. Transfer to a microcentrifuge tube.
6. Collect the cells by centrifugation (16,000 × *g*) and discard the supernatant. Wash the cells with 1 mL LiAc–TE mix. Collect the by centrifugation again and resuspend the pellet in 250 μL LiAc–TE mix.
7. For each transformation, mix 50 μL of the yeast suspension, ~150 ng digested plasmid and 10 μL sonicated salmon sperm DNA (10 mg/mL). Always do a negative control with no plasmid DNA added.
8. After mixing, add 600 μL 40% PEG mix.
9. Incubate reactions in a 30 °C shaker for at least 30 min. Longer incubation times are fine.
10. Incubate the reactions in a 42 °C water bath for 15 min to heat shock.
11. Briefly centrifuge the reaction (30 s), discarding the supernatant. Resuspend the cells in 200 mL 1× TE, pH 7.5. Select for transformants by plating on SC –histidine. Incubate in 30 °C cabinet for ~3 days.
12. Individual transformants were struck out to single colonies on SC –histidine plates. A single colony from each candidate was

frozen down for storage in 20% glycerol until ready to test for single integration of the reporter construct.

13. Perform Southern blotting (*see* below) to ensure that a single copy of the reporter construct has been integrated [37]. This will also confirm that the integration has occurred at the correct chromosomal locus.

### 3.2 TNR Expansion Assay

1. Using a sterile wooden applicator, take a small amount of cells from a frozen glycerol (20%) stock of the strain that is to be tested. Make a small patch at the top of a SC –histidine plate. Using a new wooden applicator, start toward the bottom of the patch and make a small vertical line toward the bottom of the plate, streaking a small amount of cells away from the patch. With a new wooden applicator, cross this vertical line once and zig-zag the applicator back and forth across the plate, separating your patch of cells out to single cells on the plate. Place plate into the 30 °C incubator for 3–4 days allowing the single cells to grow into ~2 mm colonies.
2. Select three individual colonies and mark them on the back of the plate, labeling them 1, 2, 3.
3. Perform colony PCR on these colonies to confirm that the repeat tract has not expanded. In this system, the starting repeat tract size is 25 repeats (75 bp). Perform PCR as follows:
  - (a) Combine, per reaction: 12.3 µL sterile dH<sub>2</sub>O, 1.6 µL 25 mM MgCl<sub>2</sub>, 2 µL 10× Taq Buffer, 2 µL dNTP mix (2.5 mM each), 0.8 µL of 5 pmol/µL SO295, 0.8 µL of 5 pmol/µL SO296, and 0.5 µL *Taq* Polymerase (*see* Note 1).
  - (b) Using a sterile 200 µL pipette tip, take a very small sample of the colony by barely touching the tip to the colony to obtain a small number of cells on the tip. Resuspend the cells in 20 µL of your PCR mix, making sure to put the tip directly into the PCR mix and twirl. Make sure that all liquid that makes its way into the pipette tip through capillary action is cleared back into the PCR tube, by pushing down on the top of the tip with your finger.
  - (c) 5 min incubation at 95 °C; 35 cycles of 2 min at 95 °C; 1 min at 53 °C, 1 min at 72 °C. End with a 10-min incubation at 72 °C.
4. Digest the PCR reactions using *Sph*I and *Afl*II. Make a master digestion mix to be added directly to the 20 µL PCR reaction, as follows. The 10 µL mix for 1 reaction, i.e., the amount to be added to 1 PCR reaction is: 0.2 µL *Sph*I (NEB), 0.2 µL *Afl*II (NEB), 3.0 µL NEB Buffer 2.1, and 6.4 µL sterile dH<sub>2</sub>O.
  - (a) Add 10 µL of digestion mix to each 20 µL PCR reaction.
  - (b) Incubate digestion reaction for 1 h and 30 min in a 37 °C water bath.



**Fig. 3** Digestion of TNR tract PCR. (a) The TNR tract length can be determined by PCR, using primers upstream and downstream of the repeat sequence (SO295 and SO296), resulting in a PCR fragment of 188 bp. The amplified regions of DNA immediately upstream and downstream of the repeat tract can be digested with *SphI*, freeing the TNR tract. (b) Lane 2 shows that *SphI* digestion results in three distinct bands, with a doublet at 75 bp (the 25 repeat TNR tract) and 73 bp and a 41 bp fragment. To eliminate the doublet for easier assessment of the TNR tract size, the 73 bp fragment is further digested with *AflII*, resulting in a 59 bp fragment, as seen in Lane 3. A 14 bp fragment is not visible on this gel. Lane 1 shows the low molecular weight marker (LMWM, NEBiolabs)

- (c) Store digested reactions at  $-20^{\circ}\text{C}$ , if necessary.
5. Run the digested reaction on a 12% polyacrylamide,  $0.5\times$  TBE gel to confirm the tract size (Fig. 3):
  - (a) Make and pour gels using a full piece of glass ( $16\text{ cm} \times 14\text{ cm}$ ), notched piece of glass, thin spacers, and a 20-well comb.
  - (b) Clean the inside of the glass plates with ethanol and a Kimwipes to ensure that they are free of any dust or debris. Insert spacers between the two cleaned pieces of glass and clamp the outside using six binder clips; two clips per side. Insert the 20-well comb. The outside edges of the glass can be sealed with 0.5% agarose in  $\text{dH}_2\text{O}$ , making sure to add extra agarose to the corners.
  - (c) Make the gel solution (30 mL, 2 gels): 16.16 mL of sterile  $\text{dH}_2\text{O}$ , 12 mL of acrylamide–bis solution (29:1), 1.5 mL of  $10\times$  TBE, 0.3 mL of 10% ammonium persulfate (APS), and 32  $\mu\text{L}$  of TEMED. Make sure to add the components in the order that they are listed, with the APS and the TEMED added last. The solution will begin to solidify



- within 7–10 min. The gel solution can be added into the sealed glass plates using a 5 mL serological pipette, placing the tip of the pipette in the corner of the clamped plates.
- (d) After gels solidify, remove the bottom spacer and submerge into  $0.5\times$  TBE in the gel apparatus, making sure that there are no bubbles between the glass plates below the gel. Assemble the gel apparatus and fill with  $0.5\times$  TBE.
  - (e) Using a gel loading tip, rinse each well to make sure that there is nothing in it.
  - (f) Add 3  $\mu\text{L}$  of  $10\times$  polyacrylamide loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, and 30% glycerol) to each sample. Mix well.
  - (g) Load 10  $\mu\text{L}$  of your digested sample into each well. You should have  $\sim 20$   $\mu\text{L}$  remaining, which can be stored at  $-20$   $^{\circ}\text{C}$ .
  - (h) Make sure to load a plasmid control (last lane) as well as the Low Molecular Weight Marker (10  $\mu\text{L}$  of loading dye, 1–2  $\mu\text{L}$  of marker).
  - (i) Run gels for 2 h at 250 V.
  - (j) Stain the gel for 10 min in 250 mL of deionized water containing 0.5  $\mu\text{g}/\text{ml}$  ethidium bromide.
6. After identifying an unexpanded tract from a colony, remove this colony from the plate using a sterile flat end toothpick, and resuspend it in 100  $\mu\text{L}$  of sterile  $\text{dH}_2\text{O}$  in a microcentrifuge tube. Vortex the tube to resuspend the colony, and serial dilute the resuspension to  $10^{-4}$  (10  $\mu\text{L}$  of undiluted resuspension into 990  $\mu\text{L}$  of sterile deionized water, vortex, 10  $\mu\text{L}$  of  $10^{-2}$  dilution into 990  $\mu\text{L}$  of sterile deionized water to get 1 mL of  $10^{-4}$  dilution). Make sure to vortex in between dilutions.
  7. Plate 20  $\mu\text{L}$  of the  $10^{-4}$  dilution onto a SC –histidine plates and incubate at 30  $^{\circ}\text{C}$  for 3–4 days, allowing the colonies to grow to  $\sim 2$  mm in diameter.
  8. Select ten colonies to test in the TNR expansion assay. Using a ruler, measure the diameter of the colonies on the plate, selecting colonies that are  $\sim 2$  mm in size. It is important to make sure that all of your selected colonies are approximately the same size; colonies of similar size will have undergone roughly the same number of cell divisions.
  9. Remove 2 mm colonies from the plate using a sterile flat ended toothpick and resuspend in 100  $\mu\text{L}$  of sterile deionized water and resuspend by twisting the toothpick in the microcentrifuge tube. After the colony is resuspended in water, discard the toothpick, close the lid on the microcentrifuge tube, label,

and vortex to further resuspend the colony. Dilute the resuspension to  $10^{-4}$  (*see step 3*), making sure to save each of your serial dilutions. Each colony that has been resuspended will be plated on both selective and permissive plates to determine an accurate expansion frequency.

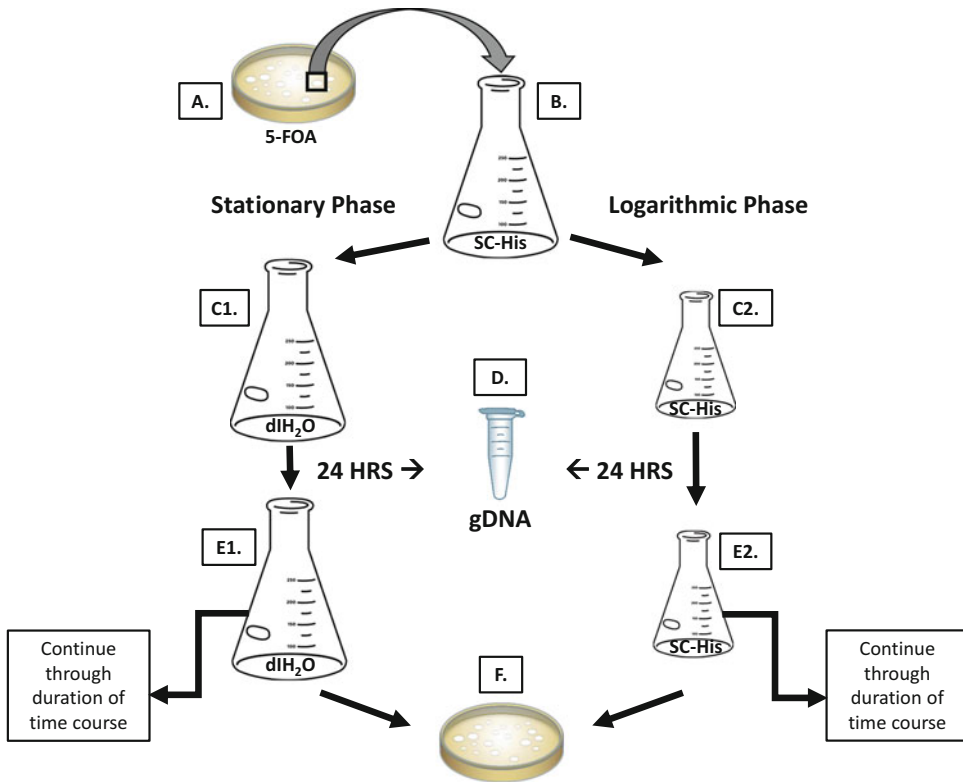
10. Plate 20  $\mu\text{L}$  of the  $10^{-4}$  dilution onto SC –histidine and spread using sterile glass beads, making sure that the resuspension is evenly spread on each plate (*see Note 3*). These are the permissive plates; all cells will grow. These plates are important to determine the total number of cells tested.
11. Plate 75  $\mu\text{L}$  of the undiluted resuspension onto SC –histidine plates containing 5-FOA. These are the selective plates, only allowing growth of cells that lack *URA3* expression. In this system, this typically means that the strain contains an expanded TNR tract greater than or equal to 29 repeats (*see Note 4*). Spread the resuspension on the plate using sterile glass beads.
12. Place both the selective and permissive plates into a 30 °C incubator and allow them to grow for 3–4 days until the majority of colonies on the plates reach a diameter of ~1–2 mm so that they are easy to count. It may be necessary to allow the selective plates to grow a day longer to allow for sufficient growth in the presence of 5-FOA.
13. Remove plates from the incubator and count the number of colonies on all plates. After counts have been obtained for both selective and permissive plates, the numbers for each colony will be subjected to analysis in order to determine the expansion frequency of the strain.

### 3.3 Mutation Rate Calculations

1. To calculate apparent expansion rates in the TNR expansion assay, colonies on selective and permissive plates are counted. Expansion rates are calculated by iterative formulas, as described by Drake [40]. The 95% confidence intervals are determined from tables of confidence intervals for the median [41, 42].
2. To calculate the mutation rates for the time course experiments (*see Subheadings 3.4–3.6*), the probability of a change in tract size was treated as a binomial distribution with  $p$  = the proportion of tracts with a change in length and  $q$  = the proportion of tracts with no length change.
3. The rate of tract change was defined as the number of changes observed/the number of cells examined per generation. It is therefore important to determine the doubling time (or generation time) of each strain prior to calculating mutation rates. Growth curve experiments in both liquid media and on plates are used to determine the number of generations for these calculations (*see Subheading 3.9*). To calculate 95% confidence intervals (95% C.I.) on the mutation rates, the F statistic are used [43, 44].

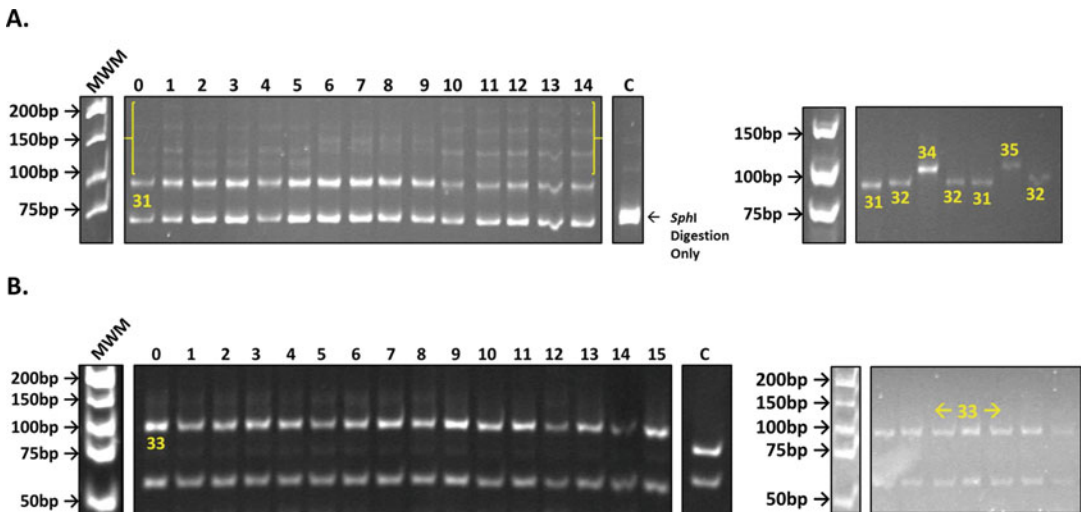
**3.4 Liquid Time Course Protocol**

1. Perform a liquid time course to assess tract length and tract length changes in a large, heterogeneous cell population. By performing assays with replicative and stationary phase cultures, it is possible to determine the contribution of DNA replication to tract instability in this system (Fig. 4).
2. Perform TNR expansion assay (*see* Subheading 3.2) to select for colonies with expanded repeat tracts.



**Fig. 4** Schematic of the liquid time course experiment protocols. (a) Individual colonies with TNR tract expansions are selected on plates containing 5-FOA; the tract length increase was confirmed by colony PCR. (b) A single colony with an expanded tract is used to inoculate a 106 mL large culture that was grown to saturation over a 72-h period (18 generations) at 30 °C. Five microliter of this starting culture is removed and genomic DNA (gDNA) was isolated to obtain time point zero. Then, parallel stationary (c1) and logarithmic (c2) phase cultures are established and propagated for 14 days by taking 200  $\mu$ L of the 24-h saturated culture and adding it to a new flask containing 10 mL of SC –His media (e1, e2). (d) gDNA was prepared from each culture every 24 h (6 generations) and subjected to PCR to evaluate tract length. (f) A sample from each log- and stationary-phase culture is diluted  $10^{-4}$  and plated to isolate individual colonies that are subjected to colony PCR to assess tract length. This is an end point experiment that allows for the observation of additional expansions as a function of time as well as general trends in the cell population. Colony PCR on colonies plated from specific time points permits the analysis of individual tract lengths from single colonies. Fitness effects of tract lengths are mitigated by limiting the number of cell divisions (6 per time point), although we previously determined that the tract length does not affect growth rate with our strains [17]. This experiment is similar to that of a mutation-accumulation experiment and indicates that additional expansions are observable in a 2-week time span

3. To confirm the presence of an expansion, perform colony PCR on colonies from selective plates (*see* Subheading 3.2, **step 3**).
4. Digest the DNA with *SphI* and *AfIII* (*see* Subheading 3.2, **step 4**) and run samples on a 12% polyacrylamide, 0.5× TBE gel (*see* Subheading 3.2, **step 5**). This will also allow you to determine the initial size of the expanded tract (*see* **Note 5**).
5. Resuspend confirmed expanded colonies in 106 mL of liquid SC –histidine (*see* **Note 6**) and grow the culture to saturation in a 30 °C shaker for 3 days (72 h) (Fig. 4).
6. From this saturated culture, remove 5 mL to isolate genomic DNA (*see* Subheading 3.5) that will be used as a template for PCR to determine the average tract size in the culture. This is the Day 0 time point.
7. Then take 200 μL and add it to a fresh flask containing 10 mL of liquid SC –histidine (1:50 dilution) to propagate the next time point. Be sure to use sterile technique when sampling the culture or transferring culture to new flasks. Place this flask in the 30 °C shaker for 24 h. This is your logarithmic phase, or replicative culture. The cells will grow and divide in this culture (Fig. 5a).

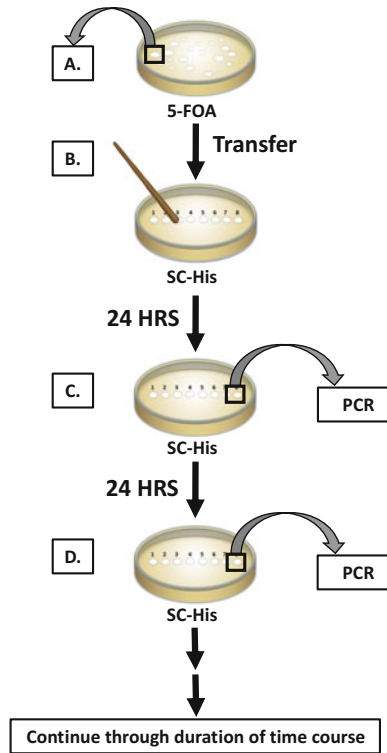


**Fig. 5** Liquid time course experiments of wild-type (CTG)<sub>25</sub> logarithmic-phase (**a**) and stationary-phase (**b**) cultures to determine population tract dynamics, as described in Fig. 4. The initial expansion size is indicated in yellow at time point zero. The numbers across the top of the gel indicate the day of the time course. The brackets at the top of the gel (**a**, *left panel*) indicate the progressive accumulation of a larger TNR tract with the concomitant loss of the initial expansion size toward the end of the time course. Stationary-phase repeat tracts (**b**, *left panel*) remain stable. A sample from Day 14 log-phase (**b**, *right panel*) and stationary-phase (**b**, *right panel*) culture were diluted and PCR was performed on individual colonies to assess individual tract lengths within the population. The length of the tracts was determined using a standard curve measured from the known bands of the low molecular weight marker (LMWM) run with each gel (*see* **Note 5**). The number above or below each tract indicates the number of repeats within each tract. C, PCR on a TNR plasmid control is performed alongside each time course as a marker for the unexpanded 75 bp tract (25 repeats). Expansions are observed in the actively replicating log-phase cultures (**a**), while stationary-phase tracts remained stable through the duration of the time course (**b**)

8. Using sterile technique, pour the remainder of the saturated culture into a 50 mL conical tube. Collect cells by centrifugation at  $3000 \times g$  for 5 min. Remove the supernatant, making sure to keep the pellet formed at the bottom of the tube. Add 10 mL of sterile dH<sub>2</sub>O to the tube, pipetting the pellet back into solution. Collect the resuspended cells by centrifugation ( $3000 \times g$ , 5 min.) and remove the supernatant. Repeat this wash step with 10 mL of sterile dH<sub>2</sub>O, again dumping out the supernatant. Resuspend the pellet in 10 mL of sterile dH<sub>2</sub>O and add to 90 mL of sterile dH<sub>2</sub>O in a 500 mL sterile flask. This will serve as your stationary phase culture. The cells will survive but will exhibit very little growth or division without nutrients. Incubate this stationary phase culture at 30 °C for 24 h (Fig. 5b).
9. Every 24 h, remove both logarithmic phase and stationary phase cultures from the 30 °C shaker. Remove 5 mL from each culture to extract genomic DNA, which will be used as template for PCR to determine TNR tract size. Remove 10 μL from each culture and add it to a microcentrifuge tube containing 990 μL of sterile deionized water, resulting in a  $10^{-2}$  dilution of your culture. Vortex the dilution and serial dilute it down to  $10^{-4}$  by removing 10 μL of the  $10^{-2}$  dilution and adding it to another microcentrifuge tube containing 990 μL of sterile deionized water. Vortex, and plate 20 μL of the  $10^{-4}$  dilution directly on synthetic media plates lacking histidine and spread using the glass bead method. Incubate plates at 30 °C for 3 days. To determine the range of expansions within the population, as opposed to the average, colony PCR can be performed on colonies to determine tract sizes from individual colonies from the liquid culture. This can be done at each time point throughout the time course. We typically plate cells on Days 1, 7, and 14 to monitor the health of the culture as well as tract size (Fig. 5).
10. Take 200 μL of your newly saturated logarithmic phase cultures and add them to a flask containing 10 mL of fresh SC –histidine and incubate in a 30 °C shaker for 24 h.
11. The cells in the stationary phase cultures should be washed and resuspended in fresh sterile dH<sub>2</sub>O every 48 h. Using sterile technique, collect the cells by centrifugation at  $3000 \times g$  for 5 min. Resuspend each pellet in fresh sterile dH<sub>2</sub>O, making sure to keep the total volume accurate: (100 mL) – (days  $\times$  5 mL) = current volume. Changing the water in these cultures helps prevent cell death.
12. Repeat **steps 7–9** throughout the entirety of the time course.
13. Throughout the time course, use the isolated genomic DNA (gDNA) as a template for PCR to determine the average tract size in the culture (Fig. 5).

### 3.5 Genomic DNA Preparation

1. Collect cells by centrifugation at  $3000 \times g$  and discard supernatant. You should have a pellet at the bottom of the tube.
2. Add 500  $\mu\text{L}$  of sterile  $\text{dH}_2\text{O}$  to each tube and vortex to resuspend the pellet. Transfer resuspension to a microcentrifuge tube.
3. Centrifuge for 30 s at  $16,000 \times g$ .
4. Remove tubes from microfuge and pour out the supernatant. Vortex the pellet in the residual liquid to resuspend.
5. Add 200  $\mu\text{L}$  Chromosome Prep Buffer.
6. Add 200  $\mu\text{L}$  of well-mixed phenol–chloroform (1 part phenol and 1 part chloroform, vortex to mix, and spin down at  $1200 \times g$  for 5 min. Store at  $4^\circ\text{C}$ , cover tube with aluminum foil, it is light sensitive).
7. Add 0.3 g of acid-washed glass beads (*see* **Note 7**).
8. Vortex tubes at  $400 \times g$  in the Mixmate (Eppendorf) in cold room for 3 min.
9. Add 200  $\mu\text{L}$  of  $1 \times \text{TE}$ , pH 8.0 to each vortexed tube.
10. Centrifuge for 5 min at  $16,000 \times g$ .
11. While waiting for your samples to spin down, set up clean microcentrifuge tubes and add 10  $\mu\text{L}$  of 4 M ammonium acetate to each tube.
12. When your samples have finished the centrifugation, transfer the supernatant into the ammonium acetate-containing tubes. Add 1 mL of 95% ethanol (EtOH) to each tube and invert two times. Your reactions are stable at this point.
13. Spin down your samples for 2 min in the microfuge at  $16,000 \times g$ .
14. Pour out EtOH, making sure not to pour out the pellet at the bottom of the tube. Add 1 mL of 70% EtOH to wash each pellet. Gently invert your tubes a few times.
15. Briefly centrifuge your samples for 30 s at  $16,000 \times g$ . Remove residual EtOH with your pipette.
16. Leave the tops on your samples open and let stand for 15–30 min at room temperature for the EtOH to evaporate and your samples to dry. Alternatively, you can use a vacuum apparatus to dry your samples; allow 10 min to dry under vacuum. Before moving on to the next step, make sure that all EtOH is evaporated from your samples, as the presence of EtOH can affect downstream applications.
17. Resuspend your dry pellet in 50  $\mu\text{L}$  of  $1 \times \text{TE}$  (pH 8.0) with 50  $\mu\text{g}/\text{mL}$  RNase. NEVER vortex your gDNA. This gDNA can be used as a template for PCR, using 1  $\mu\text{L}$  of genomic DNA sample, instead of a colony, in each 20  $\mu\text{L}$  PCR reaction (*see* Subheading 3.2, step 3).



**Fig. 6** Schematic of the colony time course experiment protocol. **(a)** Individual colonies with TNR tract expansions are selected on plates containing 5-FOA; the tract increase is confirmed by colony PCR in each case. **(b)** A single colony with an expanded tract is transferred to nonselective medium (SC – His) and allowed to continue to grow. Plates are then placed back into the 30 °C incubator. **(c)** Each colony is subjected to colony PCR every 24 h (~10 generations), sampling from around the perimeter of the colony, where the cells continued to grow. Plates are returned to the 30 °C incubator, and colony PCR performed again after 24 h **(d)**. This approach allows examination of the dynamics of a single TNR tract as a function of time. Because the PCR is derived from a sub-sample of an actively dividing colony, this is necessarily examining a mixed population, as evidenced by the multiple bands observed in some lanes, although the colony is derived from the same original cell

### 3.6 Colony Time Course Protocol

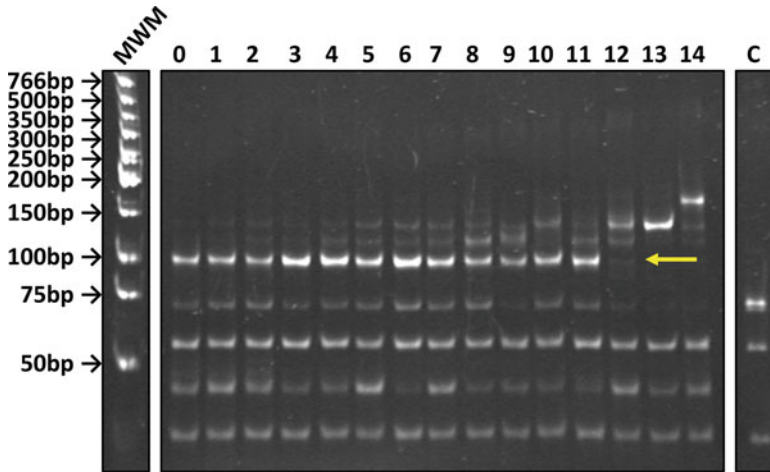
1. Perform a colony time course experiment to assess tract dynamics in a single colony, derived from a single cell, over time (Fig. 6).
2. Perform TNR expansion assay to select for colonies with expanded repeat tracts.
3. Perform colony PCR on colonies from selective plates (*see* Subheading 3.2, **step 3**), digesting with *SphI* and *AflII* (*see* Subheading 3.2, **step 4**), and run samples on a 12% polyacrylamide gel to determine the initial size of the expanded tract (*see* Subheading 3.2, **step 5**).

4. After confirming the size of the initial TNR tract expansion, select colonies for the time course experiment. Different sized expansions can be selected to test, i.e., threshold length tracts or affected range tracts. Mark these colonies on the back of the plate.
5. Using a sterile flat end toothpick, swipe the entire colony off from selective plate containing 5-FOA and deposit it onto a new SC –histidine plate and label the colony appropriately. It is important to remove confirmed colonies from the selective plate to allow for dynamic expansion and contraction events. Maintaining the colonies on plates containing 5-FOA will prevent growth of cells that sustain contractions that result in recovery of *URA3* expression, thereby potentially creating bias toward expansions.
6. After transferring colonies to new plates, perform colony PCR on each of the colonies to obtain the “time point zero” for the time course experiment (*see* Subheading 3.2, steps 3 and 4). Place the plates back into the 30 °C incubator for 24 h.
7. After 24 h, remove the plates and perform colony PCR on each individual colony (*see* Subheading 3.2, steps 3 and 4). When taking a small amount of the colony with a sterile tip for colony PCR, you may sample from either the same portion of the colony (i.e. 12 o’clock), or different portions of the colony (i.e., 12, 3, 6, and 9 o’clock) at different time points. Because the cells are always sampled from the perimeter of the colony, the cells from the colony continue to grow and fill in the sampled space. The growth rate on solid media can be determined (*see* Subheading 3.12) to determine the generation time of each strain on these plates (*see* Note 8).
8. Return plates to the 30 °C incubator for 24 h.
9. Repeat steps 6 and 7 for 14 days. If necessary, transfer colonies to new plates by swiping the entire colony off of the old plate and depositing it onto a fresh synthetic medium plate lacking histidine.
10. Colony PCR reactions for each time point should be digested with *SphI* and *AflIII* to assess tract length over time (Fig. 7).

### 3.7 Microcolony Time Course Protocol

1. Perform a microcolony time course to examine tract dynamics starting from a single cell, i.e., a completely homogeneous starting point (Fig. 8).
2. Perform TNR expansion assay to select for colonies with expanded repeat tracts.
3. Perform colony PCR on colonies from selective plates, digesting with *SphI* and *AflIII* (*see* Subheading 3.2, steps 3 and 4),

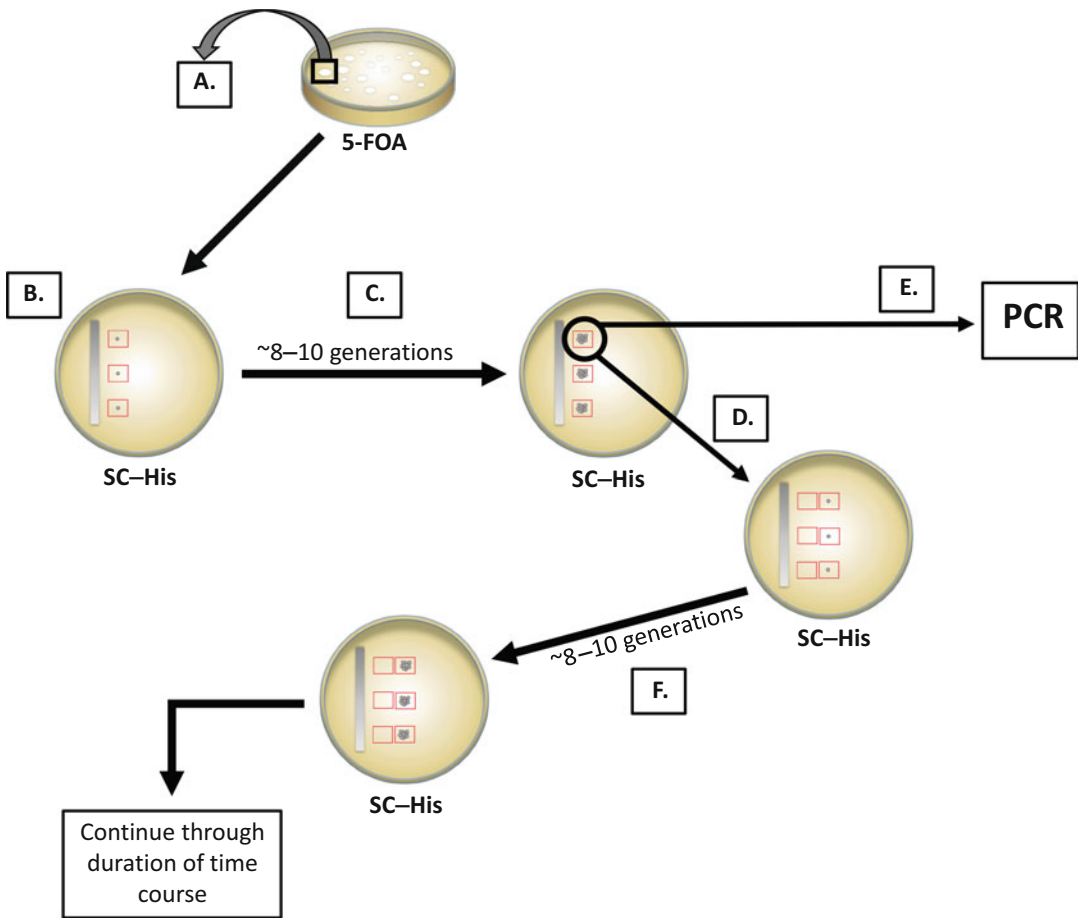




**Fig. 7** Progressive expansion events wild-type colony time course. Expansion events were selected in the wild-type (CTG)<sub>25</sub> background and confirmed by PCR. Individual colonies were followed over a 14-day time period; colony PCR was performed on each colony every 24 h to amplify the TNR tract (*see* Fig. 6). At the end of the time course, each PCR reaction was digested and resolved on a 12% polyacrylamide gel and stained with EtBr. The numbers across the top of the gel indicate the day of the time course. The lanes marked C in each panel indicate the 75-bp tract amplified from the TNR plasmid control. An uncropped gel is shown in this figure to highlight the different non-tract digestion products that appear below the TNR tract. The *yellow arrow* indicates the initial expansion product

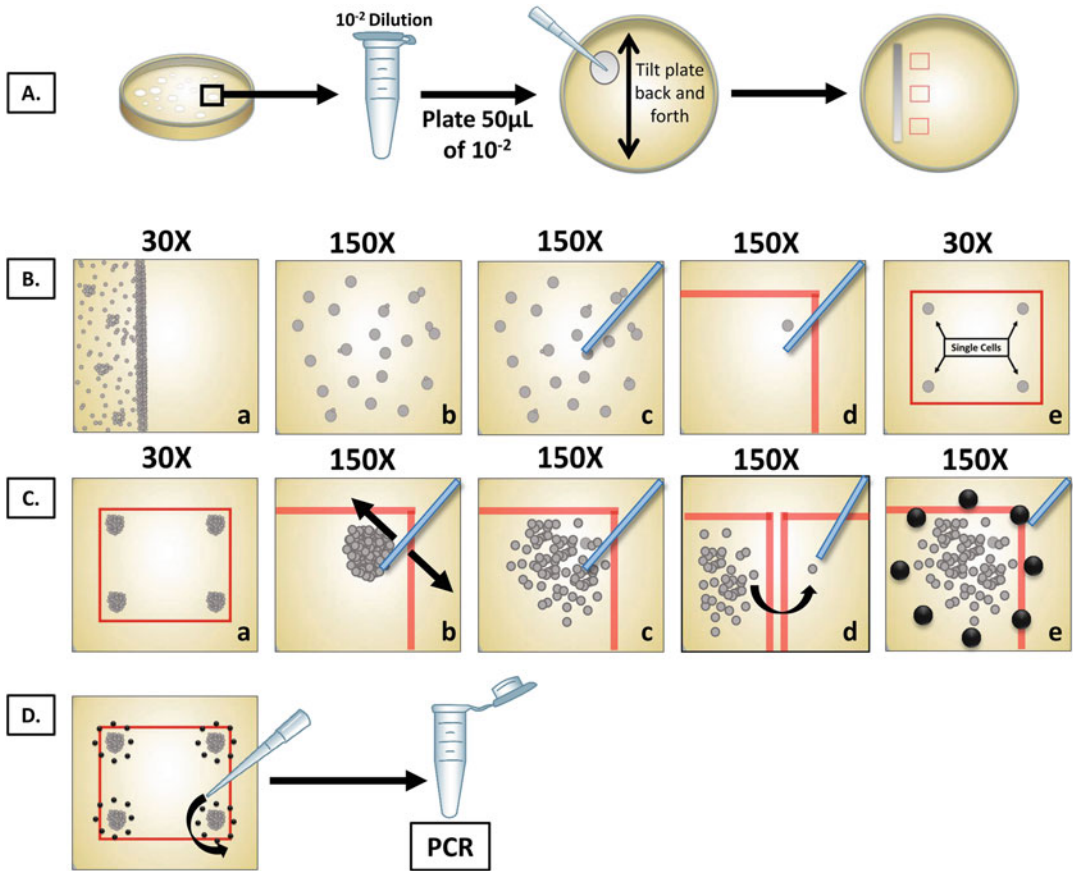
and run samples on a 12% polyacrylamide gel to determine the initial size of the expanded tract (*see* Subheading 3.2, steps 5).

4. After confirming the size of the initial TNR tract expansion, select a 2 mm colony from the selective plate and resuspend it in 200  $\mu$ L of sterile dH<sub>2</sub>O (Fig. 8).
5. Plate 50  $\mu$ L of this resuspension onto the top left side of a fresh SC –histidine. Tilt the plate so that the resuspension drips down toward the bottom of the plate in a straight line. Tilt the plate back and forth so that the cells are evenly distributed along the streak. Allow this streak to dry for approximately 5–7 min (Fig. 8b, 9A).
6. After streak has nearly dried, draw three small boxes vertically to the right of the streak of cells and label them 1, 2, and 3. This is where you will place your single cells for the time course (Fig. 8b, 9C).
7. Using a tetrad dissection microscope, i.e., a microscope with a micromanipulator (*see* Note 9), focus on individual cells under 30 $\times$  magnification (Fig. 9B). Switch to 150 $\times$  magnification and identify single cells within the streak of cells. Choose unbudded cells so that you are starting with a homogeneous/single tract.



**Fig. 8** Schematic of the microcolony time course experiment. (a) Individual colonies with TNR tract expansions were selected on plates containing 5-FOA; the tract increase was confirmed by colony PCR in each case. (b) Each confirmed expanded colony is then resuspended in 200  $\mu\text{L}$  of sterile deionized water and serially diluted down to  $10^{-2}$ . Fifty microliters of this dilution is spotted and smeared onto a nonselective media plate (SC-His). Single cells from the plated smear are then isolated using a micromanipulator and placed into boxes drawn onto the back of the plate. (c) Plates are then placed into a 30  $^{\circ}\text{C}$  incubator and allowed to grow for approximately 15 h, resulting in a microcolony that has undergone  $\sim 8$ – $10$  generations ( $\sim 250$ – $1000$  cells/microcolony). (d) Using the micromanipulator, a single cell is removed from the microcolony to propagate the next time point, while the remainder of the microcolony is swiped off the plate and used in colony PCR to determine tract length (e). Plates are then returned to the 30  $^{\circ}\text{C}$  incubator for  $\sim 15$  h to allow the new microcolony to grow (f), continuing the aforementioned steps through the duration of the time course.

8. Carefully swipe the single cell off the plate with the micromanipulator and place it into one of the boxes you drew vertically next to the streak of cells. Place a single cell in each of the three boxes you have drawn. Make sure not to puncture the plate with the cell and glass rod, this will deposit the cell into the agar where it will not grow well (Fig. 9B).

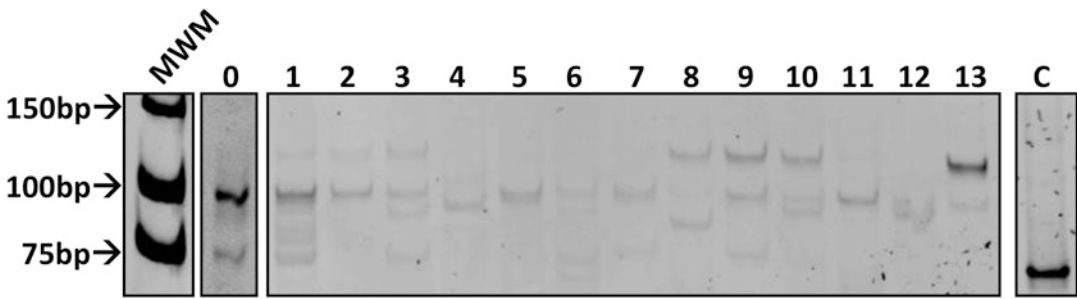


**Fig. 9** (A) After confirming an expanded TNR tract through colony PCR, the colony is swiped off the plate, resuspended in sterile deionized water and serial dilute down to  $10^{-2}$ . Fifty microliter of the  $10^{-2}$  dilution is then plated onto SC –His and the plate tilted back and forth to smear the cells to singles and the plate is brought to the micromanipulator. (Ba) The plate is placed onto the micromanipulator microscope and the smear of cells is observed at  $30\times$  magnification. Once a patch of single cells is identified, the magnification of the microscope is increased to  $150\times$  (Bb), and single, nonbudding cells are removed from the plate using the glass rod of the micromanipulator (Bc) and placed into the corners of a box drawn on the back of each plate (Bd). Four individual cells can be placed in each box (Be). (Ca) Plates are then placed into a  $30\text{ }^{\circ}\text{C}$  incubator for  $\sim 15$  h, resulting in a microcolony with 250–1000 cells in each corner of the box. The micromanipulator is used to smear the microcolony under  $150\times$  magnification (Cb), allowing for an accurate count of cells within each microcolony (Cc). A single cell from the microcolony is selected and placed in a newly drawn box to propagate the next time point (Cd). The perimeter of the remaining microcolony is marked by puncturing the media with the micromanipulator glass rod, allowing for easy identification away from the microscope (Ce). (D) The remainder of each microcolony is swiped off the plate using a  $200\text{ }\mu\text{L}$  pipette tip and resuspended in PCR master mix to assess tract length changes

9. After filling all three boxes with single cells, incubate each plate at  $30\text{ }^{\circ}\text{C}$ , allowing cells to undergo 8–10 rounds of replication, resulting in a microcolony of approximately 250–1000 cells. Determination of the generation time for each strain (*see below*) will reveal how long the cells should be incubated. It

is a good idea to check the plates periodically throughout the day to count cells in the microcolony and confirm the growth rate. Typically, 15–20 h are required for 8–10 generations in our strain background on synthetic media [17].

10. After 8–10 generations, draw three new boxes in a different color marker adjacent to your initial boxes. This new box is where you will transfer four individual cells to propagate your next microcolony.
11. Using the micromanipulator, carefully smear the microcolony with the needle to obtain a better estimate of the cell count. Then select four individual colonies for your next time point and place one individual cell in each of the four corners of the next box. Four cells are necessary to assure you have at least one microcolony that grows to test for your next time point (Fig. 9C).
12. After four cells have been placed in the new box, return to the original microcolony and puncture the agar around the smeared microcolony. These holes surrounding the microcolony will help identify the location of the microcolony on the plate and should be visible without use of the microscope. This will be useful removing the microcolony from the plate for PCR (Fig. 9C, D).
13. Repeat **steps 9–11** for the rest of your time courses, making sure to mark each microcolony that will be swiped off the plate and used for template in PCR.
14. After single cells have been transferred to generate the next microcolony, use a sterile 200  $\mu\text{L}$  pipette tip to harvest each microcolony from the plate and resuspend it in 20  $\mu\text{L}$  of the colony PCR mix (*see* Subheading 3.2, **step 3**). Make sure that all of the PCR mix in the tip is pushed back out into the PCR mix. You can do this by pushing down on the top of the tip with your finger.
15. Make sure that your tubes are well labeled.
16. Using the microscope, check that each microcolony was harvested. There will be a small number of cells that did not make it onto the tip; this is ok. If there appears to be a substantial number of cells still present on the plate, mark the area surrounding the microcolony again and reswipe the plate using a sterile 200  $\mu\text{L}$  pipette tip (*see* **Note 10**).
17. Place plates back into the 30  $^{\circ}\text{C}$  incubator for another 8–10 generations (Fig. 8f).
18. After incubation, once again draw boxes horizontally on the plate, making sure to carefully label the time point. Look at each of the propagated microcolonies and observe their size. Out of the four microcolonies that you have grown, chose the



**Fig. 10** Dynamic changes in TNR tracts starting from a single cell. Expansion events are selected in the wild-type (CTG)<sub>25</sub> background and confirmed by PCR. Individual cells from these colonies are isolated and allowed to undergo 8–10 rounds of replication, resulting in a microcolony approximately 250–1000 cells in size (see Fig. 8). A single cell from the microcolony is then taken to propagate another microcolony for the time course. The remainder was swiped off the plate and used in a PCR reaction to determine tract length. PCR reactions are digested and resolved on a 12% polyacrylamide gel and stained with EtBr; the images have been inverted for ease of viewing. The lanes marked C in each panel indicate the 75-bp tract amplified from the TNR plasmid control. The numbers across the top of the gels indicate the time point. In this microcolony time course, the initial expansion is shown at time point zero, and dynamic changes in the length of the TNR tract are observed as a function of time, with the final tract length larger than the starting tract length. Multiple bands are observed at some time points, likely due to cellular heterogeneity

single microcolony that fits the criteria the best (250–1000 cells in size).

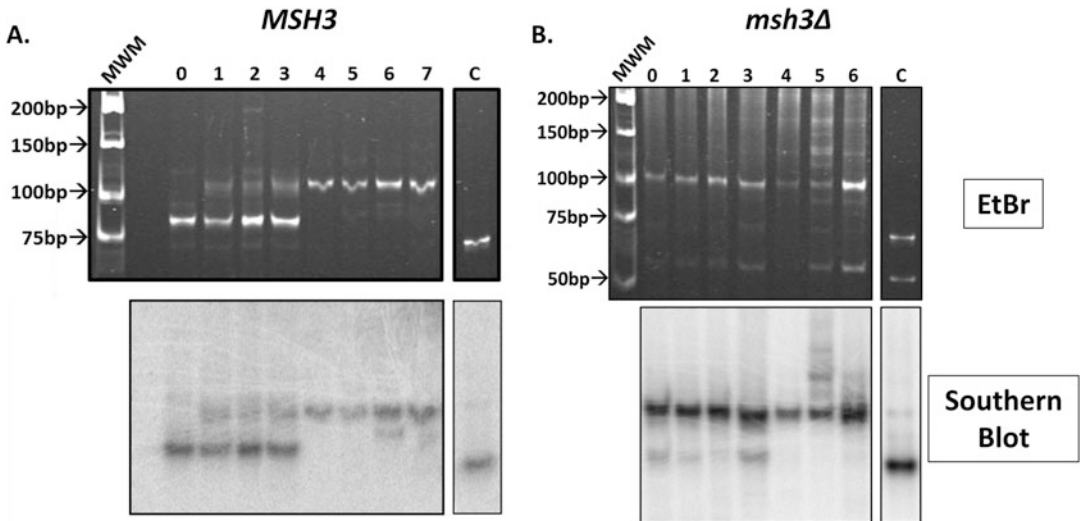
19. Repeat **steps 10–15** for each microcolony for the duration of the time course (Fig. 8b–f).
20. Microcolony PCR reactions for each time point should be digested with *SphI* and *AflII* (see Subheading 3.2, **steps 3** and **4**) to assess dynamic tract behavior (Figs. 8e, 10, 11).

### 3.8 Southern Blotting—Making the Probes

Southern blotting is used to confirm single integration of the TNR reporter construct at the correct chromosomal location. It is also used to confirm that the PCR products obtained following PCR in each of the different time course experiments actually contain a TNR tract.

Two different probes are used.

1. Southern blotting is used to confirm single integration of the tract (a) and to determine tract size following an assay (b).
  - (a) For a probe to determine proper integration of the reporter construct, a plasmid preparation of pBL94 [6, 37] is used. The plasmid is digested with *BamHI* to release a 1.2 kb fragment that contains *lys2* [37].
  - (b) For a tract specific probe to confirm the presence of the TNR tract following PCR of time course samples, a plasmid preparation of pBL169 [17, 36] is used. PCR amplify the TNR tract in 2 or more 50  $\mu$ L PCR reactions using primers SO295 and SO296 (see Subheading 2). Digest



**Fig. 11** Southern blotting using a tract specific probe. Wild-type (CTG)<sub>25</sub> (a) and *msh3Δ* (CTG)<sub>25</sub> (b) colony PCR time courses were performed as described. Digested PCR samples were resolved on a 12% polyacrylamide gel, stained with EtBr, and imaged (a, b: top panels, as in Fig. 7). DNA from the gel was then transferred to a neutral Nytran membrane, crosslinked, and subjected to Southern blot using a tract-specific TNR probe (a, b: bottom panels). Wild-type Southern blot (a, bottom panel) shows progressive expansion of the TNR tract over time. In contrast, the Southern blot of time course performed in the absence of the MMR factor *MSH3* (*msh3Δ*) (b, bottom panel) confirms that the TNR tract is stable in this background. Numbers across the top indicate days of the time course. The lanes marked C in each panel indicate the 75-bp tract amplified from the TNR plasmid control

- each of the PCR reactions with the restriction enzyme *SphI* to isolate the repeat tract.
2. Both probes are electrophoresed and purified prior to use.
    - (a) For the integration probe, gel purify the 1.2 kb fragment after electrophoresis through a 1% agarose gel.
    - (b) For the TNR-specific probe, gel purify the 75 bp PCR product after electrophoresis through a 12% polyacrylamide, 0.5× TBE gel, and staining with ethidium bromide (see Subheading 3.2, step 5).
  3. For both probes, place the excised bands into a clean microcentrifuge tube, making sure to include as little agarose or polyacrylamide as possible. Weigh each of the excised bands on a scale, using an empty microcentrifuge tube as your zero. Use a gel extraction kit (e.g., Omega Bio-Tek E.Z.N.A. Gel Extraction Kit) to remove your fragment from the gel. For the small, tract-specific probe, pool the multiple bands into a single column to increase the concentration of your final extracted DNA fragment. In the final elution step of the gel extraction protocol, use a small volume of sterile dH<sub>2</sub>O (30 μL) to increase the overall DNA content of the extraction. Sterile



dH<sub>2</sub>O is used instead of the provided elution buffer to assure proper labeling using a random primer DNA labeling kit.

4. Radiolabel the purified probes with the Random Primer DNA Labeling kit. For this specific labeling kit, 150 ng of template DNA is required, and it is designed for the use of 1.85 MBq (50  $\mu$ Ci) [ $\alpha$ -<sup>32</sup>P] dCTP (~111 TBq/mmol, 370 MQq/mL).
5. Follow the protocol provided in the Random Primer DNA Labeling Kit. Once the tract specific probe is properly labeled, it can be stored in a designated radioactivity approved freezer at -20 °C until it is ready to be used.

### 3.9 Southern Blotting: Gel and Transfer

1. To assess both tract integration and tract size, DNA is first electrophoresed through a gel.
  - (a) For confirmation of integration, isolate gDNA and digest with *Bam*HI. Electrophorese through a 0.8%, 0.5 $\times$  TBE agarose gel.
  - (b) For time course tract confirmation, digest PCR reactions with *Sph*I and *Aff*II and electrophorese the products through a 12% polyacrylamide, 0.5 $\times$  TBE gel and stain the gel with ethidium bromide (*see* Subheading 3.2, step 5). Take a picture of the gel so that you can orient yourself later.
2. In a clean Pyrex dish, denature the gel at room temperature for 60 min in 500 mL of denaturation buffer. Make sure that the gel is completely immersed in the buffer. The entire gel needs to be at the high pH for the DNA to be denatured uniformly.
3. Neutralize the gel in 500 mL of neutralization buffer at room temperature for 30 min. Remove the buffer after 30 min and repeat the neutralization step. At the end of the two washes, the pH of the gel should be 8 or less. This can be tested by touching a pH strip to the side of the gel.
4. Place two weighted tip racks in the middle of a baking pan (reservoir). Fill reservoir with 1 $\times$  SSC to ½ in. deep, making sure that the buffer does not reach the top of the tip boxes.
5. Cut two sheets of Whatman 3MM paper the width of the gel; these will be used as the wick to bring the 1 $\times$  SSC buffer through your gel, effectively transferring the DNA to a neutral Nytran membrane. Drape these pieces of Whatman paper over the top of the tip boxes, hanging the ends into the 1 $\times$  SSC-filled reservoir. Wet the entire wicks with 1 $\times$  SSC.
6. Upon flooding the surface of the wick with 1 $\times$  SSC, remove any trapped air bubbles by rolling a Pasteur pipette across surface. Bubbles in the transfer setup could affect the efficient transfer of DNA to the membrane. Before putting your gel onto the transfer apparatus, make sure to notch one corner of

the gel to identify the orientation of the gel when transferred. A corner should also be notched on the neutral Nytran membrane. Carefully invert and lay gel on top of the wick. Remove any air bubbles between gel and the wick. It is helpful to continuously wet the gel with  $1\times$  SSC. Invert the gel because the DNA is often more concentrated near the bottom.

7. Place strips of cut plastic (sheet protectors) or rolled plastic wrap on all four sides of the gel to prevent lateral diffusion of your DNA.
8. Cut a piece of neutral Nytran membrane to fit the size of the gel and label the back of the membrane that is not in contact with the gel. Wet the membrane briefly in deionized water and then allow it to soak in  $1\times$  SSC for 5 min.
9. Flood the gel surface once again with  $1\times$  SSC and carefully lay the neutral Nytran membrane on top, making sure to match up the notch of the gel with the notch of the membrane, again being careful to remove any air bubbles that may form.
10. Cut four pieces of Whatman paper to the exact size of the gel. Soak these pieces for 5 min in  $1\times$  SSC, before laying them on top of the gel. Once again, make sure that all bubbles are removed by rolling a Pasteur pipette over top of the transfer set up.
11. Place a stack of paper towels (4" tall) cut slightly larger to cover the entirety of the gel transfer area. Place a weight on top of the paper towels. In our transfer system, we use two large aluminum blocks to evenly distribute weight across the gel. Wrap the entire reservoir and transfer set-up with plastic wrap to help prevent evaporation of the  $1\times$  SSC buffer. Allow the transfer to proceed overnight (~15–20 h) [37].
12. After the overnight transfer, unwrap the reservoir and remove the gel. Image the gel to make sure that all of the DNA has transferred from the gel to the Nytran membrane.
13. Dry the membrane between to pieces of filter paper under the fumehood.
14. Once the membrane is sufficiently dry, wrap the membrane in plastic wrap.
15. Crosslink the wrapped blot at 1600  $\mu$ J, using a Stratalinker or equivalent UV crosslinker. The DNA should be stably maintained on the neutral Nytran membrane and can be stored at  $-4^\circ\text{C}$  until ready to hybridize with a tract specific probe.

**3.10 Southern Blotting: Hybridization Protocol Using  $\alpha^{32}\text{P}$ -dCTP Labeled Tract-Specific Probe**

1. Following the transfer, remove the radiolabeled probe from the radioactive materials freezer and let it thaw on ice behind a Plexiglass shield.
2. Heat the thawed probe to  $95^\circ\text{C}$  for 5 min and then place on ice for 5 min.



3. Add 5  $\mu\text{L}$  of radiolabeled probe [ $\approx 0.008$  mCi] to 10 mL of Hybridization Buffer in a 250 mL screwcap bottle that contains the neutral Nytran membrane blot that is to be probed.
4. Screw the bottle cap on tight and wrap the top several times with Parafilm to prevent any leaking.
5. Add the sealed Southern Blotting tube into a rotating HYBAID hybridization oven at 65 °C and allow it to incubate overnight (~15–20 h).
6. The following day, remove the screwcap bottle from the HYBAID oven. Open up the screwcap bottle and dump the Hybridization Buffer and unbound radiolabeled probe into an appropriate radioactive liquid waste container, making sure the blot itself stays within the screwcap bottle.
7. Add 250 mL of Wash Buffer to the bottle to wash blot. Screw the cap on tight and once again wrap the top of the bottle several times with Parafilm to prevent leaking. Put the sealed tube back into the HYBAID oven at room temperature for 20 min.
8. After 20 min, remove the bottle from the HYBAID oven and empty the wash buffer into an approved radioactive liquid waste container.
9. Repeat **steps 6–8** for a second wash. Depending on how fresh the probe is, it may be necessary to allow for longer washes to improve the quality of the final Southern Bolt image. Upon completion of the second wash, survey the Hybaid oven with the Geiger counter to assure there is no radioactivity present in or around the oven.
10. Remove the washed blot from the screwcap bottle and allow it to air-dry on a piece of Whatman paper on the radioactivity bench behind plexiglass.
11. After the probed blot has dried, wrap the blot with plastic wrap and expose it to a Phosphorimager screen (Molecular Dynamics). Make sure to note the orientation of the blot on the screen so that it can be correctly interpreted later on. Radioactive label that has undergone less than one half-life can be exposed for 4–8 h and will show clear resolution when imaged using a STORM PhosphorImager (Fig. 11). After the blot is imaged and the results obtained, the blot should be disposed in an approved dry radioactive waste container.

### **3.11 Generation Time in Liquid Culture**

To calculate rates of change, and to normalize and compare individual time courses, it is necessary to determine the number of generations a culture or colony has undergone during a particular time period. Generation times, or doubling times, can be determined in liquid media or upon growth on agar plates.

1. Streak out the TNR-containing strain of interest on SC –histidine plates to generate single colonies. SC –histidine ensures that the colonies maintain the TNR reporter system. Place plates in the 30 °C incubator and allow them to grow for ~3 days, until colonies are ~2 mm in diameter.
2. Select ~2 mm colony using a sterile wooden stick and resuspend the colony in 10 mL of liquid SC –histidine in a 125 mL sterile flask. Place the flask in a 30 °C shaker overnight to allow saturation of the culture.
3. After ~16 h, remove the flask from the shaker and measure the density of the culture. Dilute the saturated culture 1:10 (100 µL of culture +900 µL of media) in SC –histidine in a microcentrifuge tube (*see Note 11*). Using a spectrophotometer, measure and record the optical density at 600 nm (OD<sub>600</sub>) of the dilution. Multiply this number by 10 (to account for the dilution) to determine the OD<sub>600</sub> of the saturated culture.
4. Dilute the saturated culture to an OD<sub>600</sub> of 0.1 with liquid SC –histidine in fresh 250 mL flask. The final culture volume should be 50 mL.
5. From the newly diluted culture, remove 1 mL using a pipetman and measure the OD<sub>600</sub>. This will be used as the time point zero.
6. Place the diluted culture in the 30 °C shaker. After 45 min, remove the flask from the shaker and remove 1 mL to measure the OD<sub>600</sub>. Optical density should be measured in this manner every 45 min for the duration of 10 h.
7. At the end of the 10 h period, allow cultures to saturate overnight in the 30 °C shaker.
8. The next morning, **steps 1–10** can be repeated for a second set of time points. At the end of the 2-day period, optical density measurements can be plotted on a graph as a function of time to extrapolate a generation (or doubling) time.

### **3.12 Generation Time on Solid Media**

1. Streak out the strain of interest to single colonies on SC –histidine plates and allow the colonies to grow to ~2 mm in a 30 °C incubator. Select a 2 mm colony from the plate using a sterile flat end toothpick and resuspend it in 200 µL of sterile deionized water.
2. Plate 50 µL of this resuspension onto the top left side of a fresh synthetic medium plate lacking histidine. Tilt the plate so that the resuspension drips down toward the bottom of the plate in a straight line. Tilt the plate back and forth so that the cells are evenly distributed along the streak. Allow this streak to dry for approximately 5–7 min.

3. After the streak of cells has nearly dried, draw a single box to the right of the streak of cells. This is where you will be putting your single cells for the growth curve.
4. Using a microscope equipped with a micromanipulator (e.g., tetrad dissection microscope), invert the plate holder so that the streak is on your right hand side.
5. First, focus on the line of cells under  $30\times$  magnification. Switch to  $150\times$  magnification and identify single cells within the streak of cells. Select cells that are unbudded to avoid skewing the apparent generation time.
6. Using the micromanipulator, carefully move a single cell from the plate to one of the boxes you drew vertically next to the streak of cells. Place four individual cells, one per corner, into each box. Make sure not to puncture the plate with the cell; this will affect the growth rate.
7. To ensure representative growth rates, repeat **steps 2–6** using several different colonies from the same strain. This ensures that the final growth curve results represent cells from several independent colonies from the same strain background.
8. After filling all of the boxes with single cells, incubate each plate at  $30\text{ }^{\circ}\text{C}$ .
9. Regularly check the individual cells to observe budding of daughter cells. Budding cells should be probed with the micromanipulator to determine whether the cells have completely divided. When division is complete, the daughter cell is easily removed from the mother cell with the needle.
10. As soon as the daughter cell is completely separated from the mother, remove the daughter cell and place it just below the mother cell on the plate. The single cells selected for the growth curve experiment are unlikely to be at precisely the same stage of the cell cycle. Therefore the time required to complete the first cell division should not be included in the growth curve. Once the first daughter cell is separated from the mother, the time for the growth curve begins (*see Note 12*).
11. Every 30 min, for 10 h, remove the plates from the incubator and observe the cells morphology, looking for daughter cell budding. Note the amount of time it took for the daughter cell to completely separate from the mother and restart the timer. In order to record doubling times accurately, it is best to follow fewer cells at a time, and repeat the experiment several times with different starting colonies.
12. At the end of the 10 h period, median and mean doubling times can be calculated.

---

## 4 Notes

1. We use lab-purified *Thermus aquaticus* (*Taq*) polymerase. Briefly, *Taq* is overexpressed in *Escherichia coli* (BL21), harvested by centrifugation and the pellet is frozen in liquid nitrogen. The pellet is thawed and resuspended in Buffer A (50 mM HEPES pH 8.0, 50 mM glucose, 1 mM EDTA) containing lysozyme (Worthington) (4 mg/mL) and then incubated at room temperature for at least 15 min, until the cells are lysed and appear gooey. Buffer B (10 mM HEPES pH 8.0, 50 mM KCl, 1 mM EDTA, 0.5% Tween 20, and 0.5% NP-40) is added and the lysate incubated in a 75 °C water bath for 45 min, with periodic inversion. The lysate was centrifuged at  $20,000 \times g$  and the supernatant is transferred to a new tube. The supernatant is treated with DNaseI (Worthington), in the presence of MgCl<sub>2</sub> and MnCl<sub>2</sub>, at 75 °C for 45 min, centrifuged and the supernatant kept. The supernatant is dialyzed again three changes of dialysis buffer (20 mM Tris pH 7.5, 200 mM KCl, and 1 mM EDTA), then diluted 2× with storage buffer (50 mM Tris pH 7.5, 50 mM KCl, 1 mM DTT, 0.5% Tween 20, and 0.5% NP-40) containing 50% glycerol and then 2× again with storage buffer containing 75% glycerol. The final protein is aliquoted and stored at -80 °C.
2. This elegant system was developed, verified, and characterized by the Lahue lab [35, 36, 45]. It is beautifully and comprehensively described in a previous Methods in Molecular Biology publication [37].
3. Pour 8–10 spreading glass beads on each plate. Shake/rotate the plate, keeping it on the benchtop surface to move the beads around to cover the plate. Sometimes it is helpful to add 50 µL of sterile dH<sub>2</sub>O to each plate before adding the glass beads as it helps with the spreading process. A stack of ~5 plates can be plated at the same time.
4. Mutations in the *URA3* locus will also result in resistance to 5-FOA [37]. It is therefore necessary to check for false positives by performing colony PCR, digesting the DNA and electrophoresing the DNA to image the gel, as described above.
5. The following approach can be used to determine TNR tract lengths:
  - (a) After running digested PCR reaction on a 12% polyacrylamide gel, stain and image gel (*see* Subheading 3.2, **step 5**). The digested TNR sequences should be clearly visible. For the quantification of tract length, it is helpful to image your gel as straight as possible, making sure to have a level bottom edge.
  - (b) From a predetermined point on the gel (i.e., the bottom of each well), measure the distance, in millimeters (mm),

between the bands of known size in the low molecular weight marker used on the gel. These measurements (y-axis) are plotted versus the known size of each band (x-axis) to generate a standard curve from which tract lengths can be interpolated.

- (c) Using the same starting point of reference that was used in **step 2**, measure the known TNR tract band in each lane of the gel. Use the standard curve graph to determine the tract length in base pairs of each fragment. Divide this number by three to report the number of repeats within the tract.
6. Synthetic media lacking histidine is used throughout all the time courses to ensure that the TNR reporter construct is maintained [37].
7. For acid-washed beads, we purchase unwashed Sigma 425–600  $\mu\text{m}$  beads (G9268) and acid-wash in the laboratory. Pour the beads into a 1 L beaker containing 5 M nitric acid in the fumehood. Stir occasionally for 1–2 days. Wash with  $\sim 5$  volumes of  $\text{dH}_2\text{O}$  and then with 5 volumes of  $18 \Omega \text{ dH}_2\text{O}$ . Remove as much liquid as possible and bake at  $450^\circ\text{F}$  for  $\sim 2$  h. Store the beads in a  $-20^\circ\text{C}$  freezer.
8. It is worth taking the extra time to do a few colony PCR reactions per colony per time point in case one or more colony PCR reactions are not successful due to the use of too many cells or too little cells in solution.
9. A 25  $\mu\text{m}$  or 50  $\mu\text{m}$  glass needle (Cora Styles Lab Supplies) is attached to the dissection arm using putty. This is what will be used to move cells around on the plate.
10. Removing marked microcolonies from plates with pipette tips is easiest and most efficient when the pipette tip is held and swiped across the plate at a low angle, i.e., close to parallel to the plate.
11. A 1:10 dilution is necessary to get an accurate reading of your saturated culture.
12. Although tedious, it is best to only have 4–5 single cells per plate. Observing too many cells on a single plate, each with potentially different starting times, may result in the plate being out of the incubator for too long. Ideally, the plate should be incubating at  $30^\circ\text{C}$  for 30 min at a time.

---

## Acknowledgements

We are particularly grateful for the generosity of Dr. Robert Lahue in providing reagents and for many useful discussions and comments in developing this protocol. We are also grateful for discussions with Dr. Catherine Freudenreich, Dr. Robert Bambara, Dr.

Lata Balakrishnan and Dr. Eric Alani. We also thank Alyssa Crick Williams for her assistance with the artwork. Work in the Surtees lab is supported by the National Institutes of Health (GM087459 to J.A.S.) and the American Cancer Society (RSG-14-235-01 to J.A.S.). J.A.S. is an ACS Research Scholar.

## References

1. Arana ME, Kunkel TA (2010) Mutator phenotypes due to DNA replication infidelity. *Semin Cancer Biol* 20:304–311
2. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5:435–445
3. Umar A, Kunkel TA (1996) DNA-replication fidelity, mismatch repair and genome instability in Cancer cells. *Eur J Biochem* 238:297–307
4. Strand M, Prolla TA, Liskay RM, Petes TD (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365:274–276
5. Gordenin DA, Kunkel TA, Resnick MA (1997) Repeat expansion [mdash] all in flap? *Nat Genet* 16:116–118
6. Kantartzis A, Williams GM, Balakrishnan L, Roberts RL, Surtees JA, Bambara RA (2012) Msh2-Msh3 interferes with Okazaki fragment processing to promote trinucleotide repeat expansions. *Cell Rep* 2:216–222
7. Sia E, Kokoska R, Dominska M, Greenwell P, Petes T (1997) Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol Cell Biol* 17:2851–2858
8. Romanova NV, Crouse GF (2013) Different roles of eukaryotic MutS and MutL complexes in repair of small insertion and deletion loops in yeast. *PLoS Genet* 9:e1003920
9. Kunkel TA, Erie DA (2015) Eukaryotic mismatch repair in relation to DNA replication. *Annu Rev Genet* 49:291–313
10. Crouse GF (2016) Non-canonical actions of mismatch repair. *DNA Repair* 38:102–109
11. Owen BAL, Yang Z, Lai M, Gajec M, Badger J d, Hayes JJ, Edelman W, Kucherlapati R, Wilson TM, McMurray CT (2005) (CAG)<sub>n</sub>-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat Struct Mol Biol* 12:663–670
12. Manley K, Pugh J, Messer A (1999) Instability of the CAG repeat in immortalized fibroblast cell cultures from Huntington's disease transgenic mice. *Brain Res* 835:74–79
13. van den Broek WJAA, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJTA, Wieringa B (2002) Somatic expansion behaviour of the (CTG)<sub>n</sub> repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum Mol Genet* 11:191–198
14. Foiry L, Dong L, Savouet C, Hubert L, Te Riele H, Junien C, Gourdon G (2006) Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DMI transgenic mice. *Hum Genet* 119:520–526
15. Gannon A-MM, Frizzell A, Healy E, Lahue RS (2012) MutS $\beta$  and histone deacetylase complexes promote expansions of trinucleotide repeats in human cells. *Nucleic Acids Res* 40:10324–10333
16. Halabi A, Ditch S, Wang J, Grabczyk E (2012) DNA mismatch repair complex MutS $\beta$  promotes GAA-TTC repeat expansion in human cells. *J Biol Chem* 287:29958–29967
17. Williams GM, Surtees JA (2015) MSH3 promotes dynamic behavior of Trinucleotide repeat tracts in vivo. *Genetics* 200:737–754
18. Schmidt MHM, Pearson CE (2016) Disease-associated repeat instability and mismatch repair. *DNA Repair* 38:117–126
19. Mirkin SM (2007) Expandable DNA repeats and human disease. *Nature* 447:932–940
20. Kovtun IV, McMurray CT (2008) Features of trinucleotide repeat instability in vivo. *Cell Res* 18:198–213
21. Usdin K, House NCM, Freudenreich CH (2015) Repeat instability during DNA repair: insights from model systems. *Crit Rev Biochem Mol Biol* 50(2):1–26
22. Paulson HL, Fischbeck KH (1996) Trinucleotide repeats in neurogenetic disorders. *Annu Rev Neurosci* 19:79–107
23. Kay C, Collins JA, Miedzybrodzka Z, Madore SJ, Gordon ES, Gerry N, Davidson M, Slama RA, Hayden MR (2016) Huntington disease reduced penetrance alleles occur at high frequency in the general population. *Neurology* 87:282–288
24. Langbehn DR, Hayden MR, Paulsen JS (2010) CAG-repeat length and the age of onset in

- Huntington disease (HD): a review and validation study of statistical approaches. *Am J Med Genet B Neuropsychiatr Genet* 153B:397–408
25. Concannon C, Lahue RS (2014) Nucleotide excision repair and the 26S proteasome function together to promote trinucleotide repeat expansions. *DNA Repair* 13:42–49
  26. Zhang L, LeeFlang EP, Yu J, Arnheim N (1994) Studying human mutations by sperm typing: instability of CAG trinucleotide repeats in the human androgen receptor gene. *Nat Genet* 7:531–535
  27. LeeFlang EP, Zhang L, Tavaré S, Hubert R, Srinidhi J, MacDonald ME, Myers RH, de Young M, Wexler NS, Gusella JF, Arnheim N (1995) Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. *Hum Mol Genet* 4:1519–1526
  28. LeeFlang EP, Tavaré S, Marjoram P, Neal COS, Srinidhi J, MacFarlane H, MacDonald ME, Gusella JF, de Young M, Wexler NS, Arnheim N (1999) Analysis of Germline mutation spectra at the Huntington's disease locus supports a mitotic mutation mechanism. *Hum Mol Genet* 8:173–183
  29. Martorell L, Gámez J, Cayuela ML, Gould FK, McAbney JP, Ashizawa T, Monckton DG, Baiget M (2004) Germline mutational dynamics in myotonic dystrophy type 1 males: allele length and age effects. *Neurology* 62:269–274
  30. Castel AL, Cleary JD, Pearson CE (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol* 11:165–170
  31. Higham CF, Morales F, Cobbold CA, Haydon DT, Monckton DG (2012) High levels of somatic DNA diversity at the myotonic dystrophy type 1 locus are driven by ultra-frequent expansion and contraction mutations. *Hum Mol Genet* 21:2450–2463
  32. Morales F, Couto JM, Higham CF, Hogg G, Cuenca P, Braida C, Wilson RH, Adam B, del Valle G, Brian R, Sittenfeld M, Ashizawa T, Wilcox A, Wilcox DE, Monckton DG (2012) Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum Mol Genet* 21:3558–3567
  33. Higham CF, Monckton DG (2013) Modelling and inference reveal nonlinear length-dependent suppression of somatic instability for small disease associated alleles in myotonic dystrophy type 1 and Huntington disease. *J R Soc Interface* 10(88):20130605
  34. Du J, Campau E, Soragni E, Jespersen C, Gottesfeld JM (2013) Length-dependent CTG-CAG triplet-repeat expansion in myotonic dystrophy patient-derived induced pluripotent stem cells. *Hum Mol Genet* 22:5276–5287
  35. Miret JJ, Pessoa-Brandão L, Lahue RS (1997) Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol Cell Biol* 17:3382–3387
  36. Miret JJ, Pessoa-Brandão L, Lahue RS (1998) Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* 95:12438–12443
  37. Dixon M, Bhattacharyya S, Lahue R (2004) Genetic assays for triplet repeat instability in yeast. In: Kohwi Y (ed) *Trinucleotide repeat protocols*, vol 277. Humana Press, New York, pp 29–45
  38. Alani E, Sokolsky T, Studamire B, Miret JJ, Lahue RS (1997) Genetic and biochemical analysis of Msh2p-Msh6p: role of ATP hydrolysis and Msh2p-Msh6p subunit interactions in mismatch base pair recognition. *Mol Cell Biol* 17:2436–2447
  39. Gietz D, Jean AS, Woods RA, Schiestl RH (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* 20:1425
  40. Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci U S A* 88:7160–7164
  41. Dixon W, Massey F (1969) *Introduction to statistical analysis*. McGraw Hill, New York
  42. Nair KR (1940) Table of confidence interval for the median in samples from any continuous population. *SankhyA: Indian J Stat* (1933-1960) 4:551–558
  43. Foster PL, Judith LC, Paul M (2006) Methods for determining spontaneous mutation rates. In: *Methods in enzymology*, vol 409. Academic Press, New York, pp 195–213
  44. Zar JH (1999) *Biostatistical analysis*, 4th edn. Prentice Hall, Upper Saddle River, NJ
  45. Rolfmeier ML, Dixon MJ, Pessoa-Brandão L, Pelletier R, Miret JJ, Lahue RS (2001) Cis-elements governing trinucleotide repeat instability in *Saccharomyces cerevisiae*. *Genetics* 157:1569–1579

# Chapter 31

## The Detection and Analysis of Chromosome Fragile Sites

Victoria A. Bjerregaard, Özgün Özer, Ian D. Hickson, and Ying Liu

### Abstract

A fragile site is a chromosomal locus that is prone to form a gap or constriction visible within a condensed metaphase chromosome, particularly following exposure of cells to DNA replication stress. Based on their frequency, fragile sites are classified as either common (CFSs; present in all individuals) or rare (RFSs; present in only a few individuals). Interest in fragile sites has remained high since their discovery in 1965, because of their association with human disease. CFSs are recognized as drivers of oncogene activation and genome instability in cancer cells, while some RFSs are associated with neurodegenerative diseases. This review summarizes our current understanding of the nature and causes of fragile site “expression”, including the recently characterized phenomenon of telomere fragility. In particular, we focus on a description of the methodologies and technologies for detection and analysis of chromosome fragile sites.

**Key words** Common fragile site, Rare fragile site, Telomere fragility, Mitosis, DNA replication stress, Karyotype, Fluorescence in situ hybridization (FISH), Ultrafine DNA bridge (UFB), Micronuclei, 53BP1 body

---

### 1 Introduction

The faithful replication of DNA is fundamental to all organisms as it ensures that the genetic material is passed on to the next generation of cells. In human cells, the genetic information is stored in 23 pairs of chromosomes (22 pairs of autosomes and one pair of sex chromosomes). The replication and segregation of these 46 chromosomes is orchestrated by the cell division cycle. This cycle comprises interphase, during which the chromosomes are replicated, and mitosis when the chromosomes are segregated. During mitosis, the duplicated chromosomes are first condensed (compacted) and then evenly distributed to the two daughter cells. Mitosis is subdivided into six subphases: prophase, prometaphase, metaphase, anaphase, telophase, and cytokinesis. In metaphase, which is when chromosomes are fully condensed and adopt the iconic X-shaped structure, it is well documented that some chromosomal loci tend to form a gap, break or constriction that is not stained by the most commonly used DNA dye, DAPI. These gaps/breaks are generally



referred to as chromosomal fragile sites, and are particularly prevalent in cells exposed to agents that partially inhibit DNA synthesis [1]. Based on their frequency in population, these loci are classified as either common fragile sites (CFSs), which are present in all individuals, or rare fragile sites (RFSs), which exist in less than 5% of individuals [2, 3]. To date, more than 200 CFSs [4], and up to 33 RFS [5] have been identified in human cells.

Interest in chromosomal fragile sites has remained high since their discovery over 50 years ago [6]. This is because CFSs are recognized as a driver of genome instability in cancer cells [7], while some RFSs are associated with neurodegenerative diseases, such as the Fragile X syndrome [8–11]. One model for the appearance (usually termed “expression”) of fragile sites is that they result from a failure in chromatin condensation during early mitosis due to the incomplete replication of the locus [5, 12]. This failure to fully condense the locus, and hence to make it insufficiently compact to be visualized using DAPI staining, helps to explain why the apparently broken fragile site is never detached fully from the host chromosome. Instead, the site appears to be loosely connected, perhaps by undetectable DNA “tethers”; see below. It is also relevant that all of the known cell culture conditions that promote expression of fragile sites perturb DNA replication.

Great progress has been made in recent years toward understanding why and how fragile sites arise. This progress can be summarized in the following areas: (1) cataloging of the genomic features common to different fragile sites (i.e., CG/AT content, repeats features, location on the chromosomes, or evolutionary conservation locus) [13, 14]; (2) defining the cell culture conditions that induce fragile site expression [15–18]; (3) examining the replication program within fragile sites [19]; (4) analyzing how gene transcription can influence the replication of fragile sites [20, 21]; (5) identifying the proteins that are recruited to fragile sites in either S phase or mitosis [22–25]; and, more recently, (6) investigating how fragile site replication is completed, which is now known to occur in early mitosis [26]. This review discusses our current understanding of the nature and causes of chromosome fragility, with a particular focus on the methodologies used for the detection and analysis of these unstable chromosomal loci.

---

## 2 The Induction and Detection of Fragile Sites

The first chromosomal fragile site to be defined in human cells was described by Dekaban et al. in 1965 through an analysis of the karyotypes of blood lymphocytes [6]. In the following years, cytogeneticists remained puzzled by the fact that these fragile sites were observed in some laboratories but not in others. It was not until 1977 that it was shown that the type of medium in which the blood

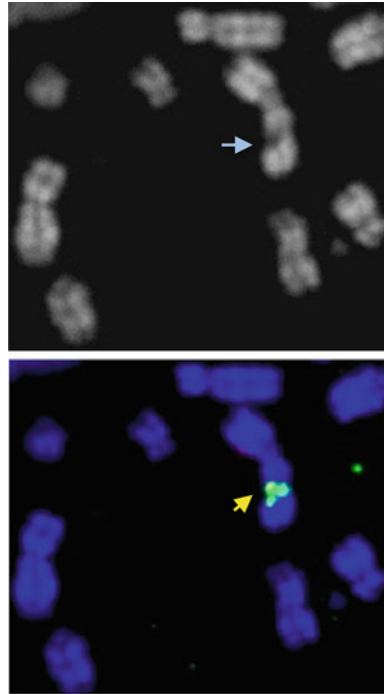
lymphocytes were cultured played a role in the induction of fragile site expression [27]. By 1984, 17 heritable fragile sites had been identified, amongst which 14 were known to be induced by cell culture medium deficient in folic acid [28], or by the so-called “thymidylate stress” (e.g., by exposure to 5-fluorodeoxyuridine, FUdR, a potent inhibitor of thymidylate synthetase) [29]. At that time, it was noticed that thymidylate stress often produces fragile site “hot spots” that occur nonrandomly in the genome [30]. It was subsequently discovered that aphidicolin (APH), an inhibitor of replicative DNA polymerases (but not polymerases  $\beta$  or  $\gamma$ ), could also induce fragility at several nonrandom sites [15]. The term “common fragile site” was then introduced to describe the sites that can be induced by both APH and thymidylate stress, while “rare fragile sites” were defined as those induced by thymidylate stress, but not APH [15]. However, soon after, it was discovered that some CFSs could also be induced by FUdR, although at a much lower frequency [16]. It was therefore suggested that the fragile site classification should be based on their frequency in the population, rather than on the form of induction [16]. To date, it is known that many agents can induce the expression of CFSs and/or RFSs. These include APH, BrdU, 5-azacytidine, caffeine, and FUdR for CFSs [1, 16], and FUdR, thymidine, distamycin A, BrdU, APH, and 5-azacytidine for RFSs [18].

Conventionally, fragile sites are detected in blood lymphocytes that are cultured for approximately 72 h in medium supplemented with 2% phytohemagglutinin [4]. Various fragile-site “inducers” can then be added to the cell culture medium according to the type of fragile site of interest. For example, to induce the expression of CFSs in lymphocytes, the cells are generally incubated with APH (0.2  $\mu$ M) for 24 h, and then arrested for 2 h with the mitotic spindle inhibitor Colcemid (0.04  $\mu$ g/ml) to trap the cells in metaphase when the chromosomes are fully condensed [4]. To observe the chromosomes under a microscope, the cells are swollen with hypotonic buffer, fixed with methanol–acetic acid (3:1), dropped onto slides, and stained by a DNA dye (usually DAPI). For more detailed analysis of chromosome structure, such as gaps or constrictions, G-banding (digestion of chromosomes with trypsin followed by Giemsa staining that yields a series of lightly and darkly stained bands) or Q-banding (a fluorescence based technique using quinacrine staining) are commonly performed [31].

With the increase in interest in fragile site research, a number of variations have been introduced to the above standard procedure. First, CFSs are now widely studied in human fibroblasts and various cancer cells [24–26, 32], as well as in mouse or rat cells [33–36]. As a result of this change, the conditions and procedures for cell culture, fragile site induction, and cell harvesting may vary from study to study. For example, to harvest mitotic cells from adherent cell lines, a method called “mitotic shake-off” is frequently used to

enrich for cells in prometaphase [23]. More recently, because of the need to define which proteins bind to and process CFSs, immunofluorescence (IF) analysis, EdU incorporation, or Fluorescence in situ hybridization (FISH), or a combination of these, are widely performed on metaphase chromosomes [23–25]. Such developments require that the chromosome preparation step be modified such that the detection of proteins bound to a fragile site is still feasible. For example, in the conventional cytogenetic protocol, the mitotic cells are fixed with methanol–acetic acid (3:1). This can interfere with IF analysis, which requires that protein conformation be preserved as much as possible for antibody binding. To overcome this obstacle, a commonly used method is to centrifuge mitotic cells onto slides (instead of dropping them) using a “cytospin” apparatus, and then to fix the metaphases with paraformaldehyde for analysis by IF and then by FISH [37]. It is important to bear in mind also that the protocol for this type of analysis can vary according to the antibodies and FISH probes used [22–24]. Therefore, each step of this protocol needs to be optimized. Figures 1 and 2 show examples of metaphase chromosomes analyzed by IF and an EdU incorporation assay (to mark sites of active DNA synthesis), or by IF combined with FISH to define the location of the FRA16D CFS locus in human cells.

The detection of fragile sites using FISH has enabled researchers to define the fragility of specific loci in greater detail, particularly in cases where fragility might be difficult to detect using simple DNA staining (Fig. 3). In addition, this method has confirmed the identity of a new class of fragile sites: telomere fragile sites. Telomeres are the sequences located at the ends of eukaryotic chromosomes. Telomeres in human cells contain a repeated TTAGGG sequence that ranges in length from 4 to 14 kb [38]. Interest in telomeres has surged since the discovery of the telomere-specific reverse transcriptase, telomerase [39], and, the fact that the maintenance of the telomeric repeat sequence, either by telomerase [40] or by an alternative mechanism (ALT) [41], is essential for the long-term proliferation of cells in culture and in vivo. In the past decade, emerging evidence has shown that telomeric repeats pose a challenge to the DNA replication process, which can generate the so-called “telomere fragile site” (TFS) instability that is similar to what is seen at APH-induced CFSs. Indeed, pronounced TFS expression is seen in cancer cells or immortalized cells, especially when the cells are treated with APH [42] (Fig. 4). The widespread usage of a peptide nucleic acid (PNA) FISH probe that binds specifically to the telomere repeat sequence has been instrumental in the analysis of this type of fragile site, as conventional chromosome staining methods cannot detect minor changes at the tip of chromosomes [43]. Equally importantly, the usage of telomeric PNA probes is compatible with IF analysis [42, 44–47].

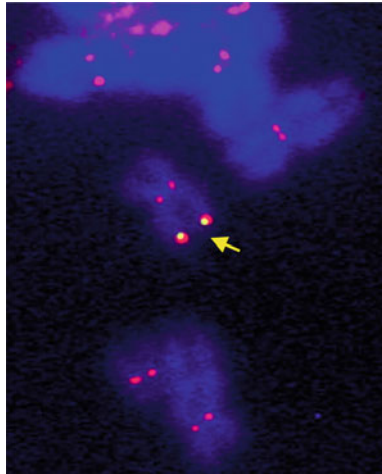


**Fig. 1** A representative example of the colocalization of FANCD2 and EdU with a common fragile site. The U2OS cells were treated with APH (0.4  $\mu$ M), synchronized using R03306 and then pulse-labeled with EdU as described previously [26]. Metaphase cells were collected by mitotic shake-off, and then spun onto coverslips for metaphase chromosome preparation. The metaphase chromosomes were stained with FANCD2 antibody (*green*), EdU (click IT reaction) (*red*), and observed using DAPI (*blue*). In the *upper panel*, the *blue arrow* indicates a chromosome fragile site. In the *lower panel*, the *yellow arrow* indicates the colocalization of FANCD2 and EdU at the fragile site, as revealed by the yellow signal in this merged image

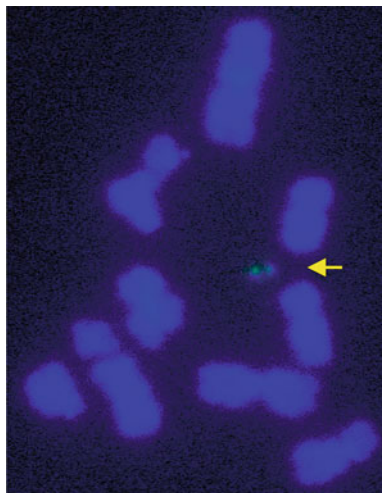
---

### 3 The Characterization of Fragile Sites

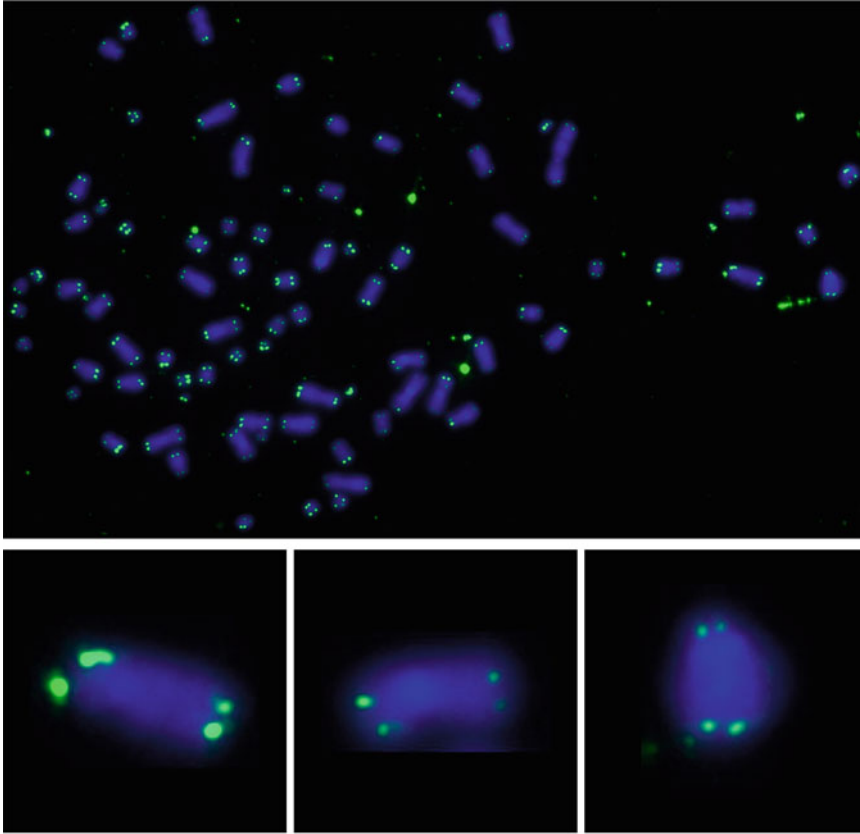
CFSs cover megabase-pair-long regions and tend to contain or overlap with large genes. In addition, they are located in evolutionally conserved chromosomal regions, such as in syntenic regions of mouse and human chromosomes [1]. Several hypotheses have been put forward to explain how completion of CFS replication is delayed compared to that in other regions of the genome. For example, evidence has been provided that CFS expression can be caused by either a lack of replication origins within the locus [19], or by collisions between replication forks and the long gene transcripts located there [48]. Furthermore, the molecular events leading to the fragile site expression are beginning to be revealed, following the findings that key components of DNA repair and



**Fig. 2** A representative example of the analysis of colocalization of FANCD2 with a specific common fragile site: FRA16D. GM06865 lymphocytes were cultured and treated with APH as described previously [4]. Metaphase spreads were prepared by the “Cytospin” method. The metaphase chromosomes were subjected to IF analysis with FANCD2 antibody (*red*) followed by FISH analysis using a BAC probe that binds specifically to the FRA16D locus (*green*). DNA was observed with DAPI (*blue*). The *yellow arrow* indicates FANCD2 “twin foci” colocalized with FRA16D FISH probes where the merged signal is *yellow*



**Fig. 3** A representative example of the analysis of a rare fragile site, FRAXA, using FISH. GM09237 lymphocytes, which contain over 900 CGG repeats in the promoter region of *FMR1* gene, were cultured and treated with FUDR as described previously [58]. Metaphase spreads were prepared by the conventional “dropping” method. The metaphase chromosomes were subjected to FISH analysis using a BAC probe that binds specifically to FRAXA (*green*). DNA was stained with DAPI (*blue*). The *yellow arrow* indicates the FRAXA FISH probe signal that is apparently detached from the X chromosome



**Fig. 4** A representative example of the analysis of telomere fragility. U2OS cells were treated with APH (0.4  $\mu\text{M}$ ) and RO3306 as described previously [26]. Metaphase cells were collected by mitotic shake-off and dropped onto slides for metaphase FISH analysis with a PNA telomere probe following the manufacturer's instructions (Panagene, Korea). The chromosomes were stained with DAPI (*blue*) and the telomeres were detected with FAM labeled PNA telomere probe (*green*). The blown-up images below show examples of atypical telomere signals

DNA damage response pathways, including  $\gamma\text{H2AX}$  [22], phospho-DNA-PKcs [22], and SMC1 [49], are recruited to CFSs.

We showed recently that the Fanconi anemia proteins, FANCD2 and FANCI, bind specifically to CFS loci in cells treated with a low dose of APH irrespective of whether the chromosome is broken or not in metaphase [23]. These proteins form a focus on each sister chromatid at the location of the CFS. Most interestingly, when cells harboring these foci are analyzed in anaphase, the foci are often linked by a DNA bridge called an ultrafine DNA bridge (UFB). It is possible that at CFS loci where DNA replication is disrupted or otherwise delayed, the sister chromatids are still interlinked by regions of underreplicated DNA when cells enter anaphase. This would then lead to the formation of a UFB. Most of the time, these replication intermediates can be processed or resolved by proteins including PICH, BLM, and RPA [50]. However, some of the bridges/UFBs may lead to detrimental consequences, such

as the formation of micronuclei or 53BP1 bodies in the next G1 phase [51]. We showed recently that underreplicated CFSs could be detected and processed by structure-specific endonucleases (e.g., the MUS81-EME1 complex) in early mitosis, and that this promotes CFS expression [25]. This discovery is consistent with the notion that the expression of CFSs is not an accidental event, but instead is a programmed and regulated process. Moreover, we have proposed that CFS expression is beneficial for the maintenance of genome stability because it prevents much more hazardous events, such as irreversible chromosome missegregation. Most recently, we have demonstrated that, following treatment of cells with APH, entry into the mitotic prophase triggers the recruitment of MUS81-EME1 to CFSs, and the nuclease activity of MUS81 then promotes POLD3-dependent DNA synthesis at CFSs. When analyzed on metaphase chromosomes, the sites of new DNA synthesis in mitosis correspond to the expressed CFSs, indicating that the gap/break in the chromosome is the location of a repair event that delays condensation of the locus. This DNA repair-based replication “salvage” pathway serves as a mechanism to minimize chromosome missegregation and nondisjunction, particularly in cancer cells with elevated levels of chromosome instability (CIN +ve) [26].

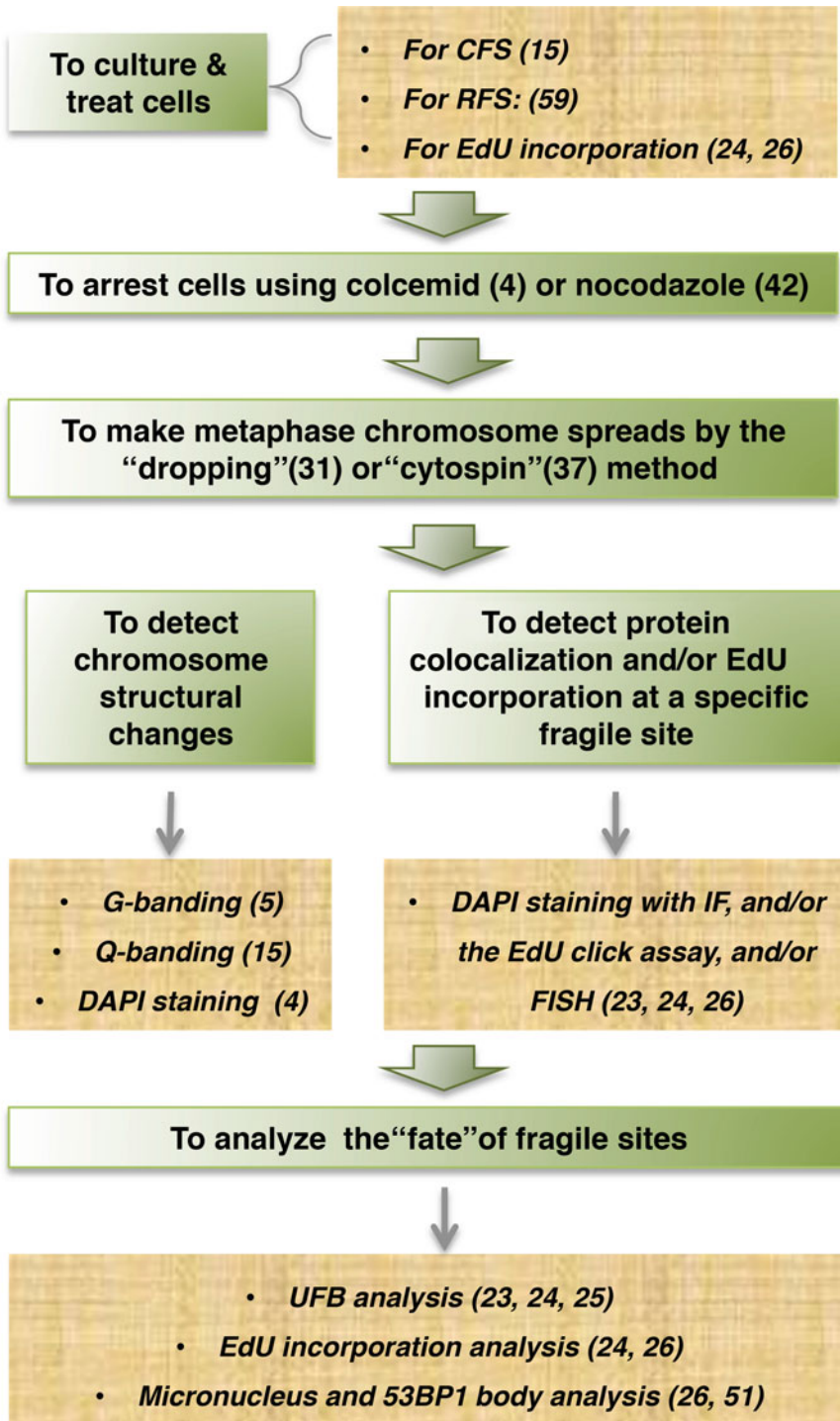
In the case of RFSs, most of the studies have been focused on the FRAXA (fragile X) locus that is associated with Fragile X syndrome. This disorder is characterized by a spectrum of intellectual disabilities, as well as atypical physical characteristics (including an elongated face, large or protruding ears, and large testicles), and behavioral characteristics such as stereotypic movements (e.g., hand-flapping), and social anxiety. It is well established that Fragile X syndrome is caused by the expansion of the CGG trinucleotide repeat that affects the expression of the Fragile X mental retardation 1 gene (*FMR1*) [52, 53]. In the past three decades, great progress has been made in deciphering the mechanism underlying the expansion of the repeats that denote such folate-sensitive RFSs [54–56]. However, the mechanism underlying the fragility of these loci and other CGG trinucleotide repeat regions remains unknown. Undoubtedly, the methodologies that have been developed to define the fragility of CFSs can be applied to the analysis of this type of fragile site.

---

## 4 Concluding Remarks

In the past five decades, a considerable amount has been learned about chromosome fragile sites in the human genome. The methods involved in the detection and analysis of chromosome fragile sites have progressed from the original cytogenetic-based techniques to fluorescence based ones (i.e., IF and FISH) (summarized in Fig. 5). This has allowed us to understand why fragile sites are





**Fig. 5** A flowchart of the methodology involved in the induction, detection, and analysis of human chromosome fragile sites. References for the methods applied are indicated where appropriate



unusually “fragile”, and the strategies utilized by cells to cope with this fragility to ensure the faithful replication and segregation of DNA during the cell division cycle. In addition to the methods mentioned above, great strides have been made in imaging technologies that allow for more sensitive, faster, and reliable detection of genomic loci. Moreover, high content array CGH technology has enabled researchers to detect DNA copy number changes in an interval of 1 marker per 1737 bases cross the whole human genome with a input of DNA as little as 250 ng (e.g., via the CytoScan<sup>®</sup> HD platform developed by Affymetrix).

It seems likely that chromosome fragile sites will remain an intriguing feature of human cells and a fascinating topic for research for the following reasons: (1) CFSs exist in every individual and many are in conserved chromosome regions in mammals. This begs the question as to why that is the case, and whether there is any advantage for the cells to maintain those apparently fragile loci during evolution? (2) Several different types of DNA synthesis inhibitors can induce the expression of either CFS or RFS, and hence it would be interesting to find out why that is the case. Is this because of the DNA sequence differences at each fragile site, or because different inhibitors initiate a different DNA damage response? (3) Errors arising during the repair or replication at fragile sites following replication stress (that can be either endogenous, such as oncogene associated, or exogenous, such as folate deficiency) lead to chromosome translocations or an abnormal number of chromosomes. This can, in turn, drive the development of several debilitating disorders (e.g. cancer, cognitive dysfunction, and infertility). It will be a high priority to identify the causes of those errors. (4) The recently identified fragility at telomeric sites raises questions regarding which factors increase or decrease the expression of TFSs in cancer cells, considering that there is evidence to show that APH can increase the expression of TFSs in ALT cells [57]. The answer to this question would potentially allow us to develop novel anticancer drug treatments given that cancer cells have shown unusual telomere length maintenance either via an active telomerase or the ALT mechanism.

---

## Acknowledgment

We thank Theresa Wass, Malgorzata Clausen, and Wei Wu for experimental help. Work in the authors' laboratory is supported by a Faculty Ph.D. fellowship from the University of Copenhagen (V.A.B.), the European Union (O.O., I.D.H., and Y.L.), The Danish National Research Foundation (DNRF115; I.D.H. and Y.L.), and The Nordea Foundation (I.D.H.).

## References

1. Durkin SG, Glover TW (2007) Chromosome fragile sites. *Annu Rev Genet* 41:169–192
2. Kremer EJ et al (1991) Mapping of DNA instability at the Fragile-X to a trinucleotide repeat sequence P(Ccg)N. *Science* 252 (5013):1711–1714
3. Sutherland GR, Baker E, Richards RI (1998) Fragile sites still breaking. *Trends Genet* 14 (12):501–506
4. Mrasek K et al (2010) Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones. *Int J Oncol* 36(4):929–940
5. Sutherland GR (2003) Rare fragile sites. *Cytogenet Genome Res* 100(1–4):77–84
6. Dekaban A (1965) Persisting clone of cells with an abnormal chromosome in a woman previously irradiated. *J Nucl Med* 6(10):740–746
7. Richards RI (2001) Fragile and unstable chromosomes in cancer: causes and consequences. *Trends Genet* 17(6):339–345
8. Bell MV et al (1991) Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. *Cell* 64 (4):861–866
9. Fu YH et al (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67(6):1047–1058
10. Pieretti M et al (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66(4):817–822
11. Verkerk AJ et al (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65 (5):905–914
12. El Achkar E, Gerbault-Seureau M, Muleris M, Dutrillaux B, Debatisse M (2005) Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites. *Proc Natl Acad Sci U S A* 102(50):18069–18074
13. Handt O, Sutherland GR, Richards RI (2000) Fragile sites and minisatellite repeat instability. *Mol Genet Metab* 70(2):99–105
14. Functammasan A, Walsh E, Chiaromonte F, Eckert KA, Makova KD (2012) A genome-wide analysis of common fragile sites: what features determine chromosomal instability in the human genome? *Genome Res* 22 (6):993–1005
15. Glover TW, Berger C, Coyle J, Echo B (1984) DNA polymerase-alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human-chromosomes. *Hum Genet* 67 (2):136–142
16. Rao PN, Heerema NA, Palmer CG (1988) Fragile sites induced by Fudr, caffeine, and aphidicolin - their frequency, distribution, and analysis. *Hum Genet* 78(1):21–26
17. Kuwano A, Murano I, Kajii T (1990) Cell type-dependent difference in the distribution and frequency of excess thymidine-induced common fragile sites - lymphocytes-T and skin fibroblasts. *Hum Genet* 84(6):527–531
18. Sutherland GR (1991) Chromosomal fragile sites. *Genet Analysis Techn Appl* 8(6):161–166
19. Letessier A et al (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 470(7332):120–123
20. Le Tallec B et al (2013) Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. *Cell Rep* 4 (3):420–428
21. Wilson TE et al (2015) Large transcription units unify copy number variants and common fragile sites arising under replication stress. *Genome Res* 25(2):189–200
22. Schwartz M et al (2005) Homologous recombination and nonhomologous end-joining repair pathways regulate fragile site stability. *Genes Dev* 19(22):2715–2726
23. Chan KL, Palmari-Pallag T, Ying S, Hickson ID (2009) Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol* 11(6):753–760
24. Naim V, Wilhelm T, Debatisse M, Rosselli F (2013) ERCC1 and MUS81-EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis. *Nat Cell Biol* 15 (8):1008–U1270
25. Ying S et al (2013) MUS81 promotes common fragile site expression. *Nat Cell Biol* 15 (8):1001–1007
26. Minocherhomji S et al (2015) Replication stress activates DNA repair synthesis in mitosis. *Nature* 528(7581):286–290
27. Sutherland GR (1977) Fragile sites on human chromosomes: demonstration of their dependence on the type of tissue culture medium. *Science* 197(4300):265–266
28. Sutherland GR (1979) Heritable fragile sites on human chromosomes I. Factors affecting expression in lymphocyte culture. *Am J Hum Genet* 31(2):125–135
29. Glover TW (1981) FUDR induction of the X chromosome fragile site: evidence for the

- mechanism of folic acid and thymidine inhibition. *Am J Hum Genet* 33(2):234–242
30. Sutherland GR (1983) The fragile X chromosome. *Int Rev Cytol* 81:107–143
  31. Keagle MB, Gersen SL (2013) Basic cytogenetics laboratory procedures. In: Keagle MB, Gersen SL (eds) *The principles of clinical cytogenetics*, 3rd edn. Springer Science+Business Media, New York, pp 53–65
  32. Le Tallec B et al (2011) Molecular profiling of common fragile sites in human fibroblasts. *Nat Struct Mol Biol* 18(12):1421–1423
  33. Elder FFB, Robinson TJ (1989) Common fragile sites do not correlate with cancer chromosome breakpoints in the laboratory mouse and rat. *Cancer Genet Cytogen* 41(2):246–246
  34. Elder FFB, Robinson TJ (1989) Rodent common fragile sites - are they conserved - evidence from mouse and rat. *Chromosoma* 97(6):459–464
  35. Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T (2006) Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res* 16(10):1222–1230
  36. Helmrich A et al (2007) Identification of the human/mouse syntenic common fragile site FRA7K/Fra12C1—relation of FRA7K and other human common fragile sites on chromosome 7 to evolutionary breakpoints. *Int J Cancer* 120(1):48–54
  37. Sullivan BA, Warburton PE (1999) Studying the progression of vertebrate chromosomes through mitosis by immunofluorescence and FISH. In: Warburton PE (ed) *Chromosomes structural analysis: a practical approach*. Oxford University Press, Oxford, pp 81–101
  38. de Lange T et al (1990) Structure and variability of human chromosome ends. *Mol Cell Biol* 10(2):518–527
  39. Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* 43(2 Pt 1):405–413
  40. Bodnar AG et al (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science* 279(5349):349–352
  41. Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, Reddel RR (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* 3(11):1271–1274
  42. Sfeir A et al (2009) Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* 138(1):90–103
  43. Lansdorp PM et al (1996) Heterogeneity in telomere length of human chromosomes. *Hum Mol Genet* 5(5):685–691
  44. Vannier JB, Pavicic-Kaltenbrunner V, Petalcorin MIR, Ding H, Boulton SJ (2012) RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell* 149(4):795–806
  45. Vannier JB et al (2013) RTEL1 is a replisome-associated helicase that promotes telomere and genome-wide replication. *Science* 342(6155):239–242
  46. Wan BB et al (2013) SLX4 assembles a telomere maintenance toolkit by bridging multiple endonucleases with telomeres. *Cell Rep* 4(5):861–869
  47. Sarek G, Vannier JB, Panier S, Petrini JHJ, Boulton SJ (2015) TRF2 recruits RTEL1 to telomeres in S phase to promote T-loop unwinding. *Mol Cell* 57(4):622–635
  48. Helmrich A, Ballarino M, Tora L (2011) Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell* 44(6):966–977
  49. Musio A et al (2005) SMC1 involvement in fragile site expression. *Hum Mol Genet* 14(4):525–533
  50. Liu Y, Nielsen CF, Yao Q, Hickson ID (2014) The origins and processing of ultra fine anaphase DNA bridges. *Curr Opin Genet Dev* 26:1–5
  51. Lukas C et al (2011) 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol* 13(3):243–253
  52. Colak D et al (2014) Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 343(6174):1002–1005
  53. Gerhardt J et al (2014) The DNA replication program is altered at the FMR1 locus in fragile X embryonic stem cells. *Mol Cell* 53(1):19–31
  54. Lee DY, McMurray CT (2014) Trinucleotide expansion in disease: why is there a length threshold? *Curr Opin Genet Dev* 26:131–140
  55. Su XA, Dion V, Gasser SM, Freudenreich CH (2015) Regulation of recombination at yeast nuclear pores controls repair and triplet repeat stability. *Genes Dev* 29(10):1006–1017
  56. Usdin K, House NC, Freudenreich CH (2015) Repeat instability during DNA repair: insights from model systems. *Crit Rev Biochem Mol Biol* 50(2):142–167
  57. Root H et al (2016) FANCD2 limits BLM-dependent telomere instability in the alternative lengthening of telomeres pathway. *Hum Mol Genet* 25(15):3255–3268
  58. Kumari D et al (2009) The role of DNA damage response pathways in chromosome fragility in Fragile X syndrome. *Nucleic Acids Res* 37(13):4385–4392

## Imaging of DNA Ultrafine Bridges in Budding Yeast

Oliver Quevedo and Michael Lisby

### Abstract

DNA ultrafine bridges (UFBs) are a type of chromatin-free DNA bridges that connect sister chromatids in anaphase and pose a threat to genome stability. However, little is known about the origin of these structures, and how they are sensed and resolved by the cell. In this chapter, we review tools and methods for studying UFBs by fluorescence microscopy including chemical and genetic approaches to induce UFBs in the model organism *Saccharomyces cerevisiae*.

**Key words** Anaphase bridges, UFBs, Fluorescence microscopy, Yeast, Ultrafine bridges

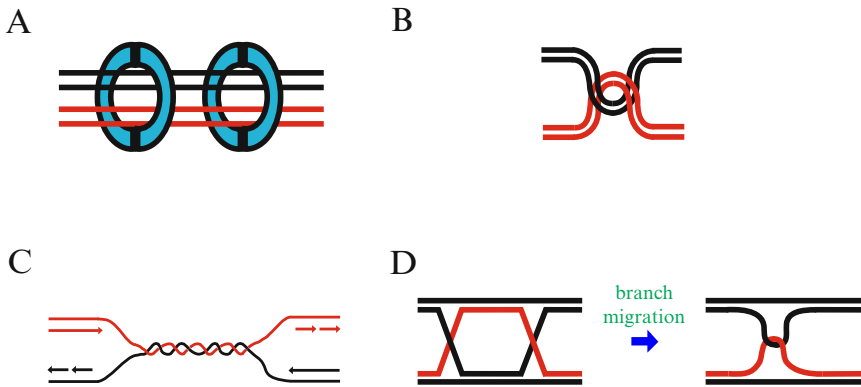
---

### 1 Introduction

The failure to fully separate sister chromatids during mitosis can lead to the appearance of DNA anaphase bridges, which is a potential source of genomic instability, a hallmark of cancer cells [1–4]. Thus, removal of sister chromatid linkages is critical in order to avoid the formation of anaphase bridges, and specialized mechanisms exist in the cell to ensure that sister chromatids are no longer held together upon anaphase entry.

So far, two main types of anaphase bridges have been described [5, 6]: chromatin bridges and ultrafine bridges (UFBs). Contrary to the former, UFBs do not contain histones and other chromatin components and cannot be visualized by staining with conventional DNA dyes such as DAPI or Hoechst. Nevertheless, they can be visualized by monitoring a subset of proteins known to bind UFBs, such as the BLM helicase (Sgs1 in budding yeast), PICH, or FANCM (Mph1 in budding yeast), among others [7–9]. Strikingly, UFBs are observed in up to 30% of dividing cells in unperturbed conditions [8].

Four main types of linkages between sister chromatids are thought to lead to the appearance of anaphase bridges [10] (Fig. 1): (A) linkages mediated by the cohesin complex, (B) catenations and other topological entanglements, (C) the presence of



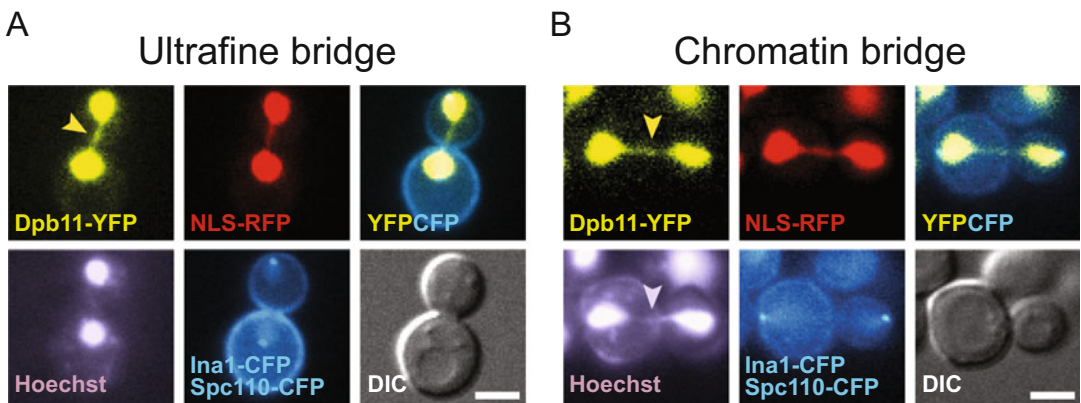
**Fig. 1** Schematic overview of the main types of linkages that lead to the appearance of anaphase bridges during mitosis. **(a)** Sister chromatids held together by the cohesin complex. **(b)** Catenations between sister chromatids. **(c)** Underreplicated regions. **(d)** Holliday junctions that can branch migrate to hemicatenanes

underreplicated regions at mitotic entry, and (D) the presence of recombination intermediates that have not been removed before anaphase/telophase. Hence, anaphase bridges may contain single-stranded gaps, stalled replication forks, or replication/recombination intermediates. As mentioned above, the linkages established between sister chromatids are removed by a number of mechanisms involving separase (removal of the cohesin complex), condensin and type II topoisomerases (removal of catenations and other topological constraints), DNA helicases such as BLM/Sgs1 and the type I topoisomerase Top3 (branch migration and removal of hemicatenanes), structure-selective endonucleases (removal of recombination intermediates), and unscheduled DNA synthesis during mitosis promoted by proteins such as TopBP1/Dpb11 (removal of underreplicated regions). These mechanisms work in a coordinated and regulated manner according to cell cycle progression, and their misregulation leads to the persistence of anaphase bridges [11]. Notably, failures in the removal of cohesion and the presence of recombination intermediates during mitosis lead mainly to the accumulation of chromatin bridges [10, 12–17], whereas catenanes and topological entanglements can potentially lead to both chromatin bridges and UFBs [6, 10, 12–19], and the presence of underreplicated regions during mitosis leads mainly to UFBs [6, 15, 20].

In humans, two types of UFBs have been described [6]: (1) Fanconi anemia (FA)-associated UFBs, originating from regions of the genome in which the completion of replication can be problematic, such as the chromosomal common fragile sites (CFSs), and induced by inhibition of replication [5, 21, 22]; and (2) non-FA-associated UFBs, originated from DNA catenations and induced by inhibition of Topoisomerase 2 [6]. Whether a similar distinction can be made in *S. cerevisiae* using the FANCM homologue, Mph1, remains to be tested.

### 1.1 Budding Yeast as a Model to Study Anaphase Bridges

*Saccharomyces cerevisiae* has a long track-record as a powerful genetic and cell biological model system [23]. However, when it comes to imaging of anaphase bridges, the closed mitosis of budding yeast poses a technical challenge. In contrast to higher eukaryotes, there is no disassembly of the yeast nuclear envelope during mitosis. Thus, a nuclear protein displaying a bridge-like structure connecting the separating nuclei in mitosis could reflect the nucleoplasm contained in the so-called anaphase tube rather than a protein-bound DNA bridge. To overcome the challenge of distinguishing DNA anaphase bridges from nucleoplasm, it is necessary to use additional markers that provide an indication of the segregation status of the chromosomes. For instance, since relaxation of the spindle tension only occurs after the sister chromatids have been fully resolved during anaphase/telophase, monitoring the lengthening of the mitotic spindle by time-lapse microscopy provides a good indication of sister chromatid separation. A spindle pole body (SPB) marker, such as Spc110 fused to a fluorescent protein, provides an accurate tool for monitoring the spindle length. Using this marker, together with another fluorescent protein fused to a nuclear localization signal (NLS), as a marker for the nucleoplasm, it was shown that the resolution of the Dpb11-bound DNA anaphase bridges takes place concomitant with relaxation of the mitotic spindle, whereas the anaphase tube remains until abscission [15]. Another indicator of the presence of DNA bridges in the division plane during mitotic exit is the activation of the *NoCut* checkpoint [24]. This checkpoint senses the presence of both chromatin bridges and UFBs [15, 24]. Activation of the *NoCut* checkpoint causes a delay in cytokinesis [24]. This delay can be monitored by the use of a plasma membrane marker fused to a fluorescent protein as an indicator of abscission [25] (Fig. 2). Both



**Fig. 2** Examples of anaphase bridges. Cells (ML734-9B: *MATa tTA(tetR-VP16)-tetO<sub>2</sub>-DPB11-4ala-YFP::KanMX NLS-yEmRFPrv::URA3 SPC110-CFP::KAN INA1-4ala-CFP*) were grown in SC + Ade to mid-log phase and stained with Hoechst before imaging (see Subheading 3.1 and Notes 1 and 2). (a) Ultrafine bridge. (b) Chromatin bridge. Yellow arrowheads indicate anaphase bridges. Purple arrowhead indicates Hoechst-stained chromatin bridge. Scale bar, 3  $\mu$ m

**Table 1**

**Protein markers for analysis of anaphase bridges. After tagging with a fluorescent protein (see Note 2), these proteins constitute fluorescent reporters to visualize UFBs, chromatin bridges, nucleoplasm, nucleolus, and the *NoCut* checkpoint delay**

UFB markers	Dpb11, Rfa1, Sgs1, Top3, Top2, Ddc2, Nop1 (nucleolus), Net1 (nucleolus)
Chromatin markers	Hta1, Hta2, Htz1, Nhp10, Rsc1
Anaphase tube/ nucleoplasm marker	Nup49 (nuclear pore complex), NLS-XFP <sup>a</sup> (nucleoplasm)
<i>NoCut</i> checkpoint	Ina1, 2xPH-XFP <sup>a</sup> (plasma membrane)

<sup>a</sup>XFP indicates any fluorescent protein

the PH domain and the Ina1 protein, fused to CFP, have been used to monitor abscission [15, 25]. Other markers that can be used to determine the structure of the anaphase tube are proteins that localize to the nuclear envelope, such as the nuclear pore complex subunit Nup49. This protein does not mark the nucleoplasm, but allows for visualization of the anaphase tube [15]. Thus, combining a spindle marker with a nucleoplasm/nuclear envelope marker is a good approach to distinguish proteins bound to DNA bridges in anaphase from those bridge-like structures that reflect the nucleoplasm contained in the anaphase tube. As a marker for anaphase DNA bridges, there is a set of proteins that are known to bind UFBs: Dpb11, the Sgs1-Top3-Rmi1 complex, RPA, and Top2 (Table 1). Chromatin bridges can be distinguished from UFBs by including a chromatin marker such as histone H2A or by staining DNA with DAPI or Hoechst, which do not detect UFBs. Finally, the presence of DNA in UFBs can be demonstrated by the incorporation of and staining for the nucleoside analogue 5-ethynyl-2'-deoxyuridine (EdU) [15].

## **1.2 Induction of Anaphase Bridges**

Even if UFBs can be detected in unperturbed cells, it can be useful to apply conditions that increase their frequency, either by the use of mutants or by the use of drugs that are known to interfere with DNA metabolic processes. For instance, the use of the thermosensitive allele *top2-1* at the permissive temperature was shown to increase the number of UFBs by ~4-fold in asynchronous cultures [15]. As for the use of drugs to induce an increase in the frequency of UFBs, both methyl methanesulfonate (MMS) and hydroxyurea (HU) can be used. Transient exposure (70 min) to 0.03% MMS leads to a ~5-fold increase of UFBs [15]. Similarly, exposure to 5.6 mM HU for 2 h leads to a ~4-fold increase in the frequency of yeast cells with UFBs in asynchronous culture (unpublished data).

---

## 2 Materials

1. SC + Ade medium: synthetic complete medium supplemented with 100 µg/ml adenine [26].
2. SC + Ade patches (SC + Ade/LMPA): SC + Ade with 1% low melting-point agarose (V2111; Promega).
3. DAPI solution: 10 mg/ml stock solution in deionized water (1000×).
4. Hoechst 33258 solution: 5 mg/ml stock solution in deionized water (1000×) (B2883; Sigma-Aldrich).
5.  $\alpha$ -factor mating pheromone.
6. Hydroxyurea.
7. Wax: equal volumes of molten bees' wax, petroleum jelly (VWR Scientific), and lanolin (Sigma).
8. Paraformaldehyde (PFA): For a 3.7% PFA solution, add 3.7 g of PFA to approximately 70 ml of deionized water. Heat to 60 °C in a water bath. Add a few drops of 1 M NaOH until the PFA dissolves. Cool to room temperature and add 10 ml of 10× PBS. Adjust the pH to 7.4 and bring to a final volume of 100 ml with deionized water.

---

## 3 Methods

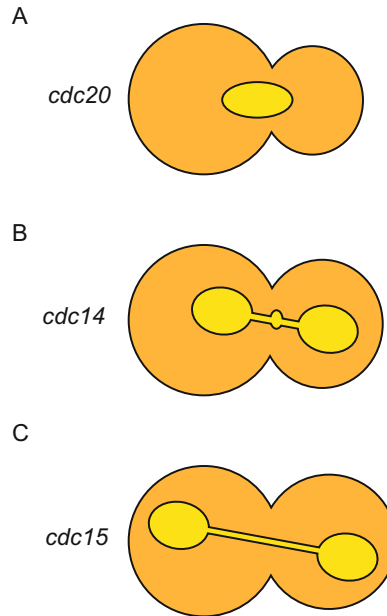
### 3.1 Live Cell Staining of Chromatin Bridges Using Conventional DNA Dyes

1. Grow cells in SC + Ade medium.
2. Transfer 1 ml of culture to a separate tube. Add DAPI (to a final concentration of 10 µg/ml) or Hoechst 33258 (to a final concentration of 5 µg/ml) and incubate cells with shaking for 10 min at the same temperature used in **step 1**.
3. Spin down cells at 1000 rpm, wash once with SC + Ade, and resuspend in 50–100 µl SC + Ade.
4. Transfer cells to a microscope slide for imaging under the microscope (*see Note 3*).

### 3.2 Visualization of Anaphase Bridges in Metaphase/Anaphase/Telophase Using *cdc20*, *cdc14* or *cdc15* Thermosensitive Mutants

Since their discovery in the 1970s by Hartwell and colleagues in their study of the genetic control of the cell division cycle [27], the thermosensitive alleles of genes encoding proteins controlling different aspects of the cell cycle progression have been extensively used to synchronize yeast cultures. Examples of those alleles are the *cdc20-1* allele that has been used to block the cell cycle in metaphase [27]; the *cdc14-1* allele used to synchronize cells at the anaphase-to-telophase transition [27]; and the *cdc15-2* allele used to block cells in telophase [27] (Fig. 3).





**Fig. 3** Schematic representation of Dpb11-YFP localization in a metaphase, anaphase, or telophase cell cycle arrest. (a) Metaphase arrest by disruption of the Cdc20 function. (b) Anaphase arrest by disruption of the Cdc14 function. In this case it is possible to observe bulges in the bridge that have been shown to colocalize with unresolved regions of the ribosomal DNA [37]. (c) Telophase arrest by disruption of the Cdc15 function. Yellow, Dpb11-YFP signal. Orange, cytoplasm

1. Grow cells in SC + Ade until early log phase (OD<sub>600</sub> = 0.2–0.5) at 25 °C (permissive temperature) with moderate shaking.
2. Incubate cells at 37 °C (restrictive temperature) with moderate shaking for 3 h. Then, collect a sample and proceed as in **steps 3** and **4** above (*see* Subheading 3.1) in order to image directly under the microscope. If a DNA dye is used for chromatin staining, proceed as stated above (*see* Subheading 3.1). It is crucial that the incubation temperature is 37 °C during the staining.
3. To release cells from the cell cycle arrest, transfer the culture back to 25 °C with moderate shaking and collect samples at appropriate times for imaging under the microscope, proceeding as stated before. If a DNA dye is used for chromatin staining, proceed as stated above (*see* Subheading 3.1). In this case, the incubation temperature during the staining should be 25 °C.

**3.3 Visualization of Anaphase Bridges in Metaphase/Anaphase/Telophase Using the Top2-1 Thermosensitive Mutant**

As mentioned above, haploid cells carrying the thermosensitive *top2-1* allele have a higher frequency of UFBs [15], making this mutant suitable for studying this type of bridges. However, it is not possible to use a combination of *top2-1* and the thermosensitive *cdc20/cdc14/cdc15* mutants to block the cell cycle in metaphase/anaphase/telophase, since the incubation at the restrictive temperature required for these cell cycle arrests would inactivate the

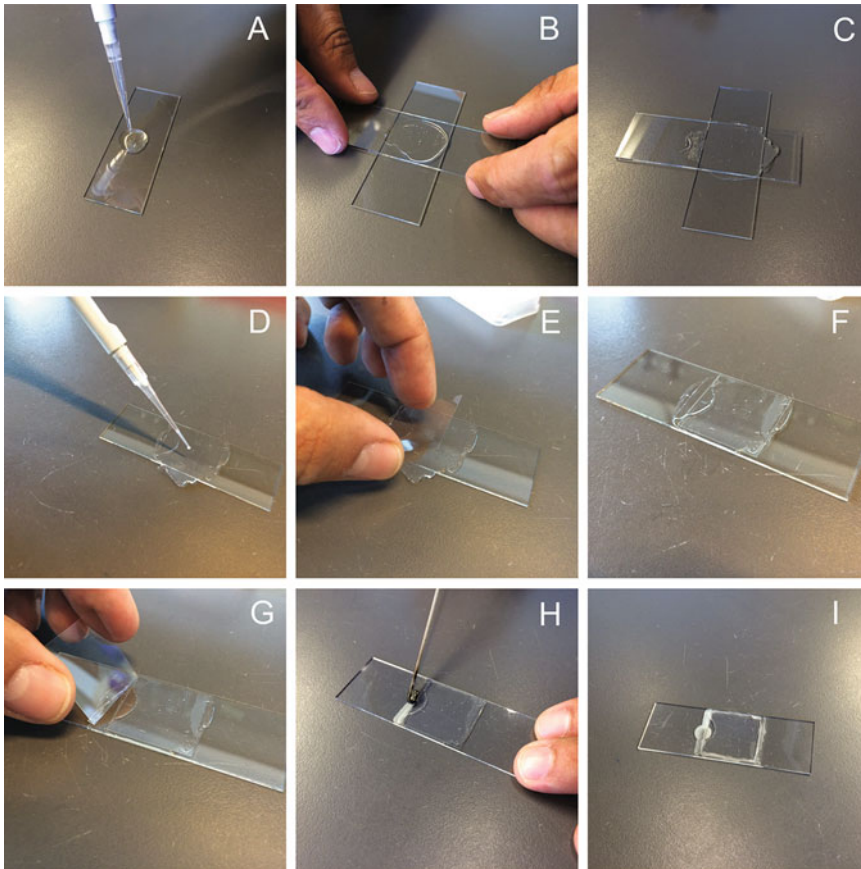
Topoisomerase 2 [12], leading to an increase of chromatin bridges instead of UFBs [15]. Thus, an alternative strategy to inactivate *CDC20*, *CDC14*, and *CDC15* must be applied. For instance, the methionine repressible *MET3* promoter has been used to switch off the expression of *CDC20* and thus synchronize cells in metaphase followed by release into anaphase upon removal of methionine [28]. For the disruption of Cdc14 and Cdc15 function, different conditional non-thermosensitive alleles have been successfully used in yeast, such as the mouse ornithine decarboxylase for the depletion of Cdc14 [29] and the Cdc15-as1 mutant protein that can be inhibited by the ATP analog “PPI analog 8” [30, 31]. In these cases, the protocol to visualize UFBs is similar to the one described (*see* Subheadings 3.1 and 3.2), but keeping the temperature at 25 °C and inducing the disruption of any of the conditional alleles immediately after the release from G<sub>1</sub> (*see* Subheading 3.4).

### **3.4 Time-Course to Monitor the Formation and Resolution of Anaphase Bridges Induced by Mild Replication Stress**

1. Grow *MATa* haploid cells in SC + Ade until early log phase (OD<sub>600</sub> = 0.2–0.5) at 25 °C/30 °C with moderate shaking.
2. Add the  $\alpha$ -factor pheromone to the culture to synchronize cells in G<sub>1</sub>. If the strain that is being used is *bar1Δ*, then use a final concentration of 50 ng/ml and an incubation time of 3 h to allow all cells to reach the G<sub>1</sub> arrest. If the strain used is *BARI*, then use a final concentration of 5  $\mu$ g/ml. Incubate cells for 1 h and then add fresh  $\alpha$ -factor (half of the amount used before). Incubate cells for an additional hour and verify the arrest under the microscope.
3. To release cells from the G<sub>1</sub> arrest, wash cells twice with 1 volume of fresh medium. Then, resuspend pellet in 1 volume fresh medium containing 5.6 mM HU and incubate cells at 25 °C/30 °C with moderate shaking. Collect samples at appropriate time points for imaging. If a DNA dye is used for chromatin staining, proceed as stated above (*see* Subheading 3.1). Spin cells at 1000 rpm, wash once with SC + Ade, and resuspend cells in 0.1 volume of SC + Ade.
4. Transfer cells to a microscope slide for imaging under the microscope (*see* Note 3).

### **3.5 Preparation of Samples for Time-Lapse Experiments**

1. Prepare a microscope slide with a patch of SC + Ade solid medium. For preparation of the patch, add 200  $\mu$ l of melted SC + Ade containing 1% low melting point agarose to the center of a microscope slide. Then cover with another slide and wait 10 min for the agarose to solidify. Finally, remove the top slide very carefully in order to avoid breakage of the SC + Ade patch (Fig. 4).



**Fig. 4** Scheme for the preparation of SC + Ade patches on slides for time-lapse microscopy. **(a)** Add 200  $\mu$ l melted SC + Ade/LMPA to the slide; **(b, c)** cover with another slide, spreading the SC + Ade as much as possible without breaking the patch. Allow to solidify for 10 min; **(d)** carefully remove the top slide, trying to avoid breaking the patch, and add 2.8  $\mu$ l of cell suspension; **(e, f)** cover the cells with a coverslip; **(g)** remove the excess of SC + Ade/LMPA; **(h, i)** use wax to carefully seal the edges of the coverslip and proceed to imaging

2. If a synchronized culture is to be used, proceed first with the synchronization of the culture in the desired cell cycle stage as stated above (*see* Subheadings 3.2–3.4).
3. Immediately after release, take 500  $\mu$ l of cell culture and spin down cells at 1000 rpm.
4. Discard supernatant leaving 50–100  $\mu$ l of medium in the tube. Then, resuspend the pellet in the remaining medium.
5. Take 2.8  $\mu$ l of cell suspension and carefully pipette onto the SC + Ade patch on the microscope slide.
6. Cover with a coverslip. Finally, seal the edges using melted wax and proceed to imaging.

### 3.6 EdU Staining of Anaphase Bridges

1. Grow cells in SC + Ade until early log phase ( $OD_{600} = 0.2\text{--}0.5$ ) at 25 °C/30 °C with moderate shaking.
2. Add 20  $\mu\text{M}$  EdU to the medium and incubate cells for 3 h with moderate shaking (*see Note 4*).
3. Collect cells by centrifugation at 1000 rpm.
4. Fix cells by resuspension in 3.7% PFA for 15 min at room temperature with shaking.
5. Quench PFA with 0.3 M glycine for 15 min at room temperature.
6. Wash fixed cells in PBS.
7. Stain with 10  $\mu\text{g}/\text{ml}$  DAPI or 5  $\mu\text{g}/\text{ml}$  Hoechst 33258 for 10 min.
8. Proceed to detection as described in the Click-iT EdU Alexa Fluor 594 imaging kit manual (Invitrogen).

---

## 4 Notes

1. Microscopy was performed on a wide-field AxioImager Z1 (Carl Zeiss MicroImaging, Inc) equipped with a 100 $\times$  objective lens (Zeiss PLAN-APO, NA 1.4), a cooled Orca-ER CCD camera (Hamamatsu, Japan), differential interference contrast (DIC), and a Zeiss HXP120C illumination source. Images were acquired and processed using Volocity (PerkinElmer) software. Images were pseudocolored according to the approximate emission wavelength of the fluorophores.
2. Proteins were tagged with the cyan fluorescent protein (CFP, clone W7) [32], yellow fluorescent protein (YFP, clone 10C) [33], and red fluorescent proteins (RFP, clone yEmRFP) [34] as described [35].
3. The thickness of the coverslip should match the objective lens and the volume of the sample should match the area of the coverslip to restrict all cells to the same focal plane. In this study, a cell suspension of 2.8  $\mu\text{l}$  was mounted with a no. 1.5 coverslip (24 mm  $\times$  24 mm).
4. For incorporation of EdU, cells should be engineered to express thymidine kinase and a nucleoside transporter [36].

---

## Acknowledgments

This work was supported by The Danish Agency for Science, Technology and Innovation, the Villum Foundation, and the Danish National Research Foundation (DNRF115). We thank Camilla Colding and Giedrė Bačinskaja for comments on the manuscript.

## References

1. Draviam VM, Xie S, Sorger PK (2004) Chromosome segregation and genomic stability. *Curr Opin Genet Dev* 14:120–125
2. Gollin SM (2005) Mechanisms leading to chromosomal instability. *Semin Cancer Biol* 15:33–42
3. Bannon JH, Mc Gee MM (2009) Understanding the role of aneuploidy in tumorigenesis. *Biochem Soc Trans* 37:910–913
4. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144:646–674
5. Chan KL, Hickson ID (2009) On the origins of ultra-fine anaphase bridges. *Cell Cycle* 8:3065–3066
6. Chan KL, Hickson ID (2011) New insights into the formation and resolution of ultra-fine anaphase bridges. *Semin Cell Dev Biol* 22:906–912
7. Baumann C, Körner R, Hofmann K et al (2007) PICH, a centromere-associated SNF2 family ATPase, is regulated by Plk1 and required for the spindle checkpoint. *Cell* 128:101–114
8. Chan KL, North PS, Hickson ID (2007) BLM is required for faithful chromosome segregation and its localization defines a class of ultra-fine anaphase bridges. *EMBO J* 26:3397–3409
9. Kaulich M, Cubizolles F, Nigg EA (2012) On the regulation, function, and localization of the DNA-dependent ATPase PICH. *Chromosoma* 121:395–408
10. Díaz-Martínez LA, Giménez-Abián JF, Clarke DJ (2008) Chromosome cohesion - rings, knots, orcs and fellowship. *J Cell Sci* 121:2107–2114
11. Holm C, Goto T, Wang JC et al (1985) DNA topoisomerase II is required at the time of mitosis in yeast. *Cell* 41:553–563
12. Holm C, Stearns T, Botstein D (1989) DNA topoisomerase II must act at mitosis to prevent nondisjunction and chromosome breakage. *Mol Cell Biol* 9:159–168
13. Strunnikov AV, Hogan E, Koshland D (1995) SMC2, a *Saccharomyces cerevisiae* gene essential for chromosome segregation and condensation, defines a subgroup within the SMC family. *Genes Dev* 9:587–599
14. Uhlmann F, Lottspeich F, Nasmyth K (1999) Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* 400:37–42
15. Germann SM, Schramke V, Pedersen RT et al (2014) TopBP1/Dpb11 binds DNA anaphase bridges to prevent genome instability. *J Cell Biol* 204:45–59
16. Spence JM, Phua HH, Mills W et al (2007) Depletion of topoisomerase IIalpha leads to shortening of the metaphase interkinetochore distance and abnormal persistence of PICH-coated anaphase threads. *J Cell Sci* 120:3952–3964
17. Gallina I, Christiansen SK, Pedersen RT et al (2016) TopBP1-mediated DNA processing during mitosis. *Cell Cycle* 15:176–183
18. Bhalla N, Biggins S, Murray AW (2002) Mutation of YCS4, a budding yeast condensin subunit, affects mitotic and nonmitotic chromosome behavior. *Mol Biol Cell* 13:632–645
19. Broderick R, Nieminuszczy J, Blackford AN et al (2015) TOPBP1 recruits TOP2A to ultra-fine anaphase bridges to aid in their resolution. *Nat Commun* 6:6572
20. Pedersen RT, Kruse T, Nilsson J et al (2015) TopBP1 is required at mitosis to reduce transmission of DNA damage to G1 daughter cells. *J Cell Biol* 210:565–582
21. Naim V, Rosselli F (2009) The FANCD1 pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol* 11:761–768
22. Vinciguerra P, Godinho SA, Parmar K et al (2010) Cytokinesis failure occurs in Fanconi anemia pathway-deficient murine and human bone marrow hematopoietic cells. *J Clin Invest* 120:3834–3842
23. Botstein D, Fink GR (2011) Yeast: an experimental organism for 21st Century biology. *Genetics* 189:695–704
24. Norden C, Mendoza M, Dobbelaere J et al (2006) The NoCut pathway links completion of cytokinesis to spindle midzone function to prevent chromosome breakage. *Cell* 125:85–98
25. Mendoza M, Norden C, Durrer K et al (2009) A mechanism for chromosome segregation sensing by the NoCut checkpoint. *Nat Cell Biol* 11:477–483
26. Sherman F (2002) Getting started with yeast. *Methods Enzymol* 350:3–41
27. Hartwell LH, Mortimer RK, Culotti J et al (1973) Genetic control of the cell division cycle in yeast: V. Genetic analysis of *cdc* mutants. *Genetics* 74:267–286
28. Rocuzzo M, Visintin C, Tili F et al (2015) FEAR-mediated activation of Cdc14 is the limiting step for spindle elongation and anaphase progression. *Nat Cell Biol* 17(3):251–261

29. Jungbluth M, Renicke C, Taxis C (2010) Targeted protein depletion in *Saccharomyces cerevisiae* by activation of a bidirectional degron. *BMC Syst Biol* 4:176
30. Bishop AC, Ubersax JA, Petsch DT et al (2000) A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* 407:395–401
31. D’Aquino KE, Monje-Casas F, Paulson J et al (2005) The protein kinase Kin4 inhibits exit from mitosis in response to spindle position defects. *Mol Cell* 19:223–234
32. Heim R, Tsien RY (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol* 6:178–182
33. Ormö M, Cubitt AB, Kallio K et al (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* 273:1392–1395
34. Keppler-Ross S, Noffz C, Dean N (2008) A new purple fluorescent color marker for genetic studies in *Saccharomyces cerevisiae* and *Candida albicans*. *Genetics* 179:705–710
35. Silva S, Gallina I, Eckert-Boulet N et al (2012) Live cell microscopy of DNA damage response in *Saccharomyces cerevisiae*. *Methods Mol Biol* 920:433–443
36. Viggiani CJ, Aparicio OM (2006) New vectors for simplified construction of BrdU-Incorporating strains of *Saccharomyces cerevisiae*. *Yeast* 23:1045–1051
37. Quevedo O, García-Luis J, Matos-Perdomo E et al (2012) Nondisjunction of a single chromosome leads to breakage and activation of DNA damage checkpoint in G2. *PLoS Genet* 8:e1002509

## Detection of Ultrafine Anaphase Bridges

Anna H. Bizard, Christian F. Nielsen, and Ian D. Hickson

### Abstract

Ultrafine anaphase bridges (UFBs) are thin DNA threads linking the separating sister chromatids in the anaphase of mitosis. UFBs are thought to form when topological DNA entanglements between two chromatids are not resolved prior to anaphase onset. In contrast to other markers of defective chromosome segregation, UFBs cannot be detected by direct staining of the DNA, but instead can be detected using immunofluorescence-based approaches. Due to the fact that they are short-lived and fragile in nature, UFBs can be challenging to detect. In this chapter, we describe methods that have been optimized for successful detection of UFBs. We also provide guidelines for the optimization of UFBs detection depending on the antibody and the cell line to be used.

**Key words** Ultrafine anaphase bridges, Common fragile sites, Centromeres, Telomeres, Genome stability, Chromosome segregation

---

### 1 Introduction

The removal of all DNA entanglement between the sister chromatids is a prerequisite for the accurate segregation of the daughter genomes in mitosis. Nevertheless, certain genomic loci can remain entangled at the onset of anaphase. In anaphase, such persistent DNA entanglements are converted into ultra fine DNA threads bridging the two segregating genomes. These structures are the so-called “ultrafine anaphase bridges” (UFBs). In contrast to other segregation defects, such as lagging chromatin or chromatin bridges, UFBs cannot be stained by conventional DNA dyes. Instead, they can be detected by immunofluorescence, using antibodies raised against UFB-associated factors such as the SNF2-family DNA translocase, PICH, and the RecQ DNA helicase BLM [1, 2].

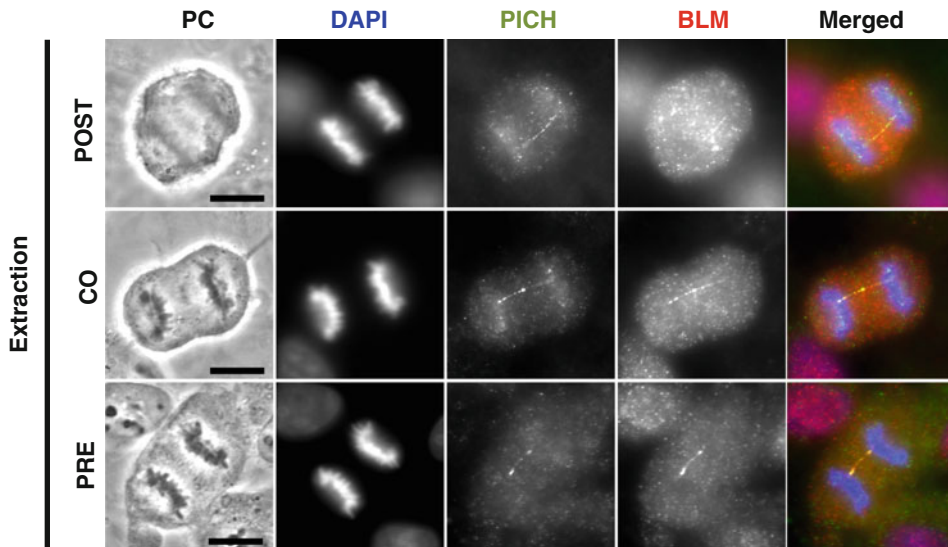
Several subclasses of UFBs have been identified. These are defined by the genomic locus from which they originate. The most common are the centromeric-UFBs (CEN-UFBs) arising due to persistent catenation of the centromeric chromatin.

CEN-UFBs are abundant at anaphase onset but are rapidly resolved by topoisomerase II $\alpha$  during anaphase. Hence, their number rapidly drops as the cells progress through anaphase, unless cells are treated with Topoisomerase II inhibitor (such as ICRF-193) [2–5]. A second source of UFBs is from common fragile sites (CFS). In contrast to CEN-UFBs, CFS-UFBs are quite rare in an unperturbed cell cycle, but can be specifically induced by exposing the cells to DNA replication stress [6]. Besides centromeres and CFSs, other critical regions of the genomes that have been shown to be prone to the formation of UFBs are telomeres and the ribosomal-DNA loci [7, 8].

There is some evidence to suggest that UFBs that are not resolved during mitosis can cause the cell to abort cytokinesis, potentially generating aneuploidy [5]. Hence, UFBs might play a role in human diseases associated with genome instability [9, 10]. Despite this, the analysis of UFBs has been hampered by a lack of standardized methods for their detection. It is noteworthy that UFBs can be challenging to visualize when routine immunofluorescence protocols are used. This is due largely to three main issues. First, UFBs are short-lived and can only be observed during the anaphase and telophase of mitosis, which each represent only a very short stage of the cell division cycle. Second, UFBs are relatively fragile, and can be broken if the cells are not handled carefully prior to, during and even after being fixed. Finally, UFBs are very thin, and hence the intensity of immunostaining can easily drop below the level of detection for a conventional fluorescence microscope.

Accordingly, successful detection of UFBs relies on classical immunofluorescence protocols, in which some critical steps are executed in a way that preserves the full range of structures present in an anaphase cell. Depending on the cell line and the antibodies used to detect UFBs, we routinely use three different permeabilization procedures that differ from one another by the timing of the application of Triton to the cells (i.e., before, during, or after the fixation step for pre-permeabilization, co-permeabilization, and post-permeabilization procedures, respectively). The importance of the permeabilization procedure choice is exemplified in Fig. 1, which shows representative images of UFBs obtained with antibodies directed against PICH and BLM. While the anti-PICH antibody allows good detection of UFBs independently of the fixation procedure used, detection using the anti-BLM antibody requires either co-permeabilized or pre-permeabilized fixed cells. It is noteworthy that the influence of the permeabilization protocol on the quality of the detected UFBs also greatly depends on the cell line studied. In this chapter, we describe in greatest detail the pre-permeabilized protocol, which is the most widely used and sensitive. The co-permeabilization and post-permeabilization protocols are also described as alternative protocols (*see Note 1*).





**Fig. 1** Representative images of UFBs detected using the anti-PICH (*green*) or anti-BLM (*red*) antibody, in HT1080 cells fixed with the post-permeabilization (*top*), co-permeabilization (*middle*), or pre-permeabilization (*bottom*) procedures. DNA was stained with DAPI (*blue*). PC shows the phase contrast images of each anaphase B cell. Images were taken with a conventional wide field fluorescence microscope. *Scale bar*, 10  $\mu\text{m}$

Finally, while UFBs can be detected in virtually all cell lines, it is noteworthy that some cell lines are more suitable for UFB analysis than others. For routine UFB analysis, it is, therefore, advisable to perform the analysis with adherent cell lines that can be grown directly on glass coverslips, such as HT1080, U2OS, or fibroblasts such as GMO00637. Nevertheless, it is possible to successfully detect UFBs when using suspension cell lines. In this case, cells should be seeded onto a poly-lysine-coated surface immediately before the fixation (*see Note 3*).

## 2 Materials

### 2.1 Slides

1. Sterile coverslips 22  $\times$  22 mm (Thickness No. 1.5).
2. Microscope slides 75  $\times$  25  $\times$  1 mm.
3. For humidified chambers: Plastic petri dishes; Whatman 3MM, and Parafilm discs at the dimensions of the petri dishes.

### 2.2 Buffers and Solutions

#### 2.2.1 Stock Solutions

1. 20 $\times$  Phosphate Buffer Saline (20 $\times$  PBS): 2.8 M NaCl, 54 mM KCl, 130 mM  $\text{Na}_2\text{HPO}_4$ , and 30 mM  $\text{KH}_2\text{PO}_4$ , pH adjusted to 7.4 with 1 M HCl. Filter and store at room temperature.
2. 37 wt. % paraformaldehyde (PFA) solution (stabilized with 10–15% methanol) (Sigma, 25,2549). Store at room temperature protected from light.

3. 10× PEM Buffer: 200 mM PIPES pH 6.8, 10 mM MgCl<sub>2</sub>, and 100 mM EGTA. Filter and store at room temperature.
4. 10% Triton X-100, diluted in ddH<sub>2</sub>O. Store at room temperature.
5. 30% (w.v) bovine serum albumin (BSA), in ddH<sub>2</sub>O. Filter and store and 4 °C.
6. Vectashield mounting medium for fluorescence, (without DAPI) (Vectors Lab, H-1000).
7. DAPI (1 mg/ml in ddH<sub>2</sub>O). Make small aliquots and store at -20 °C protected from light.
8. Transparent nail polish.

*2.2.2 Working Solutions  
(to Be Prepared for  
Immediate Use and  
Depending on the Chosen  
Permeabilization  
Procedure)*

9. Pre-permeabilization Buffer A: 0.2% Triton in 1× PEM Buffer. Prepare 1 ml per coverslip.
10. Pre-permeabilization Buffer B: 0.1% Triton, 8% PFA in 1× PEM Buffer). Prepare 2 ml per coverslip.
11. Co-permeabilization Buffer: 0.4% Triton, 8% PFA, in 1× PEM). Prepare 1 ml per coverslip.
12. Post-permeabilization Buffer: 8% PFA, in 1× PBS). Prepare 1 ml per coverslip.
13. Post-staining Fixation Buffer: 4% PFA in 1× PBS). Prepare 200 µl per coverslip.

*2.2.3 Other Solutions  
(Can Be Prepared in  
Advance)*

14. 1× PBS (1/20 dilution of the stock; in ddH<sub>2</sub>O). Prepare 20 ml per coverslip. Store at room temperature.
15. PBSAT Buffer: 3% BSA, 0.5% Triton X-100 in 1× PBS). Filter, aliquot in 50 ml Falcon tubes and store at -20 °C.
16. 1× DAPI (1/10.000 dilution of the stock; in 1× PBS). Store at 4 °C protected from light.

**2.3 Antibodies**

1. Mouse monoclonal anti PICH antibody (clone 14226-3; Millipore, 0.41540). Store a 4 °C. Recommended working dilution: 1/50 in PBSAT.
2. Rabbit polyclonal anti BLM antibody (Abcam, ab2175). Aliquot and store at -20 °C. Recommended working dilution: 1/200 in PBSAT.
3. Alexa Fluor<sup>®</sup> 488 goat anti-mouse IgG (Life Technologies, A11001). Store at 4 °C protected from light. Recommended working dilution: 1/500 in PBSAT.
4. Alexa Fluor<sup>®</sup> 568 goat anti-rabbit IgG (Life Technologies, A11001). Store at 4 °C protected from light. Recommended working dilution: 1/500 in PBSAT.

### 3 Methods

#### 3.1 Cell Culture for Adherent Cell Lines

The protocol described below has been optimized for an asynchronously growing population of HT1080 cells (*see Note 2*). Some adjustments might be required if other adherent cell lines were to be used. If suspension cells are being used, *see Note 3*.

The following protocol describes the procedure for preparing six coverslips to be fixed simultaneously.

##### DAY 1

1. Place a sterile coverslip in each well of a 6-well plate (*see Note 4*). Add 1 ml of fresh medium to the wells. Ensure that the medium covers the coverslips, and that the coverslips are well seated on the bottom of each well by applying a little pressure on each coverslip using a pipette tip or sterile tweezers.
2. From an approximately 80% confluent, T75 flask of a healthy growing population of cells, aspirate the medium, wash the cells once with PBS, add 2.5–3 ml of trypsin and incubate for 5 min at 37 °C. Once the cells have detached, neutralize the trypsin by adding 5 ml of complete medium to the flask. Transfer the cells to a 15 ml Falcon tube and harvest the cells by centrifugation at  $300 \times g$  for 5 min at room temperature.
3. Aspirate the supernatant and resuspend the cell pellet in 1 ml of fresh medium using a Gilson P1000 pipette (or similar). Pipette the suspension up and down several times, with the tip touching the bottom of the tube in order to dislodge cell aggregates (*see Note 5*).
4. Add culture medium in order to reach a cell density of  $0.3 \times 10^6$  cells/ml (*see Note 6*).
5. Distribute 1 ml of cells to each well of the 6-well plate.
6. Homogenize the cell suspension by pipetting up and down once, and immediately place the plate in the incubator.
7. Let the cells adhere to the surface and proliferate for 40–48 h until they reach approximately 80–90% confluence (*see Note 7*).

#### 3.2 Cell Permeabilization and Fixation (Pre-permeabilization Protocol)

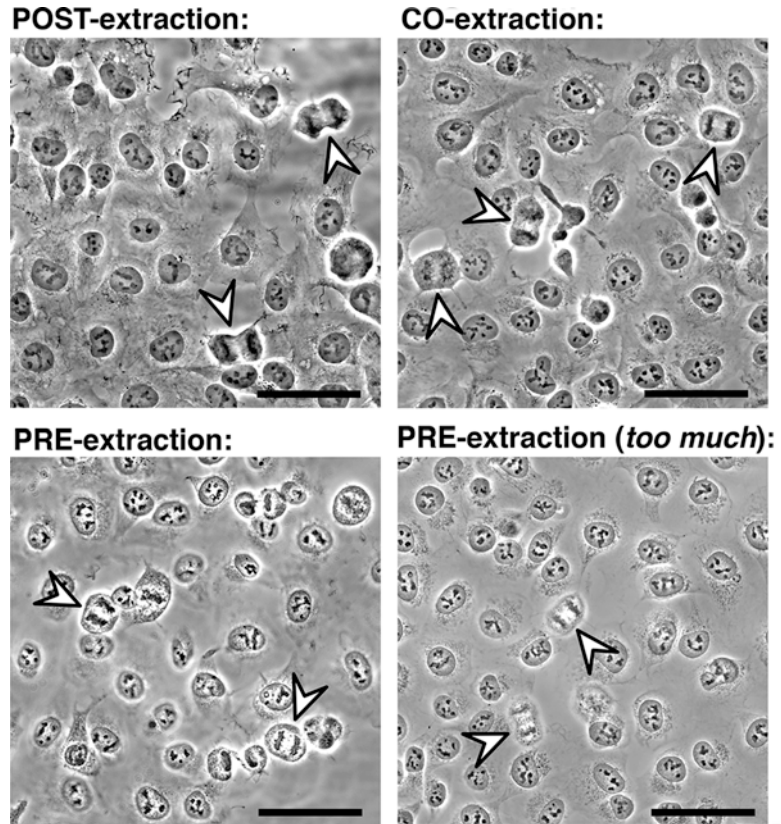
For the co-permeabilization and post-permeabilization protocols, *see Notes 8 and 9* respectively.

##### DAY 3

*Unless otherwise stated, each subsequent step is performed at room temperature (see Note 10).*

1. Using a conventional inverted light microscope, verify that the culture has reached the expected confluence (*see Note 11*) and that the mitotic cells can be observed at a frequency compatible with the subsequent analysis (*see Note 12*).

2. Prepare sufficient quantity of pre-permeabilization Buffer A and pre-permeabilization Buffer B as described in Subheading **2.2**, **items 9** and **10**.
3. Set up all reagents required for the permeabilization and the fixation of cells in a fume hood (Buffers prepared in Subheading **3.2**, **step 2**; PBS 1×; and appropriate liquid waste disposals) (*see Note 13*).
4. Take the 6-well plate out of the incubator and place in a fume hood. Avoid agitation of the cells as much as possible (*see Note 14*). Once the plate has been placed in the hood, it should not be moved until the end of the fixation procedure. Proceed immediately to fixation.
5. Using a P1000, remove 1 ml of medium from the side of each well. The coverslips are now covered by only 1 ml of medium (*see Note 15*).
6. Add 1 ml of PBS down one side of each well. Immediately take out 1 ml of medium/PBS mixture from the opposite side of each well (*see Note 16*).
7. Repeat **step 6** one more time. The cells are now immersed in 1 ml of PBS (*see Note 17*).
8. Proceed to pre-permeabilization by adding 1 ml of pre-permeabilization Buffer A down the side of each well.
9. Carefully mix the buffer in each well by transferring 1 ml from one side to the opposite side of each well. At this stage, the pipetting needs to be quick, but gentle, smooth and regular. Expect this procedure to require approximately 30 s for the entire plate.
10. Incubate for 60 s at room temperature (*see Note 18*).
11. Add 2 ml of pre-permeabilization Buffer B to each well. Maintain the same frequency and order as in **step 8** to ensure that each well is being pre-permeabilized for approximately the same amount of time.
12. Incubate for 15 min at room temperature. Do not move the plate until fixation is over.
13. Once the cells are fixed, discard the 4 ml from each well into a dedicated chemical waste container. Rinse the coverslips once with 4 ml PBS.
14. Wash the coverslips further with 2 ml PBS for 5 min. Repeat three times.
15. Further permeabilize and saturate the coverslips by incubating them overnight at 4 °C in 1 ml sterile PBSAT (*see Note 19*). The coverslips can be kept in PBSAT at 4 °C for up to 1 week before being stained (*see Note 20*).



**Fig. 2** Phase contrast images of asynchronously growing HT1080 cells fixed with the post-permeabilization (*top left*), co-permeabilization (*top right*), or pre-permeabilization (*bottom left*) procedures. An example of cells fixed following excessive pre-permeabilization is also shown (*bottom right*). In each panel, *arrowheads* indicate cells in anaphase B. Images were taken with a conventional wide field microscope. *Scale bar*, 50  $\mu\text{m}$

16. Using a conventional inverted light microscope, verify that the cells have been fixed properly (*see Fig. 2* and **Note 21**). Be particularly rigorous with this step when the pre-permeabilized procedure has been used (*see Notes 22* and **23**).

### 3.3 Staining

#### DAY 4

1. Prepare a humidity chamber by placing a wet disc of Whatman 3MM paper, covered by a disc of Parafilm into a petri dish (*see Note 24*). Ensure that the surface of the Parafilm is flat.
2. Place the coverslip onto the Parafilm with the cells facing up (*see Note 25*). Immediately cover the cells with 200  $\mu\text{l}$  PBSAT (*see Note 26*).

3. Working on ice, dilute the primary antibody in ice cold PBSAT (100  $\mu$ l per coverslip). Centrifuge the diluted antibody at  $5000 \times g$  for 2 min at 4 °C.
4. Remove as much as possible of the PBSAT from the coverslip and then immediately cover it with 75–100  $\mu$ l of the diluted antibody. Ensure that the antibody solution spreads evenly over the entire surface of the coverslip (*see Note 27*).
5. Close and seal the petri dish with Parafilm. Incubate overnight at 4 °C.

#### DAY 5

6. After the primary antibody incubation is completed, wash the coverslip with 200  $\mu$ l PBSAT for 10 min. Repeat the washing step twice more.
7. While waiting for the completion of the last wash step, dilute the secondary antibody in ice cold PBSAT (100  $\mu$ l per coverslip). This procedure should be performed on ice and the solution protected from light wherever possible. Centrifuge the diluted antibody at  $5000 \times g$  for 5 min at 4 °C.
8. Once the last wash is completed, remove as much as possible of the PBSAT from the coverslip and immediately cover it with 75–100  $\mu$ l of the diluted secondary antibody. Ensure that the antibody solution spreads over the entire surface of the coverslip (*see Note 27*).
9. Incubate for 2 h at room temperature and protect from light.
10. After the secondary antibody incubation is completed, wash the coverslip with 200  $\mu$ l PBSAT for 15 min. Repeat the wash step three times.
11. Perform one more wash with PBS for 10 min. Eventually proceed with post-staining fixation (*see Note 28*).
12. Cover the coverslip with 200  $\mu$ l of DAPI solution that is diluted 1/10,000 in PBS.
13. Incubate for 5 min at room temperature and protect from light.
14. Wash twice with PBS for 5 min.
15. Label the slides and apply a drop (approx. 10  $\mu$ l) of Vectashield mounting medium.
16. Once the washes are completed, remove the PBS from the coverslip and carefully rinse the coverslip with 200  $\mu$ l of ddH<sub>2</sub>O.
17. Drain as much of the liquid as possible from the coverslip and, using tweezers, carefully place the coverslip onto the drop of mounting medium with the cells facing down.

18. Place the slide upside down on the piece of tissue paper, and allow any excess mounting medium to drain away for 15–30 min (the coverslip is now sandwiched between the tissue paper and the slide).
19. Seal the coverslip to the slide by spreading nail polish around the edges of the coverslip.
20. The slides must be kept at 4 °C and protected from the light until analysis (*see Note 29*).
21. Using a conventional fluorescence microscope, verify the quality of the staining, particularly of mitotic cells. Besides the presence of UFBs (*see Fig. 1*), the staining can be validated by a rapid analysis of characteristic staining pattern of the used antibody (*see Note 30* for PICH, and *Note 31* for BLM).
22. Proceed to the analysis of UFBs, by counting the number of UFBs per anaphase B cell (*see Note 32*). Results are typically expressed as the proportion of anaphase B with at least 1 UFBs and as the average number of UFBs per anaphase B (*see Note 33*).

---

## 4 Notes

1. In order to maximize the chance of successful UFB detection, each time a new combination of cell line and antibody is to be used, it is advisable to test all three of the fixation protocols.

**Fixation protocol optimization:** When using a new combination of cells and antibodies, proceed to a test experiment, in which the cells are fixed using each of the three protocols, and then stained with the desired antibodies. Assess the quality of the fixation and of the staining for each fixation procedure. If two protocols show an equivalent quality of staining, always favor post-permeabilization and co-permeabilization over the pre-permeabilization protocol because these protocols are more reproducible, and preserve better the different structures.

2. The frequency of anaphases present in an asynchronously growing population of cells is more than enough to allow quantitative UFBs analysis. Nevertheless, it is sometimes preferable to analyze a population of cells synchronously progressing throughout mitosis after release from a nocodazole block. The synchronization protocol needs to be optimized according to the cell line. Briefly, after the cells have been arrested in prometaphase in the presence of nocodazole, the mitotic cells are detached from the surface by tapping the side of the flask with the palm of the hand three or four times (generally termed mitotic shake-off). The detached cells are then washed twice in warm PBS. The cells are then resuspended in prewarmed

medium and seeded on poly-lysine-coated coverslips as described for suspension cells (*see Note 3*).

3. **When using cells that grow in suspension**, proceed as follows: on the day of the fixation place a poly-lysine-coated coverslip into each well of a 6-well plate; add 1 ml of medium to each well and ensure that the coverslips are well seated on the bottom of the wells. Distribute 1 ml of cell suspension into each well (approx.  $0.5 \times 10^6$  per well) and let them settle for 15–30 min. Proceed to “fixation and staining” as described in Subheadings 3.2 and 3.3.
4. Coverslips are usually supplied pre-sterilized, and therefore do not require sterilization—as long as the box has not been opened in a non-sterile environment. If the sterility of the coverslips has been compromised, transfer them into a glass petri dish, and microwave them for 15 s at maximum intensity.
5. This step ensures that any cells clumps or aggregates are dispersed, and allows the formation of a homogeneous layer of cells on the surface of the coverslip.
6. Optimize according to the doubling time of the cell line to be used. We generally aim to allow the cells to go through two rounds of division between the seeding and the fixation steps (*see Note 5*).
7. It is important that the cells are actively proliferating at the time of fixation. A high cell density is important to maximize the number of mitotic events that can be scored per sample.
8. **Alternative protocol: Co-permeabilization.** Use the pre-permeabilization protocol through Subheading 3.2, step 7, and then add 1 ml of co-permeabilization Buffer down the side of the well. Carefully mix by transferring 1 ml from one side of the well to the other, and incubate for 15 min. Proceed to the wash steps described in the main protocol (from Subheading 3.2, step 13).
9. **Alternative protocol: Post-permeabilization.** Use the pre-permeabilization protocol through Subheading 3.2, step 5, and then fix the cells by adding 1 ml of post-permeabilization Buffer. Incubate for 15 min at room temperature and then proceed to the steps described in the main protocol (from Subheading 3.2, step 13).
10. Be aware that the pre-permeabilization and co-permeabilization efficiency can be affected by even small variations of temperature. Keep the samples at 20–25 °C wherever possible. It is possible to conduct the fixation at 4 °C, but this can under some circumstances induce microtubule depolymerization and lead to the release of the tension holding the UFBs.



11. The cells should be 80–90% confluent. If the confluence is less than 50%, incubate the culture further until it reaches a good confluence (*see* Fig. 2, for examples of expected confluence). If the cells are overgrown, it is preferable to start again from Subheading 3.1, step 1.
12. Depending on the cell type, expect approximately 5% of the cells to be mitotic (*see* Fig. 2, for examples of expected mitotic density). A high mitotic index not only maximizes the number of events that can be scored per sample, but also is a good indicator of the “healthiness” of the cells.
13. The success of this procedure essentially relies on accurate timing and smooth manipulation of the cells. Therefore, it is important to organize properly the fume hood and to ensure that all materials and buffers are carefully organized and are easily accessible.
14. It is essential to keep in mind that mitotic cells are poorly adhesive to the surface and that UFBs are short-lived and fragile in nature. Hence, the coverslips must be processed as gently as possible, before, during, and after the fixation process.
15. In order to minimize the perturbations applied on the cells while they are being washed and fixed, always add buffer from the same side of the well, and always remove buffer from the opposite side of the well. Always proceed with the different wells of the plate in the same order. By using a Gilson P1000 for this and the following steps helps to maintain a regular rhythm during periods when medium is added or removed.
16. During this and the following steps, the fact that the cells remain constantly immersed in solution is critical in order to avoid the detachment of mitotic cells from the coverslip.
17. The fraction of remaining media is negligible for the following steps.
18. The incubation time might need to be adjusted depending on the cell type used. It usually varies between 30 s and 2 min.
19. As an alternative, the coverslips can be saturated in PBSAT for 2 h at room temperature.
20. For longer storage, keep coverslips in PBS with 0.02% sodium azide. Seal the plate with Parafilm to prevent evaporation.
21. Regardless of the procedure used, verify that the mitotic cells have not been detached from the coverslips during the fixation (*see* Fig. 2 for examples of cells after each different fixation procedure). The number of mitotic cells visible at this stage should be similar to what was observed prior to fixation. Moreover, the mitotic cells should appear perfectly fixed to the coverslips. If the mitotic cells are not static upon moving

the plate, it is likely that the procedure has not been performed gently and/or quickly enough. These mitotic cells will be particularly difficult to image. It is recommended to repeat the fixation. Moving mitotic cells can also suggest a too extensive pre-permeabilization (*see* **Note 23**).

22. The extent to which the cells have successfully been pre-permeabilized is critical for UFBs analysis. Typically, the cytoplasm of interphase cells should be still visible, while the contours of mitotic cells should be clearly marked (*see* Fig. 2 for examples). If the pre-permeabilization step is too extensive, the staining of some of the bridges will be too weak to be detected (*see* **Note 23**).
23. If there are indications that the pre-permeabilization is too extensive, it is preferable to repeat the fixation with a modified pre-permeabilization step. It is most likely that a reduction in the incubation time (Subheading 3.2, step 10) will improve pre-permeabilization. Otherwise, one might consider, either to decrease the concentration of Triton in the pre-permeabilization buffer, or to select one of the other fixation procedures.
24. We usually can place two and four coverslips, respectively, into each 10 and 15 cm petri dish. Careful labeling of the samples is essential in this case.
25. During the staining part of the protocol the manipulation of the coverslip is minimized, in order to avoid mitotic cells being detached from the coverslips or UFBs from being broken or distorted.
26. It is crucial to prevent the coverslips from drying out, as this will irreversibly affect the quality of the staining. Hence, the time during which the coverslips are not covered by buffer should be minimized.
27. At this stage, if the antibody does not spread evenly over the entire surface of the coverslip, it might indicate that the coverslip has dried out from the previous step. The quality of the staining will be irreversibly impaired on any dried region of the coverslip.
28. **Post-staining fixation.** At this stage, proceeding to a second fixation step can help to increase the stability of the slides over time. For this, remove as much PBS as possible from the coverslip and cover the cells with 200  $\mu$ l post-staining fixation Buffer. Incubate for 5 min at room temperature and wash three times for 5 min with PBS. Proceed to DAPI staining as indicated in the main protocol (Subheading 3.3, step 12).
29. Normally, the slides can be kept up to 1 month at 4 °C. It is highly recommended, however, to proceed to the analysis as soon as possible, as the quality of the staining may decline

overtime. In order to increase the “working-life” of the stored slides, it is advisable to perform a post-staining fixation (*see Note 26*).

30. Besides its localization on UFBs, PICH localizes to the cytoplasm of interphases cells, and to the centromeres of mitotic cells [2, 11]. Typically, staining using the antibody directed against PICH is relatively clean but is often quite faint (*see Fig. 1*).
31. Besides this localization to UFBs, BLM localizes at PML bodies in a subset of interphase cells [12]. The antibody directed against BLM overall gives more intense staining than does the antibody directed against PICH. However, the signal from the soluble fraction of the cells can interfere with the proper visualization of BLM on UFBs. For this reason, it is recommended to proceed with the co-permeabilized or the pre-permeabilized procedure when using this antibody (*see Fig. 1*).
32. It is important to restrict analysis to the anaphase B stage of mitosis. Otherwise, UFBs will be over-represented from the very early stage of anaphase due to the presence of centromeric UFBs. Anaphase B can be distinguished from anaphase A by the distance between the segregating DNA masses, as well as by the general form of the cell (ovoid more than spherical). However, it is also important to make the distinction between anaphase B and telophase (during which the DNA is partially decondensed and the daughter cells are clearly visible) (*see Figs. 1 and 2 for examples of anaphase B cells*).
33. In an untreated, asynchronously growing population of cells, expect between 5 and 20% of the anaphase B cells to contain at least 1 UFB.

---

## Acknowledgments

Work in the authors' laboratory is supported by The European Research Council and The Danish National Research Foundation (grant DNR115).

## References

1. Chan KL, North PS, Hickson ID (2007) BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *EMBO J* 26 (14):3397–3409. doi:[10.1038/sj.emboj.7601777](https://doi.org/10.1038/sj.emboj.7601777)
2. Baumann C, Korner R, Hofmann K, Nigg EA (2007) PICH, a centromere-associated SNF2 family ATPase, is regulated by Plk1 and required for the spindle checkpoint. *Cell* 128 (1):101–114. doi:[10.1016/j.cell.2006.11.041](https://doi.org/10.1016/j.cell.2006.11.041)
3. Wang LH, Schwarzbraun T, Speicher MR, Nigg EA (2008) Persistence of DNA threads in human anaphase cells suggests late completion of sister chromatid decatenation. *Chromosoma* 117(2):123–135. doi:[10.1007/s00412-007-0131-7](https://doi.org/10.1007/s00412-007-0131-7)

4. Hengeveld RC, de Boer HR, Schoonen PM, de Vries EG, Lens SM, van Vugt MA (2015) Rif1 is required for resolution of ultrafine DNA bridges in anaphase to ensure genomic stability. *Dev Cell* 34(4):466–474. doi:[10.1016/j.devcel.2015.06.014](https://doi.org/10.1016/j.devcel.2015.06.014)
5. Nielsen CF, Huttner D, Bizard AH, Hirano S, Li TN, Palmai-Pallag T, Bjerregaard VA, Liu Y, Nigg EA, Wang LH, Hickson ID (2015) PICH promotes sister chromatid disjunction and co-operates with topoisomerase II in mitosis. *Nat Commun* 6:8962. doi:[10.1038/ncomms9962](https://doi.org/10.1038/ncomms9962)
6. Chan KL, Palmai-Pallag T, Ying S, Hickson ID (2009) Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol* 11(6):753–760. doi:[10.1038/ncb1882](https://doi.org/10.1038/ncb1882)
7. Nielsen CF, Hickson ID (2016) PICH promotes mitotic chromosome segregation: identification of a novel role in ribosomal DNA disjunction. *Cell Cycle* 15(20):2704–2711
8. Barefield C, Karlseder J (2012) The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures. *Nucleic Acids Res* 40(15):7358–7367. doi:[10.1093/nar/gks407](https://doi.org/10.1093/nar/gks407)
9. Vinciguerra P, Godinho SA, Parmar K, Pellman D, D’Andrea AD (2010) Cytokinesis failure occurs in Fanconi anemia pathway-deficient murine and human bone marrow hematopoietic cells. *J Clin Invest* 120(11):3834–3842. doi:[10.1172/JCI43391](https://doi.org/10.1172/JCI43391)
10. Nera B, Huang HS, Lai T, Xu L (2015) Elevated levels of TRF2 induce telomeric ultrafine anaphase bridges and rapid telomere deletions. *Nat Commun* 6:10132. doi:[10.1038/ncomms10132](https://doi.org/10.1038/ncomms10132)
11. Kaulich M, Cubizolles F, Nigg EA (2012) On the regulation, function, and localization of the DNA-dependent ATPase PICH. *Chromosoma* 121(4):395–408. doi:[10.1007/s00412-012-0370-0](https://doi.org/10.1007/s00412-012-0370-0)
12. Yankiwski V, Marciniak RA, Guarente L, Neff NF (2000) Nuclear structure in normal and Bloom syndrome cells. *Proc Natl Acad Sci U S A* 97(10):5214–5219. doi:[10.1073/pnas.090525897](https://doi.org/10.1073/pnas.090525897)

## A Chromatin Fiber Analysis Pipeline to Model DNA Synthesis and Structures in Fission Yeast

Sarah A. Sabatinos and Marc D. Green

### Abstract

Chromatin fibers, first described by Jackson and Pombo (J Cell Biol 140(6):1285–1295, 1998) are prepared from cells lysed on glass coverslips, and require minimal equipment to produce. Since the DNA is not previously treated with denaturing agents, proteins are left intact and may be used to model other DNA-based processes. Such an analysis can be daunting, without a rigorous method for analysis. We describe a pipeline for chromatin fiber use to model DNA replication complexes. Full protocols for chromatin fiber preparation and staining are presented. Further, we have developed an analysis algorithm for One Dimensional Data—Boolean Logic Operations Binning System (ODD-BLOBS). This freely available software defines replication and protein tracts, measures their lengths, and then correlates replicated areas with protein distributions. Our methods and analysis are tested in *Schizosaccharomyces pombe* (fission yeast) but may be applied to model replication structures across multiple organisms.

**Key words** Chromatin fibers, Replication tracts, Fission yeast, Fiber spreads, Replication tract analysis

---

### 1 Introduction

*Schizosaccharomyces pombe* is an excellent model organism to study replication dynamics and complexes, but high-resolution imaging is difficult in yeast nuclei due to their small size. In general, replication complex studies are challenging in many model organisms due to compaction of DNA within nuclei. By releasing DNA from nuclear confines into chromatin spreads or fibers, individual replicated areas can be assessed with better resolution. DNA fibers, also called combed, extended or stretched DNA, have been extensively used to examine replication dynamics such as origin firing and fork stability (e.g., [1, 2]). Methods to prepare DNA fibers typically purify and isolate DNA free of proteins, and then extend the fibers along treated glass surfaces to produce uniformly aligned, straight and stretched DNA for analysis.

Chromatin fibers, also called spread-fibers or fiber-spreads, are a related but distinct technology from DNA fibers. As with DNA

fibers, chromatin fibers can be used to detect DNA synthesis by adding nucleoside analogs such as 5-bromo-2'-deoxyuridine (BrdU). Distinct from DNA fibers, chromatin fibers retain protein components and epigenetic domains. Chromatin fibers do not require specialized equipment or extensive preparation. They are prepared by lysing whole cells on a slide and allowing the DNA to extend along a glass surface. They retain considerable detail that can be imaged using high-resolution microscopy. Chromatin fibers have been used in multiple organisms to examine origin firing and replication fork progression by examining synthesis alone (e.g., [3, 4]). By adding the dimension of protein detection, chromatin fibers have been used to characterize chromatin domains (e.g., [5–7]), histone modifications relative to DNA synthesis (e.g., [8]), and replication fork components (e.g., [9]).

We describe protocols to prepare chromatin fibers from *S. pombe*, and then detect proteins and study the data with a new analysis algorithm. The basic chromatin fiber method uses cells that have been blocked in early S-phase, and then released into media containing a nucleoside analog to complete replication. Two-color labeling during replication block and release is possible using nucleoside pairs of either 5-iodo-2'-deoxyuridine (IdU) and 5-chloro-2'-deoxyuridine (CldU) (e.g., [10, 11]), or 5-bromo-2'-deoxyuridine (BrdU) and 5-ethynyl-2'-deoxyuridine (EdU) (e.g., [4, 12]). Nucleoside analogs can then be detected by immunofluorescence (IdU, CldU, and BrdU) or click chemistry (EdU) to visualize regions of DNA replication on the spread fibers. Proteins are detected using immunofluorescence and their spatial arrangement can be compared to nucleoside analog incorporation, providing a unique opportunity to assess replication factor distributions. Further, fluorescence in situ hybridization (FISH) (e.g., [13, 14]) can be used to compare protein localization on chromatin fibers relative to specific DNA loci. Together, these methods expand the analysis of DNA replication to consider the size, composition and distribution of replication forks and factors at synthesized and/or specific areas of a genome.

We also describe an analysis algorithm to systematically correlate DNA synthesis with protein location on DNA fibers. Named “One Dimensional Data—Boolean Logic Operations Binning System” (ODD-BLOBS), our analysis algorithm models fiber data as linear distributions, investigating replicated areas and the regions surrounding the ends of replicated tracts. The ends of replicated tracts may be adjacent to replication forks, and increased replication fork proteins might be found at replicated tract tips. ODD-BLOBS quickly analyzes data from many samples to examine protein distribution patterns and frequencies. User-defined changes allow data smoothing to close signal gaps, and alter the size of tract ends. ODD-BLOBS allows users to build a model of protein distribution around replicated regions and presumed fork zones.

---

## 2 Materials

### 2.1 DNA Fibers

1. Analog-incorporating strains must be used to detect DNA synthesis. The strain must have a thymidine kinase (*tk*) gene so that BrdU is converted into a useable nucleoside triphosphate. The kinase is supplied from an integrated *hsv-tk* cassette (e.g., [15]) or on a plasmid (e.g., [10]). Transporters such as the human equilibrative nucleoside transporter 1 (*hENT1*) gene facilitate uptake of the analog and allow more efficient incorporation at much lower concentrations of BrdU in cultures [10, 15, 16].
2. 5-bromo-2'-deoxyuridine (BrdU) solution at 5 mg/mL in sterile water. If the strain has hENT1, BrdU is added at 50 µg/mL; if not 200 µg/mL is required. BrdU stock solution is stored at -20 °C and must be protected from light.
3. Hydroxyurea (HU, FW 76.5 g/mol) solution is made at 1 M in water and stored in the dark at 4 °C up to 1 week. For a reversible early-to-mid S-phase block in *S. pombe*, stock HU is diluted to 10–15 mM (see Note 1).
4. Phosphate buffered saline (PBS 1×) is made with 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> and 1.8 mM KH<sub>2</sub>PO<sub>4</sub> in deionized water. A 10× solution can be made and autoclaved or filter sterilized and stored at room temperature.
5. Zymolyase Mix: 1 M sorbitol, 60 mM EDTA, and 100 mM sodium citrate, pH 6.9–7.0 (see Note 2). Before use add 0.5 mg/mL zymolyase 20T, 1.0 mg/mL lysing enzymes and 100 mM 2-mercaptoethanol. If a large volume is required over time (e.g., time course), prepared zymolyase mix can be kept on ice between time points.
6. Zymolyase 20T (Seikagaku Corporation, Japan) is added to 0.5 mg/mL in zymolyase mix. This enzyme is supplied as a powder, stored at 4 °C, and is added to a master-mix volume of item 5 solution before use.
7. Lysing enzymes (Sigma) are added at 1.0 mg/mL in zymolyase mix to digest cell walls. This dried powder is kept at 4 °C, to be weighed and added to Zymolyase mix (item 5) when needed, similarly to Zymolyase 20T.
8. Lysing solution is made of 50 mM Tris-HCl pH 7.4, 25 mM EDTA, 500 mM sodium chloride, 0.1% Nonidet P-40 (Sigma), and 0.5% (w/v) sodium dodecyl sulfate (SDS). Before lysis, 3 mM 2-mercaptoethanol is added to a suitable volume of lysing solution, and the solution is heated to 70 °C before use.
9. 4% paraformaldehyde (PFA) fixative is obtained as a 16% stock solution, from various manufacturers (e.g., Pierce,

ThermoFisher). It is supplied in 10 mL aliquots in brown glass ampoules. To make 40 mL of a 4% solution, mix entire contents of an ampoule with 30 mL of 1× PBS. Excess 4% PFA solution can be stored at 4 °C up to 1 month.

Alternatively, 4% PFA solution can be made by mixing 2 g PFA powder in 50 mL of phosphate buffered saline and adding 4 M sodium hydroxide added to pH to 7.0–7.5. Carefully heat and stir the solution until all PFA powder is in solution and do not boil. To prevent scorching the PFA, we recommend a double boiler on a hot plate using an exterior beaker with water. Allow the solution to cool and store at 4 °C up to 1 month, or aliquot and freeze at –20 °C indefinitely (*see Note 3*).

10. Poly-L-lysine-treated coverslips. Treated coverslips can be purchased premade. Alternatively, they are easily made from #1.5 glass coverslips (*see Note 4*). Coverslips are first washed in a dilute solution of residue-free detergent (e.g., Sparkleen, Fisher) and then rinsed in several changes of water and then deionized water. The cleaned slides are soaked for 5 min in 100% ethanol and air-dried. Slides are submerged in a 0.1% to 0.01% poly-L-lysine solution (e.g., Sigma P8920) for 5 min at room temperature. Coated slides are baked at 60 °C for approximately 1 h, cooled, and are stored in an airtight container at 4 °C indefinitely.

## 2.2 Immunofluorescence

1. Phosphate buffered saline (1× PBS) as above.
2. 2 N HCl, in water. After dilution from stock HCl, do not store 2 N HCl for longer than 4 months.
3. 0.1 M sodium tetraborate ( $\text{Na}_2\text{B}_4\text{O}_7$ ) solution (pH 8.5).  $\text{Na}_2\text{B}_4\text{O}_7$  is dissolved in water and then pH adjusted to 8.5 with 4 N HCl. Filter-sterilize and store at room temperature, protected from light.
4. Blocking buffer: 10% fetal calf serum, 10% bovine serum albumin, 0.05% Tween 20 detergent (Sigma) are diluted in 1× PBS and filter sterilized. Store at 4 °C up to 1 month.
5. Primary antibody for BrdU is rat BU1/75 (ICR1) (Abcam), diluted to 1:100 in blocking buffer that contains any protein-detecting antibodies (**item 6**).
6. Primary antibody concentration for protein detection is dependent on the tag/protein of interest. All primary antibodies are diluted in blocking buffer, together with the BrdU antibody (**item 5**). We recommend a starting concentration of approximately 10× the usual concentration for western blot detection. For example, if an antibody is used at 1:1000 for a western blot, we would test 1:100 concentration of the same antibody for chromatin fiber immunofluorescence. Alternatively, follow the manufacturer's recommendations for immunofluorescence.



Validated antibodies and concentrations we have tried include: mouse PCNA at 1:100 (PC-10 clone, Santa Cruz Biotechnology), mouse anti-MYC tag at 1:100 (9E10 clone, Covance), mouse anti-HA tag at 1:100 (16B12 clone, Roche), mouse anti-FLAG at 1:100 (M2, Sigma), rabbit anti-Mcm4 at 1:100 [17], and yeast-specific anti-phospho histone H2A at 1:100 (07-745, EMD Millipore).

7. Secondary antibodies against each primary antibody are diluted in blocking buffer together at 1:500 each (2  $\mu$ L each antibody in 1 mL of block). We typically use chicken anti-rat Alexa Fluor 488 (Invitrogen), goat anti-rabbit Alexa Fluor 546 (Invitrogen) and donkey anti-mouse Alexa Fluor 647 (Invitrogen). We recommend changing the detection colors for all epitopes, i.e., goat anti-rabbit 647 and donkey anti-mouse 546 in a repeat of the same experiment and primary antibodies (*see Note 5*). Fluorescent-conjugated antibodies should be used in foil wrapped tubes, incubated in the dark and shielded from light as much as possible to prevent fluorophore degradation.
8. Antifade Mount is 50% glycerol in water with 0.1% PPD (p-phenylenediamine dihydrochloride) and DAPI (4',6-diamidino-2-phenylindole dihydrochloride) at a final concentration of 1  $\mu$ g/mL (*see Note 6*).
9. DAPI stock solution is a 1 mg/mL solution in water, stored at  $-20^{\circ}\text{C}$ . DAPI solutions should be kept in foil-wrapped tubes and shielded from the light as much as possible.
10. VALAP, a 1:1:1 (w/w/w) mixture of petroleum jelly, lanolin and paraffin wax that is used to seal coverslips onto slides (*see Note 7*). VALAP solidifies below  $60^{\circ}\text{C}$  and must be melted and mixed before use. VALAP can be stored in a beaker, melted on a hotplate at low temperature, and applied to seal cover glass onto slides using a thin wood rod.

### **2.3 One Dimensional Data: Boolean Logic Operations and Binning System (ODD-BLOBS) Analysis**

ODD-BLOBS software can be downloaded from [www.SabatinoLab.net/ODD-BLOBS](http://www.SabatinoLab.net/ODD-BLOBS). ODD-BLOBS runs in the Libre Office freeware suite ([www.LibreOffice.org](http://www.LibreOffice.org), or legacy versions of MS Office that support VBA). ImageJ (<http://imagej.nih.gov/ij/>, [18]) or some other image analysis software is required to parse chromatin fiber images into intensity profiles along line segments.

### **2.4 Specialized Equipment**

1. We use  $22 \times 30$  mm, #1.5 cover glasses for fiber preparation, which are easier to manipulate than smaller square cover glass. Cover glasses can be difficult to label and handle, but we find that the image brightness and quality is considerably improved on coverslips compared to slides.

2. Humid chamber for sample incubations. A plastic container with lid (e.g., food storage container) can be used, with damp paper towels on the bottom and a level support on top to hold samples (e.g., microfuge tube rack). The chamber can be wrapped in aluminum foil to protect samples from light.
3. Specialty incubation jars and racks are available that fit  $22 \times 30$  mm cover glasses and are useful in poly-lysine coating, washing, and incubation steps. One supplier of these is Electron Microscopy Supplies. Regularly sized Coplin jars and/or slide-washing baskets may be helpful for washing steps, but are not essential.

### **2.5 Equipment Considerations for Imaging**

Green et al. discuss imaging chromatin fibers and the reader is referred there for additional discussion. Confocal, widefield, and super-resolution modalities should all produce good results when high numerical aperture objectives and appropriate filter sets are employed [19].

---

## **3 Methods**

### **3.1 BrdU Incorporation**

1. Grow cells in an appropriate volume of minimal media to  $1\text{--}5 \times 10^6$  cells/mL (OD595 ~ 0.3–0.6). Add hydroxyurea (HU) to 12 mM, and grow for an additional 3 h at 30 °C to block in early/mid-S-phase. If desired, remove samples of the asynchronous, blocked and experimental samples to confirm HU effect using flow cytometry DNA staining (*see* [20]), and DAPI/aniline blue staining for morphology assessment [19].
2. Remove HU by filtration onto a filter, and washing with 2 volumes of media without drug. Resuspend cells in 1 volume of fresh, prewarmed (30 °C) minimal media.
3. Alternative to **step 2**: remove HU by centrifugation. Pour culture into a sterile 50 mL conical tube and centrifuge at  $500 \times g$  for 5 min. Decant media, and resuspend pelleted cells in 1 volume of fresh media. Centrifuge as above. Repeat a second wash. Resuspend cells in 1 volume of fresh, prewarmed (30 °C) minimal media.
4. Remove an unlabeled sample (10 mL culture), or partition the culture to include a non-BrdU incorporated control released for the same length of time. Add BrdU to 50  $\mu\text{g}/\text{mL}$  and grow for 15–45 min at 30 °C. We suggest a 30-min release in BrdU to start. Under these conditions, cells complete replication by 60 min post-release. Note that high BrdU concentrations and longer incubation times cause cell toxicity (e.g., [16]).

### 3.2 Preparation of Chromatin Fibers

Ahead of time, prepare poly-L-lysine-treated coverslips, zymolyase mix with enzymes and 2-mercaptoethanol, and heat lysis buffer (containing 3 mM 2-mercaptoethanol) to 70 °C. Zymolyase mix with enzymes and 2-mercaptoethanol can be stored on ice until required.

1. Grow cells as appropriate to the experiment, e.g., exponential growth, or during/after temperature or drug treatment. BrdU is added to label new DNA synthesis. Cultures should be in mid-exponential phase (OD<sub>595</sub> ~ 0.3–0.9, approximately  $1\text{--}2 \times 10^7$  cells/mL).
2. Remove 10 mL of culture to a round bottom snap-cap tube. Add sodium azide to 0.2%, mix cells, and incubate on ice for 5 min. Centrifuge at  $500 \times g$  for 5 min in a swinging bucket tabletop centrifuge. Decant supernatant.
3. Wash cells once in 1× PBS, centrifuge as above. Decant supernatant.
4. Wash cells once in 1 mL zymolyase mix without enzymes or 1 mL water, and centrifuge as above.
5. Resuspend the cells in zymolyase mix with 2-mercaptoethanol and enzymes. Incubate at 37 °C for 15 min to spheroplast cells. Check for digestion by phase contrast microscopy, both with and without a drop of 10% SDS. Cells that are digested will be spherical without added SDS. Addition of 10% SDS to the sample will cause spheroplast lysis, leaving broken cells and clear capsules (“ghosts”). Continue spheroplasting cells until >70% of cells in the sample are digested. If buffer pH is correct, this takes 10–15 min, but could require up to 30 min at 37 °C for hard-to-digest samples.
6. Centrifuge cells as above and decant.
7. Add 0.5 mL of 1× PBS to each tube. Resuspend spheroplasts by gently shaking.
8. Pipette 25 μL of the cells onto the short (22 mm) end of a poly-lysine-coated coverslip; a second line (25 μL) can be added, with approximately 2 cm in between the two lines.
9. Let cells settle approximately 10 min at room temperature (cover to protect from light). Tip the cover glass to one side and blot away excess solution. Allow the samples to dry for approximately 5 min, lying flat at room temperature, until the edges are just dry but the center of the sample lines are slightly damp.
10. Pipette 25 μL of hot (70 °C) lysing solution (containing 3 mM 2-mercaptoethanol) onto the prepared and adhered cells. Incubate flat, at room temperature for 3–5 min. Tip the coverslips to a 15° angle and slowly increase to 30°. Use a pipette tip to

channel the lysis solution down the center of the coverslip in the direction of flow. This allows the DNA fibers to stretch out of the lysed cells. Once all liquid has drained, dry samples vertically at room temperature, approximately 3 min (*see Note 8*).

11. Fix fibers in 4% paraformaldehyde for 10 min at room temperature. Alternatively, pipette fixative onto the fibers, sit 10 min, and then tip off PFA solution (*see Note 9*).
12. Rinse slips by dipping into a jar of 1× PBS. Repeat rinse. Blot the excess PBS onto paper towels, and then allow the coverslips to air-dry approximately 5 min.
13. Heat fix fibers onto the coverslips by placing onto a low-temperature heat block at approximately 30–40 °C for 5–10 min.
14. Store fixed fibers in the dark at –20 °C for several weeks.

### 3.3 *Immunofluorescence*

This immunofluorescence protocol detects both labeled DNA synthesis tracts and proteins of interest on the chromatin fibers. Fiber DNA is first denatured to facilitate BrdU–antibody interactions. After washing and blocking, primary antibodies to BrdU and the desired protein epitopes are added. Standard immunofluorescence techniques are used to wash and detect antibodies, and slides are mounted in the presence of DAPI to detect DNA.

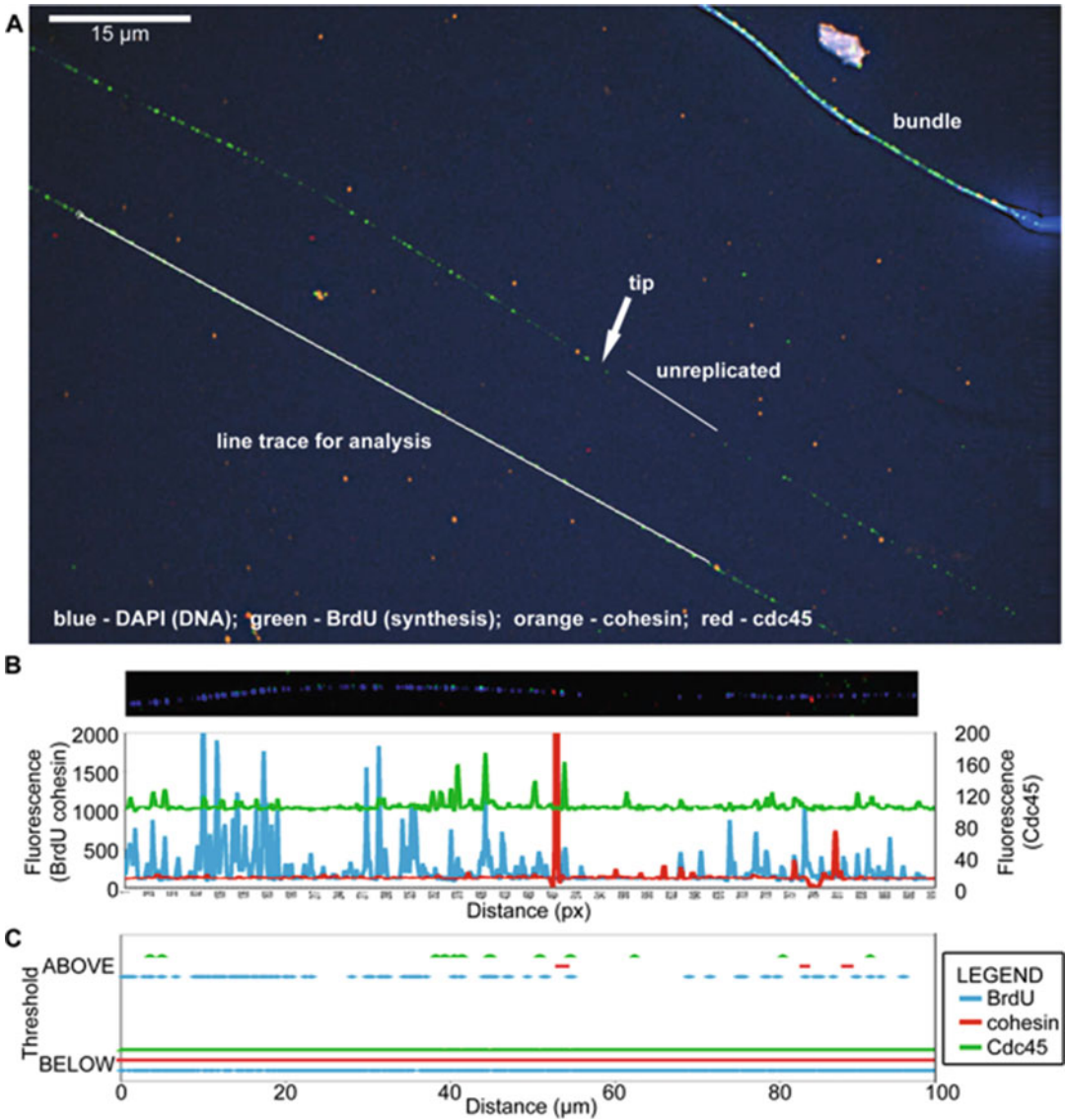
1. Wet samples in 1× PBS. Protect BrdU-incorporated samples from light as much as possible during this and subsequent steps.
2. To denature DNA, immerse samples in 2 N HCl for 15 min at room temperature. Neutralize slides in 0.1 M Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub> (pH 9) for approximately 5 min, and then wash slides three times in 1× PBS.
3. Block slides in blocking buffer. Place coverslips (fiber-side up) into a humid chamber and pipette 100 μL of blocking buffer onto the samples. Distribute and cover with a piece of Parafilm, cut to fit and cover the sample area. Avoid bubbles between the sample and the Parafilm cover (*see Note 10*). Incubate samples in blocking buffer in the dark for 30 min at 37 °C, or for 1 h at room temperature. During this incubation step, dilute primary antibodies into blocking buffer (**step 4**).
4. Remove the Parafilm covers and tip off excess block from each sample. Do not let the samples dry completely. Pipette 100 μL of primary antibody solution onto the samples and cover with Parafilm slips (as in **step 3**). Keep samples in dark, humid chamber.
5. Incubate in primary antibody for 1 h at room temperature, or overnight at 4 °C.

6. Remove Parafilm covers and blot off excess primary antibody. Transfer samples into a Coplin jar with  $1\times$  PBS and wash for 3–5 min at room temperature. Repeat wash three times total. Washes can be performed in Coplin jars or racked and washed in larger containers. Protect samples from light during wash steps. During wash incubations dilute secondary antibodies in blocking buffer for use in **step 7**.
7. After the final wash, wipe excess PBS from the backs of samples, and place in the humid chamber. Pipette 100  $\mu$ L of diluted secondary antibodies onto the samples and cover with a Parafilm coverslip (*see Note 10*).
8. Incubate in secondary antibodies for 1 h at room temperature.
9. Wash samples three times in  $1\times$  PBS, as in **step 6**. If desired, add 0.1% Tween-20 to the first and second washes; this may help to decrease background. The third wash should be  $1\times$  PBS only wash (no Tween 20).
10. Briefly air-dry sample slips in the dark at room temperature until samples are damp but not wet. Do not overdry.
11. Mount samples, fiber-side down, onto glass slides using Antifade Mount with DAPI at 1  $\mu$ g/mL. A commercial alternative is SlowFade gold with DAPI (Invitrogen), or other noncuring mounting media. Seal edges with VALAP or nail varnish.
12. Using DAPI (DNA) fluorescence, look for whole cells and then broken cells with chromatin fibers spilling out of lysed nuclei. Single-stranded DNA antibodies may be used to detect fibers and might provide a better signal for thin fibers.
13. Store slides protected from light at 4  $^{\circ}$ C, or at  $-20^{\circ}$  C.

### 3.4 Imaging DNA Fibers

Parameters related to primary–secondary antibody combinations (exposure times, etc.) must be considered for high-resolution imaging. Negative controls that were not treated with primary antibody are used to determine the baseline fluorescence from nonspecific secondary antibody interactions. By testing no-primary controls, samples exposed to primary antibody will show epitope-specific patterns using the same settings. We also recommend primary antibody dropouts to ensure that two or more primary antibody patterns are not linked to each other.

1. Well-spread Chromatin Fibers are planar, within the axial resolution of conventional and super-resolution visible-light microscopy. Thin, likely single chromatin fibers will have a very weak DAPI signal (*see Fig. 1a*), so much so it will often go unnoticed in a field if some unspread nuclei or unlysed cells are present. It may be necessary to find the fiber-spreads using other color channels, e.g., tracts of BrdU synthesis (*see Note 11*).



**Fig. 1** Example of fiber imaging and line tracing results. **(a)** Merged image of fibers prepared using the described protocol, to detect DNA (DAPI, *blue*), DNA synthesis (BrdU, *green*), cohesin (Rad21-HA, *orange*), Cdc45 (GFP tag, *red*). Bundles are brightly staining and composed of multiple DNA fibers. Single fibers have weaker DAPI signal, and the BrdU signal is frequently punctate in replicated tracts; an *arrow* indicates a replicated tip where a putative DNA replication fork site is hypothesized to occur. An unreplicated area is identified with a line. Shown is a line tracing to acquire raw fluorescent intensity values on all channels with pixel position information. All work was performed in strain FY 3841 (*h<sup>+</sup> rad21-3HA-kanMX6 cdc45-YFP-ura4<sup>+</sup> leu1-32::[hENT1<sup>+</sup> leu1<sup>+</sup>(pJAH29)] his7-366::[hsv-tk<sup>+</sup> his7<sup>+</sup>(pJAH31)] ura4- ade6-210*). **(b)** An area of a sample fiber pseudocolored for BrdU (*blue*), cohesin (*red*) and Cdc45 (*green*), with associated line tracing of raw intensities. Values are plotted for each pixel along the tracing, with BrdU and cohesin scale on the *left axis*, Cdc45 intensity scale on the *right axis*. **(c)** Thresholding was performed for all signals, converting the fluorescent intensities into ON (above threshold value) or OFF (below threshold) for each place along the line. Pixels from **(b)** were converted into distance along the line in  $\mu\text{m}$ , using the system conversion of 0.1092  $\mu\text{m}/\text{px}$  (specific to microscope acquisition parameters)

2. Samples are exposed and data acquired using multicolor z-stacks. While optical sectioning is not obviously required for such thin objects, it is helpful to capture z-chromatic aberration.
3. Multicolor z-stacks are deconvolved in softWoRx with the constrained iterative method and the manufacturer's supplied OTF for the Olympus 60× NA 1.4 PlanApo lens.

### **3.5 Parsing DNA Fiber Data into One Dimension**

ODD-BLOBS must be supplied with a linear array of intensity data for each color channel in the fiber images. This data can be parsed with almost any microscopy image software. In most programs this analysis is called “Line Profiling” or similar, and involves the user tracing a line along the fiber. The line may be straight or curved, but is 1 pixel wide. The output is a trace of intensity values for each color at each pixel along the line object that can be saved to a delimited text file with columns for each color. Automatic fiber detection is nontrivial and falls outside the scope of this article, but the manual process is simple such that naive operators can obtain quantitative amounts of One Dimensional Data from images.

In ImageJ, the Analyze > PlotProfile operation can be performed on any line object drawn on an active window (*see* Fig. 1a). A helpful macro for automating all colors in an image in one operation is ‘Multi color line profile plot’ (Kees Straatman, 2014, <http://www2.le.ac.uk/colleges/medbiopsych/facilities-and-services/cbs/lite/aif/software-1/imagej-macros#Multi%20color%20profile>), which is similarly executed on any line-drawing in an open file.

1. Open file in ImageJ, using the LOCI Bioformats Importer if needed so that channels are properly read in. Locate DAPI-stained fibers. Activate the line tool in the Main Window. Trace a line along a straight segment of fiber. Open the Analyze > Tools > ROI manager and “Add” the trace to the ROI manager. Use the ROI manager tick box “show all” to identify all traces and avoid duplication.
2. Draw additional line segment traces and “Add” them to the ROI manager. When all desired traces are made Save the ROI file under the ROI manager “more” button.
3. On the ROI manager, disable “show all” and select one line segment.
4. Using Plugins > Macros > Run, activate ‘Multi color line profile plot’
5. Export or copy/paste the intensity values to the ODD-BLOBS spreadsheet.

6. Repeat parsing and exporting line segments as required, Additional segments from the same experimental data type may be appended to one ODD-BLOBS sheet if a one (or more) row gap is left between data segments so that tracts across fibers are not artificially concatenated.
7. ODD-BLOBS reserves Column B for pixel number/identity; Column C should be DNA intensity values (e.g., DAPI); Column D should be the marker of replication tracts (e.g., BrdU intensity values); Column E—protein #1 intensity values; Column F—protein #2 intensity values. Columns G to J are reserved for pixel coordinates, if available.

### 3.6 Analyzing DNA Fiber Data

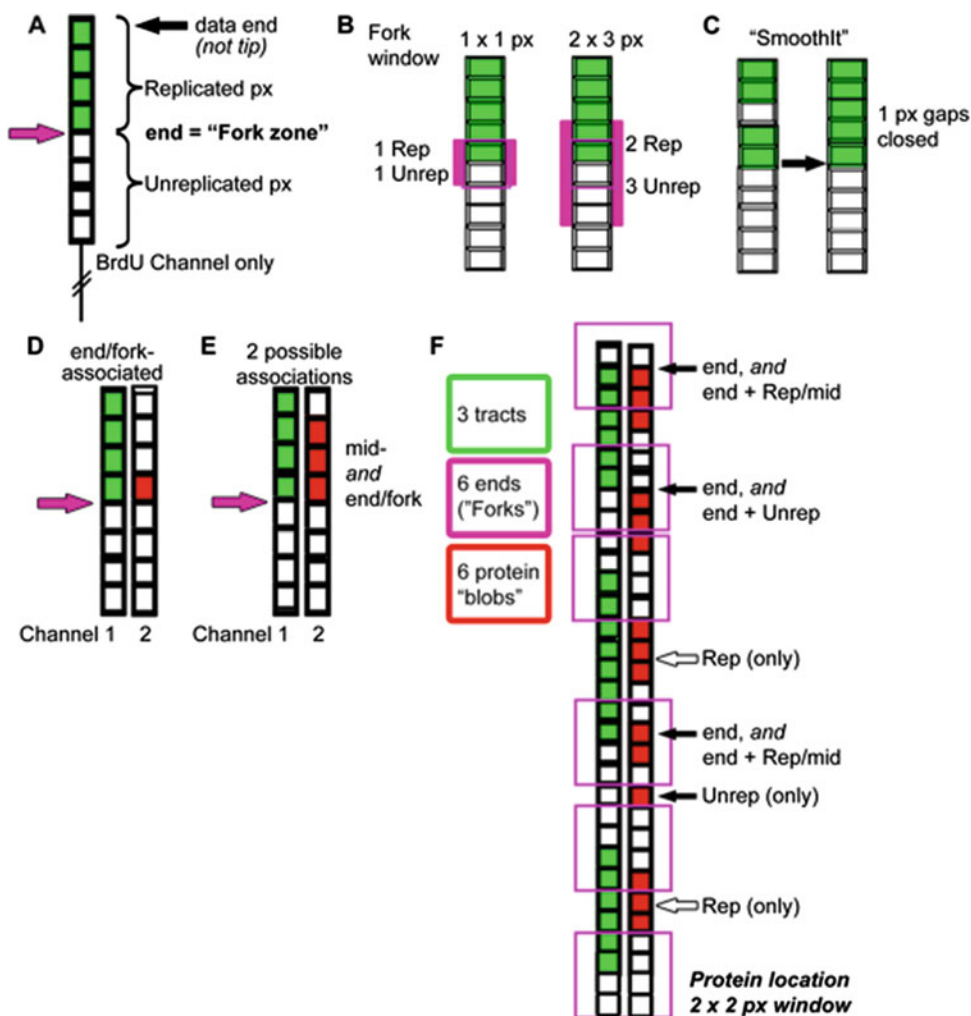
To model DNA replication forks from chromatin fibers, we hypothesize that replication forks are most likely to be found near the tips of replicated tracts. Our analysis examines analog-incorporated DNA (tracts) and associated protein (“blobs”) (*see* Fig. 2). A chromatin fiber is unreplicated (no analog signal), replicated (within an analog region), or tip-proximal (putative replication fork zone). Anywhere in a “replicated” zone might be called the “middle” of that zone, and we do not consider the middle of un-replicated tracts. Thus, replicated tracts have “Middles” and “Ends/Tips”.

The simplest DNA fiber analysis is done using analysis software (e.g., ImageJ) to measure replicated tract lengths. However, associating proteins with replication becomes tedious. Further, defining where a protein is relative to replicated/unreplicated areas and the tips is imprecise without measurement. If the pixel size is known, we can measure and model how much space around a replicated tract tip might be a region wherein a replication fork was active (*see* Fig. 2).

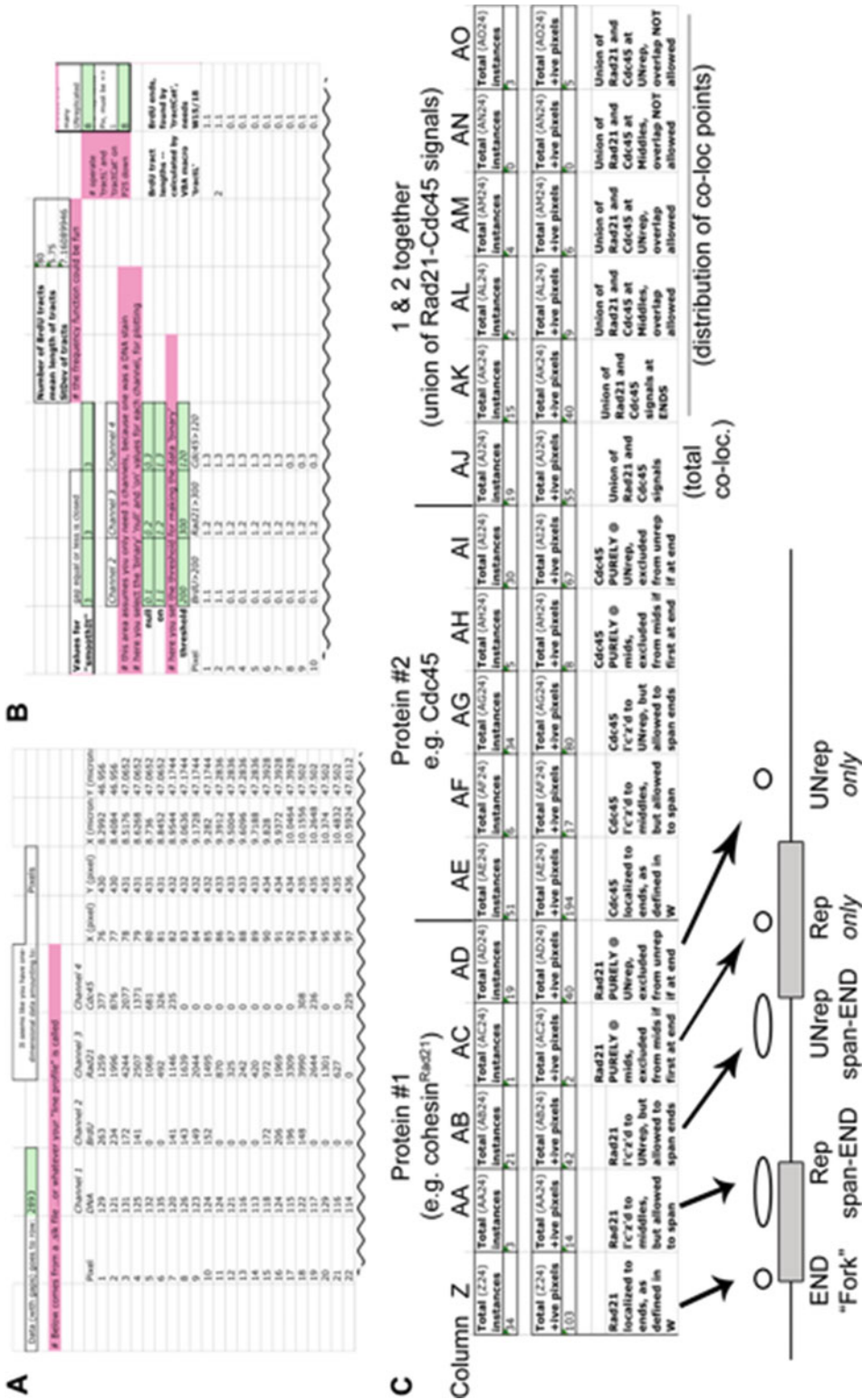
We developed ODD-BLOBS to take linear fiber data, calculate replicated tract lengths, and correlate protein location relative to tract middles or tips/ends. Deconvolved, 2D-projected images are analyzed using a line-drawing tool to acquire fluorescence intensities on all channels at a given point in the image. DNA staining by DAPI or another DNA dye indicates the line that should be traced. Data from all fluorescent channels is saved in a delimited file format (*see* Note 12). The 1D/linear data is imported to LibreOffice and an ODD-BLOBS worksheet. A threshold is set for each channel, converting signal into a binary on or off datum. ODD-BLOBS uses Boolean logic to calculate how many “On” pixel-signals make up a tract-event, and where protein is relative to synthesized tracts (*see* Fig. 2). Additional parameters allow the user to model smoothing to remove gaps in signal, and define a larger or smaller window of “tip” size for the purpose of modeling replication fork structures (*see* Fig. 3).

1. Perform basic labeling of channels in cells C24:F24 and document the housekeeping information in the comments section





**Fig. 2** The logic of ODD-BLOBS. (a) Pixels are represented as *boxes* that are ON (colored, or above threshold signal) or OFF. A group of 4 green (ON) pixels is followed by four uncolored (OFF) pixels, distinguishing a replicated tract from an unreplicated area. Between the replicated and unreplicated pixels is the tip of the replicated tract, which we call the end and putative "Fork zone". (b) "Tip/End windows" are determined as a number of pixels around the tip of the replicated tract. A  $1 \times 1$  pixel (px) window means that 1 pixel on each side of the tip encompasses the putative "Fork zone". This  $1 \times 1$  px window is the minimum for ODD-BLOBS analysis. A  $2 \times 3$  px window here means that two replicated pixels and three unreplicated pixels on either side of the end of the tract define the window of the putative "Fork zone". (c) In the case where small, nonresolvable gaps in signal occur (see **Note 15**) the "SmoothIt" algorithm closes the gaps. SmoothIt is user-defined (i.e., "n pixels") to close gaps of less than or equal to n. (d) A 1 pixel protein ("blob") within the end window, spans only the end and is called purely End or "Fork-zone" associated. Note that DNA synthesis tracts are determined on Channel 1 (green), and protein on a second Channel 2. (e) A larger protein blob may span both the End and Replicated regions. (f) A complex example of 3 tracts (green) and associated protein blobs (red). The six tract ends are defined as putative "Fork-zones". There are six identified protein blobs, under a  $2 \times 2$  pixel end-window definition. These include End only; End spanning to Replicated (end + Rep/mid-tract); End spanning to Unreplicated (end + Unrep); Replicated only (Rep only); Unreplicated only (Unrep only)



**Fig. 3** Sample ODD-BLOBS input and output. **(a)** Data from line traces on all channels, with positional information, are input into columns B to J. **(b)** User defined analysis parameters include thresholds and “values for Smoothit” defining the gap size to be closed for each channel. These may be different for each channel. “Smoothit” closes a gap equal to or less than the defined number, but does not have to be used. The number, average size, and standard deviation of replicated tracks are reported in the middle of this area. The “window size” of ends is set at right, and must be 1 or greater. **(c)** Data output of number of instances of different kinds of events and number of pixels comprising those events

C11:C12 such as epitope identity (*see Note 13*) (example, Fig. 3a).

2. Select a cutoff threshold for each channel's fluorescence intensity that determines presence or absence of signal in cells P23:R23 (example Fig. 3b, right side).
3. In cell W15, tell ODD-BLOBS how many unreplicated pixels proximal to the end of a replicated tract should be considered part of the "tip" (*see Note 14*) (Fig. 3b, right top).
4. In cell W18, tell ODD-BLOBS how many replicated pixels proximal to the end of a replicated tract should be considered part of the "tip" (Fig. 3b, right bottom). The number of pixels on either side of a tract end, **steps 3** and **4**, can be independently set and becomes a modeling parameter.
5. Optionally, use cells P15:R15 to tell ODD-BLOBS for each channel to close/smooth gaps in signal equal to or less than a given size. This smoothing parameter is provided to smooth local artifacts that could be introduced by binary thresholding (*see Notes 14* and *15*). The number of pixels to smooth in "SmoothIt" is designated for each signal (BrdU, proteins) independently (Fig. 3b), and becomes a modeling parameter.
6. On the ODD-BLOBS spreadsheet call the "SetUpSheet" macro with the supplied button. This macro performs some basic layout procedures to enumerate the linear data and adjust cell reference formulas to accommodate the number of rows of data the user has provided.
7. Optionally call the "SmoothIt" macro using the button provided.
8. Call the "FillFiberSheet" macro using the button provided.
9. ODD-BLOBS returns counts of protein tracts for all proteins individually, relative to tracts (Fig. 3c). The first five columns (Z to AD) are for above-threshold detections of Protein #1 (here Rad21, cohesin) along DAPI signal. The second set of five columns (AE to AI) describes Protein #2. The last six columns (AJ to AO) describe the total instances of Protein 1 and 2 colocalizations (column AJ), and their distributions relative to replicated tracts (AK to AO).
10. Distributions of proteins can be calculated relative to the total number of BrdU tracts (found in Columns U and AU). A parameter "Number of tracts decorated with protein" is described relative to END/"Fork", Replicated, or Replicated-spanning END associations.
11. The total number of positive pixels is calculated below the number of instances of association (Fig. 3c). This may allow modeling if a protein complex is asymmetrically located or sized in replicated/unreplicated areas.

12. Protein association around tract ends may extend into Replicated or Unreplicated areas. If a protein blob is only found in a Replicated or Unreplicated area it is described independently (Fig. 3c, Columns AC/AD; AH/AI). If the protein at a tip extends past the defined tip window into purely Replicated or Unreplicated areas (defined in **steps 3** and **4**) it is independently counted (Fig. 3c, Columns AA/AB; AF/AG). These parameters may be useful in cases where proteins are associated at the fork and spread outward such as single stranded DNA.
13. If the user wishes to rerun ODD-BLOBS with different settings in **steps 2–5** both **steps 7** and **8** should be rerun in order. We recommend having a separate spreadsheet tab for separate parameters.

---

## 4 Notes

1. Hydroxyurea is toxic. Handle with care and dispose of waste HU under appropriate safety guidelines.
2. Zymolyase activity is pH dependent. An acidic pH is required for efficient enzyme activity. If neutral or basic buffers are used zymolyase digestion requires significantly more time.
3. Paraformaldehyde is toxic in powder and liquid forms. Wear appropriate safety gear, work in a fume hood if possible and dispose of reagent and waste with appropriate caution and under regional safety guidelines.
4. #1.5 cover glass has an average thickness of 0.17  $\mu\text{m}$ , which is the thickness that most microscope objectives assume and correct for.
5. “Fluor swap” changes the secondary antibodies and colors used to detect primary antibodies. For example, if chicken anti-mouse Alexa Fluor 488 and donkey anti-rabbit Alexa Fluor 546 were used in one experiment, the two colors could be reversed (anti-mouse Alexa Fluor 546 and anti-rabbit Alexa Fluor 488). Additionally, the host species of the secondary antibody can be changed. Altering the batch and color of secondary antibodies helps to identify and eliminate spurious interactions and patterns.
6. 50% glycerol is used for the samples described in these protocols, where samples are mostly dry. Alternatively, mount with 90% glycerol can be used.
7. VALAP is an excellent choice for sealing microscope samples, since it is nonreactive and nontoxic. One preparation, made in a small glass beaker with 10–20 g of each component, will last a year or longer with moderate to heavy use. A thin bead of melted VALAP is sufficient to completely seal around the coverslip.

8. An alternative lysis method is to immerse the coverslip into a Coplin jar or conical tube containing the 70 °C lysis solution. Incubate the coverslip in the lysis solution for 30 s, and then remove slowly and vertically. To maintain temperature while lysing multiple samples, the lysing solution can be kept in a beaker with 70 °C water.
9. An alternative fixative for these samples is 30% methanol with 10% formaldehyde, although we prefer 4% paraformaldehyde.
10. Carefully pipette the blocking buffer onto the cover glass sample side up, and avoid adding bubbles. Place the Parafilm cover onto the blocking buffer to exclude bubbles, which may interfere with sample contact with blocking buffer and antibodies.
11. Finding fibers using antibody-dependent channels carries a risk of biased data collection toward epitope-positive events (i.e., only BrdU-incorporated DNA, as opposed to total DNA). This is problematic if the percentage of replicated fibers is desired, and is less of a problem if patterns are being examined. To detect all DNA in an unbiased manner, DNA-antibody methods are recommended.
12. Some image analysis programs save data as text and not numbers (e.g., \*.slk files generated in softWorx Line Tool). To convert text information into numbers in Libre Office, open the \*.slk file and select the affected columns of data to set the cell format as “Number”. Choose Edit → Find & Replace, and enter  $^[0-9]$  in the “Search for” box. In the Replace area, enter “&”. Check on “Regular Expressions” and “Current selection only”, and then click Replace All. These steps convert the data into numbers in the spreadsheet.
13. Copy and paste data into the ODD-BLOBS spreadsheet, with the following parameters: Column B—pixel number/identity; Column C—DNA intensity values; Column D—synthesis/BrdU intensity values, collected at 528 nm (green) emission; Column E—protein #1 intensity values, collected at 617 nm (orange) emission; Column F—protein #2 intensity values, collected at 675 nm (far red) emission; Columns G to J—pixel coordinates.
14. According to the Abbe limit of  $0.61(\lambda/NA_{obj})$  in visible epifluorescent microscopy, two objects will require approximately 200 nm separation to be resolved as distinct, using 488 nm excitation and a state-of-the-art objective (NA 1.4).
15. “SmoothIt” closes gaps in signals. Assuming that the system is sampling at the Nyquist rate, it is not possible to resolve a single pixel gap. Pixel gaps might be introduced by thresholding in ODD-BLOBS analysis. SmoothIt can eliminate pixel gaps that are artifacts of binary thresholding. The number of pixels smoothed can be set and used to model tract parameters.

## Acknowledgments

We wish to thank Susan Forsburg (University of Southern California, USA) for her support of this project and for providing strain FY3841. The Sabatinos lab is supported by NSERC DG RGPIN/04405-2015.

## References

1. Patel PK, Arcangioli B, Baker SP, Bensimon A, Rhind N (2006) DNA replication origins fire stochastically in fission yeast. *Mol Biol Cell* 17 (1):308–316
2. Scorah J, McGowan CH (2009) Claspin and Chk1 regulate replication fork stability by different mechanisms. *Cell Cycle* 8 (7):1036–1043
3. Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J Cell Biol* 140 (6):1285–1295
4. Sabatinos SA, Green MD, Forsburg SL (2012) Continued DNA synthesis in replication checkpoint mutants leads to fork collapse. *Mol Cell Biol* 32(24):4986–4997
5. Sullivan BA, Karpen GH (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol* 11 (11):1076–1083
6. Blower MD, Sullivan BA, Karpen GH (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev Cell* 2 (3):319–330
7. Ross JE, Woodlief KS, Sullivan BA (2016) Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. *Epigenetics Chromatin* 9:20
8. Bailis JM, Luche DD, Hunter T, Forsburg SL (2008) Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. *Mol Cell Biol* 28(5):1724–1738
9. Cohen SM, Chastain PD 2nd, Cordeiro-Stone M, Kaufman DG (2009) DNA replication and the GINS complex: localization on extended chromatin fibers. *Epigenetics Chromatin* 2 (1):6
10. Sivakumar S, Porter-Goff M, Patel PK, Benoit K, Rhind N (2004) In vivo labeling of fission yeast DNA with thymidine and thymidine analogs. *Methods* 33(3):213–219
11. Luke-Glaser S, Luke B, Grossi S, Constantinou A (2010) FANCM regulates DNA chain elongation and is stabilized by S-phase checkpoint signalling. *EMBO J* 29(4):795–805
12. Bradford JA and Clarke ST (2011) Dual-pulse labeling using 5-ethynyl-2'-deoxyuridine (EdU) and 5-bromo-2'-deoxyuridine (BrdU) in flow cytometry. *Curr Protoc Cytom Chapter 7 Unit 7 38*
13. Sullivan BA (2010) Optical mapping of protein-DNA complexes on chromatin fibers. *Methods Mol Biol* 659:99–115
14. Haaf T, Ward DC (1994) Structural analysis of alpha-satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum Mol Genet* 3(5):697–709
15. Hodson JA, Bailis JM, Forsburg SL (2003) Efficient labeling of fission yeast *Schizosaccharomyces pombe* with thymidine and BUdR. *Nucleic Acids Res* 31(21):e134
16. Sabatinos SA, Mastro TL, Green MD, Forsburg SL (2013) A mammalian-like DNA damage response of fission yeast to nucleoside analogs. *Genetics* 193(1):143–157
17. Forsburg SL, Sherman DA, Otilie S, Yasuda JR, Hodson JA (1997) Mutational analysis of Cdc19p, a *Schizosaccharomyces pombe* MCM protein. *Genetics* 147(3):1025–1041
18. Rasband WS. ImageJ. U.S. National Institutes of Health, Bethesda, Maryland, USA, <http://imagej.nih.gov/ij/>. 1997–2011
19. Green MD, Sabatinos SA, Forsburg SL (2015) Microscopy techniques to examine DNA replication in fission yeast. *Methods Mol Biol* 1300:13–41
20. Sabatinos SA, Forsburg SL (2015) Measuring DNA content by flow cytometry in fission yeast. *Methods Mol Biol* 1300:79–97



## Long-Term Imaging of DNA Damage and Cell Cycle Progression in Budding Yeast Using Spinning Disk Confocal Microscopy

Riccardo Montecchi and Etienne Schwob

### Abstract

Live cell imaging can monitor biological processes in time and space by providing quantitative measurements of cell behavior on a single-cell basis and in live conditions. However the illumination required to visualize fluorescently tagged endogenous proteins often perturbs cellular physiology, a problem particularly acute for yeast cells that are small, highly photosensitive and with scarce protein content. Analyzing the activation of the DNA damage response (DDR) in various yeast mutants or growth conditions, as well as its consequences for cell cycle progression and cell viability over extended periods of time therefore requires a special microscopy setup that does not by itself create DNA damage or perturb cell growth. Here, we provide a quick guide, strains and advice for imaging the DDR in *S. cerevisiae* for extended time (3–12 h) using spinning-disk confocal microscopy in conditions of limited photobleaching and photodamage. DDR is a conserved mechanism that allows the cell to respond to various stresses, especially those altering DNA integrity or topology. Acquiring time-lapse images of the DDR at high temporal and spatial resolution is of great interest, in particular when studying the effects of mutations or drugs which compromise genomic stability and cell cycle progression.

**Key words** *S. cerevisiae*, Yeast, Spinning disk microscopy, Yokogama CSU-X1, DNA damage response, Rad52-GFP, mCherry-Tub1, Microscopy, Recombination foci, Phototoxicity

---

### 1 Introduction

Live cell imaging is a powerful method that provides four-dimensional information by correlating spatial three-dimensional ( $x, y, z$ ) acquisition to time ( $t$ ). This contributed to reveal the dynamics of numerous biological and biochemical events, the nature of which was previously obscure [1–3]. However time-lapse microscopy still remains an invasive approach. Although cells have evolved protection mechanisms that allow them to handle

---

**Electronic Supplementary Material:** The online version of this chapter (doi: [10.1007/978-1-4939-7306-4\\_35](https://doi.org/10.1007/978-1-4939-7306-4_35)) contains supplementary material, which is available to authorized users.

stressful situations, transmitted-light microscopy exposes the sample to overheating, drying, and photodamage. The main cause of photodamage is the accumulation of light-induced reactive oxygen species (ROS), such as singlet oxygen that is extremely reactive to organic molecules, including those forming biological structures. This cellular stress will either be undetected when cells are imaged for a short period of time or remain under-appreciated when cells are seen dividing, but becomes obvious when cells are imaged over several generations [4, 5]. Many other factors besides illumination, such as temperature, humidity, cell cycle stage and the presence of intracellular chromophores or deleterious mutations will cumulatively increase the stress that cells experience during time-lapse imaging.

Researchers therefore constantly try to develop solutions to reduce the fluorescence excitation light delivered to cells during time-lapse acquisition, while trying to maintain a sufficient signal-to-noise ratio for each monitored fluorescent protein. One such solution is spinning disk microscopy where the incident excitation light is temporally and spatially segmented at high frequency so to avoid continuous sample illumination. When combined with a sensitive EM-CCD camera, this method allows imaging of live cells over long periods of time under conditions of low photobleaching and photodamage [6, 7]. Yeast cells are particularly photosensitive during live cell imaging due to their small size; low protein content; and the presence of intracellular chromophores that absorb light and generate ROS. For instance, the light irradiance that yeast cells receive during fluorescence microscopy with a 100x objective can be  $10^3$  to  $10^6$  more intense than direct sunlight [8], a dose yeast cells are certainly not equipped to deal with. Indeed Sedat and colleagues have calculated that the irradiance a yeast can tolerate without adverse long-term effects is  $\sim 0.5 \mu\text{W}/\mu\text{m}^2\cdot\text{s}$ , about  $100\times$  less than the dose given during traditional wide-field fluorescence microscopy [4]. Another problem is that microscopy laser output varies with time due to misalignment, acousto-optical tunable filter (AOTF) overheating, etc., which means that the same illumination settings will deliver different light intensities to the sample during different imaging sessions. To deal with this issue it is recommended to record and reset the actual light intensity delivered by the system using a slide power meter (e.g., Thorlabs S170C). However the maximum light irradiance a cell can cope with also depends on its genetic make up and growth condition, which calls for using a cellular proxy for photodamage.

In this chapter we provide a pragmatic solution to quickly assess photodamage during time-lapse acquisition of yeast cells within each user's environment, and a guide to perform dual fluorescence imaging for several cell generations of *S. cerevisiae* using spinning disk confocal microscopy (Yokogawa CSU-X1) in conditions that minimize the stress caused to cells. Cells are maintained immobile



and viable for at least 12 h under a thin layer of agarose in synthetic complete medium poured in glass-bottom FluoroDishes. Although cell growth and doubling time are still the most common criteria used to assess phototoxicity during time-lapse acquisition [4, 9], we noticed when using a reporter of DNA repair and recombination (Rad52-GFP) that even fast dividing cells can experience important stress. Thus we suggest monitoring the DNA damage response (DDR) as an additional sensor for microscopy-induced photodamage. Indeed, it is well known that ROS accumulation leads to a variety of damage to proteins and DNA [10, 11], which can lead to a strong activation of the DDR. Here we use a diploid yeast strain expressing GFP-tagged Rad52 from its endogenous locus that produces bright nuclear foci upon DNA damage and DSB repair [12]. The strain also carries an ectopic copy of *mCherry-TUB1* to label microtubules and monitor cell cycle progression [13]. Alternative read-outs for DNA damage, such as Rfa1-GFP, Ddc1-GFP or Mre11-GFP can also be used [14]. Wild-type strains and various mutants were observed for up to 12 h or eight generations with little or no photobleaching or photodamage using an Andor CSU-X1 spinning disk microscope equipped with an EM-CCD camera. The strain expressing Rad52-GFP is useful for everyone to rapidly determine the highest illumination condition that does not cause overt photodamage to cells in their own microscopy setup. Long-term cell imaging using spinning disk microscopy is valuable for determining the duration of the DDR and its consequences for cell cycle progression, cell division, and the fate of progeny. Protocols for determining the position, colocalization and mobility of DNA double-strand breaks (DSB) have been described recently elsewhere [15].

---

## 2 Materials

### 2.1 Yeast Culture and Sample Preparation

1. Yeast strains.
  - E3416: *MATa/α*, *RAD52-GFP/RAD52-GFP*, *URA3::mCherry-TUB1/URA3::mCherry-TUB1* (W303)
  - E3410: *MATa/α*, *cdc6-1/cdc6-1*, *RAD52-GFP/RAD52-GFP*, *URA3::mChe-TUB1/URA3::mChe-TUB1*
  - E5357: *MATa/α*, *mec1Δ/mec1Δ*, *sm11Δ/sm11Δ*, *RAD52-GFP/RAD52-GFP*, *URA3::mCherry-TUB1/URA3::mCherry-TUB1*
2. SC-D medium (synthetic complete + dextrose): 2 g/L Yeast nitrogen base (w/o amino acids, w/o NH<sub>4</sub>SO<sub>4</sub>), 5 g/L NH<sub>4</sub>SO<sub>4</sub>, 20 mL/L 50× AAA mix (10 g/L threonine; 5 g/L lysine, leucine, tryptophan, and phenylalanine; 3 g/L isoleucine and methionine; 2.5 g/L histidine, adenine, and uracil;

2 g/L arginine), buffered with 10 mM Tris-HCl pH 7.5 and 2% dextrose (*see Note 1*).

3. 1% agarose (e.g., UltraPure, InVitrogen) in SC-D medium.
4. Tissue culture dish with glass bottom. 35 mm dish diameter, 0.17 mm glass thickness, non-pyrogenic, gamma sterilized (e.g., FluoroDish, World Precision Instrument).

## 2.2 Image Acquisition

A number of suitable spinning disk confocal microscopes exist. We describe the equipment used in our group.

1. Andor Nikon Spinning disk microscope CSU-X1, equipped with 488 nm (60 mW), 561 nm (50 mW) solid state laser lines and EM-CCD camera (Andor iXon Ultra).
2. Nikon 100×/1.45 oil objective with objective heater and controller unit (Biotech).
3. Environmental chamber and control (Okolab H301-T-UNIT-BL-PLUS).
4. Acquisition software (Andor IQ3).

## 2.3 Image Processing

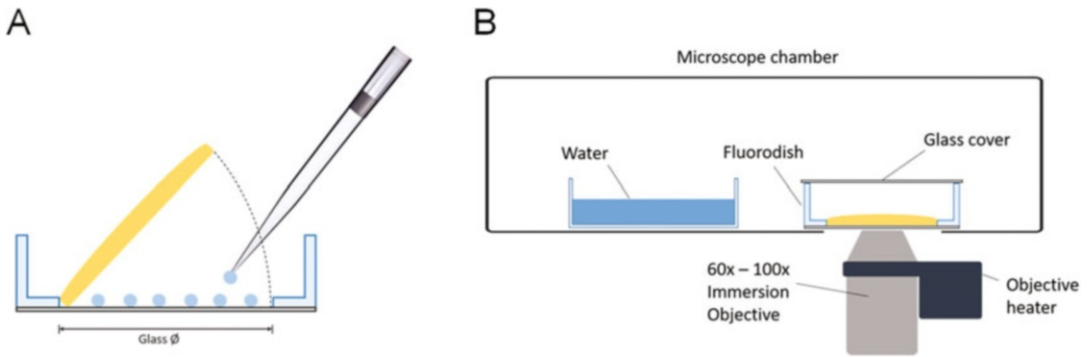
1. Fiji Is Just ImageJ software [16].
2. Pure Denoise plug-in for Fiji (<http://bigwww.epfl.ch/algorithms/denoise/>).

---

## 3 Methods

### 3.1 Yeast Culture and Sample Preparation

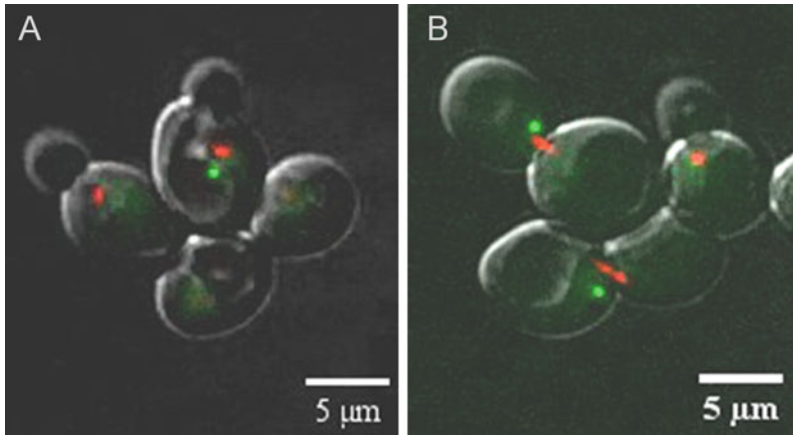
1. Grow cells at 30 °C in buffered SC-D to a concentration of 10<sup>7</sup> cells/mL.
2. Centrifuge 1 mL cells for 4 min at 5000 rpm (2400 × *g*) and resuspend in 100 μL SC-D.
3. Melt an aliquot of 1% agarose in SC-D by microwaving (*see Note 2*).
4. Pour 800 μL of the melted medium and spread it evenly on the bottom glass of a FluoroDish (*see Notes 3 and 4*). Let the agarose layer harden 5 min at 4 °C with lid open. Then take it back to RT (*see Note 5*).
5. Gently lift up the agarose layer with a 100 μL pipetman tip held in one hand.
6. With the other hand deposit 8 μL of the sample in 8–10 small drops on the glass surface of the FluoroDish (Fig. 1a).
7. Gently put the agarose layer back in place, making sure that no air bubble gets trapped.



**Fig. 1** Imaging chamber schematic. **(a)** The agarose-SCD pad (*yellow*) is lifted and several droplets of yeast cells are deposited on the glass bottom of a FluoroDish. The agarose pad is then gently put back in place so that cells are immobilized and growing in monolayer. **(b)** The FluoroDish is clamped on the plate holder within a closed imaging chamber set at the right temperature and containing a water reservoir for humidification. Covering the FluoroDish with glass improves transmitted light images. Objective heater is required to maintain the sample at the set temperature

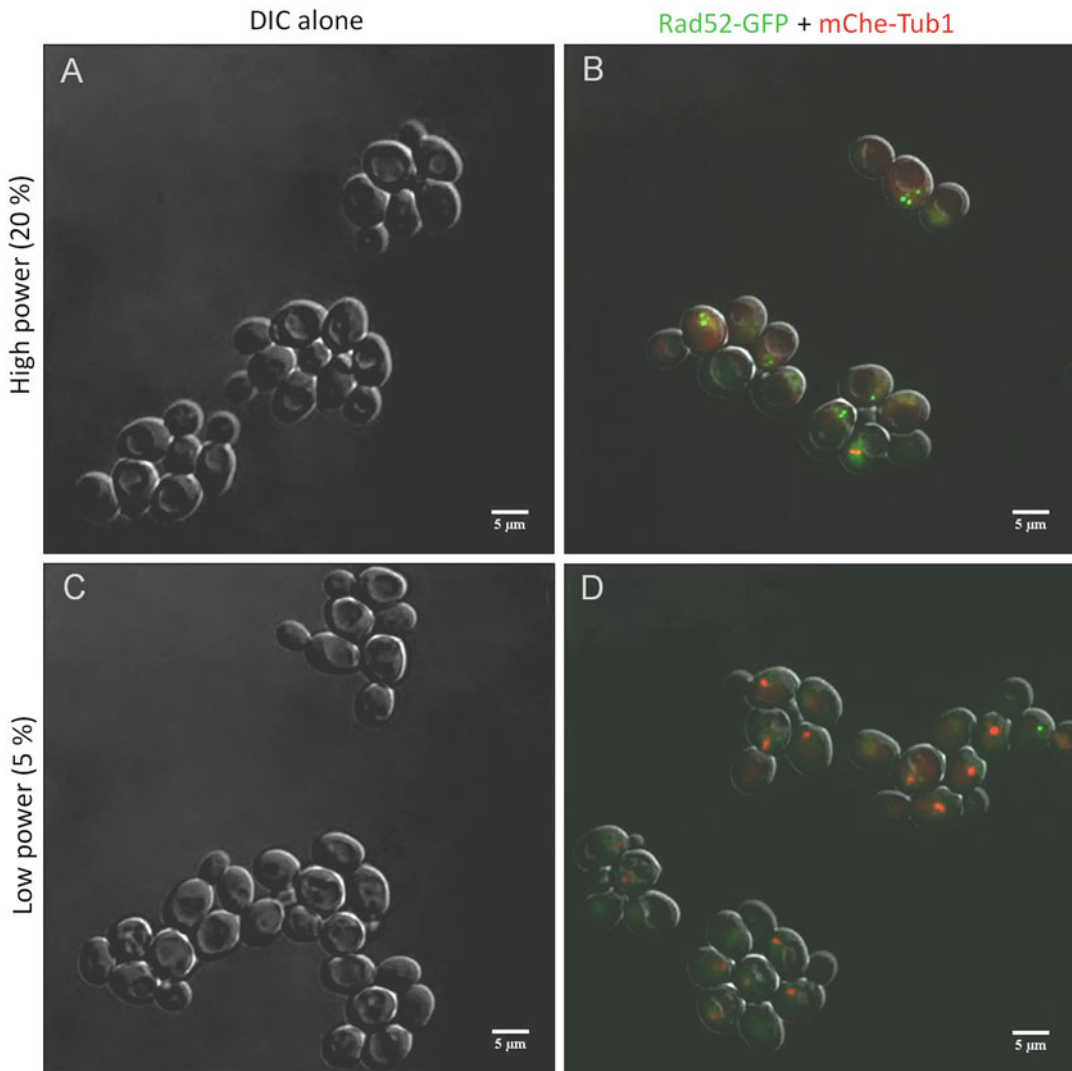
### 3.2 Image Acquisition

1. Set the temperature of the chamber and objective heater to an effective temperature of 30 °C on the sample, at least 30 min before acquisition (*see Note 6*).
2. Place the FluoroDish in a humidified chamber insert in order to prevent drying (Fig. 1b).
3. Focus on the cells using the transmitted light mode and set the Perfect Focus System (PFS) to keep cells in focus during the entire duration of the movie.
4. For each field of view, program the acquisition loop to take one image with Differential Interference Contrast (DIC) and 11 z-stacks of 0.67  $\mu\text{m}$  step-size in the GFP and mCherry channels (*see Note 7*). It is highly recommended that a DIC-alone field-of-view is taken as a negative control for laser-induced photodamage.
5. When using the *RAD52-GFP*, *mCherry-TUB1* tester strain (E3416), set the green and red laser power at the minimum intensity (mW) necessary to barely see the nucleoplasmic Rad52-GFP signal (soluble pool) and the mCherry-Tub1 labeled mitotic spindle (*see Note 8*).
6. Perform a 4–12 h acquisition by taking one image (11 z-stacks) every 1 or 2 min.
7. Process the images using SUM slices or MAX projection for optimal signal intensity.
8. Use the Pure Denoise plug-in for Fiji to denoise images and increase the signal to noise ratio.



**Fig. 2** Persistent Rad52-GFP foci in large budded cells as a marker of photodamage and/or DDR activation. **(a)** Wild-type diploid yeast cells (E3416) grown and imaged at 30 °C for 6 h (shown frame is at 96 min) using a spinning disk confocal microscope under low illumination conditions. *Green*, Rad52; *red*, tubulin. Only a fraction of the cells in S phase (small bud, no spindle) contain Rad52 foci. **(b)** *cdc6-1* temperature-sensitive cells (E3410) grown and imaged in the same conditions show persistent Rad52 foci in the majority of metaphase-arrested cells (large bud, short spindle). Shown frame is at 146 min. A similar phenotype is seen in WT cells exposed to higher laser illumination (see Movie 2)

9. Rad52 forms nuclear foci in a fraction of S-phase cells [12], but these foci are faint, transient (2–15 min) and appear in small budded cells (Fig. 2a).
  - The presence of persistent Rad52-GFP foci (>30 min) in large budded cells with short spindle is a sign of important photodamage (Fig. 2b).
  - In unstressed cells the mitotic spindle remains in its metaphase state (1–2 μm) only for a short time before undergoing rapid extension during anaphase (see Movie 1). A stressed cell will show a slightly elongated metaphase spindle present at the bud neck for considerably longer time before engaging anaphase (see Movie 2).
  - Large, bright, and long-lasting Rad52-GFP foci gradually accumulate in number over time in cells experiencing photodamage, because photodamage is cumulative.
  - Cells blocked in G2 or metaphase will have a bud that keeps growing in size, whereas dead cells will stop increasing their cell mass. After many hours of growth without division, cells will eventually burst.
  - Some damaged cells can undergo mitosis and divide in spite of the presence of a bright Rad52-GFP focus, indicating that cell division is not a good indicator for the absence of stress.



**Fig. 3** Comparison of DIC alone and laser-illuminated field of views reveals long-term toxicity of high light irradiance. (**a, b**) Wild-type yeast cells (E3416) were imaged every 2 min for 4 h at 30 °C on a spinning-disk confocal microscope either without (**a**) or with  $11 \times 300$  ms exposure with 488 nm light (GFP) and 561 nm light (mCherry), both set at 20% laser power (**b**). The right panel shows that cells receiving laser light contain Rad52 foci and divide less rapidly than those exposed to transmitted light only. (**c, d**) Cells were treated identically, except that laser power was set at 5%. Cells divided equally well regardless of laser light exposure and contained little or no Rad52 foci, indicating that these conditions are adequate for live cell imaging with limited photodamage

- At the end of the movie (or after 6–12 h of growth) compare the number of cell divisions in the DIC-only field of view to that in the laser-illuminated regions (Fig. 3). If cell density is roughly equal then the conditions of imaging are adequate with little or no induced photodamage.

11. Using these conditions it is then possible to study the effect of drugs (e.g., Movie 3, WT + camptothecin) or mutants (Movie 4, *mec1Δ* + camptothecin) on the activation of the DDR and cell cycle progression. However, keep in mind that drugs and mutations can have additive effects with the light-induced photo-damage, and may show a different response in non-illuminated conditions. Conducting experiments in parallel with live or fixed cells that are not permanently exposed to strong light is therefore recommended.

---

## 4 Notes

1. Use SC-D medium for better signal–noise ratio, avoiding the high autofluorescence of cells grown in rich medium (YPD). Keeping the pH near neutrality (pH 7.5) improves GFP stability [17].
2. Prepare a solution of 1% agarose in SC-D beforehand, and keep 1 mL aliquots at 4 °C. Briefly microwave to melt before pouring. If drugs are to be added to the medium, let it first cool down to ~50 °C.
3. Start pouring from the border between glass and plastic; this helps even distribution of the medium. A small agarose bump will not affect the acquisition, as light is collected from the bottom.
4. We found FluoroDish with agarose top layer to be good solution for long and short acquisition times of budding yeast cell division. Yeast cells are more immobile than on concanavalin-treated glass, while growing for a long time in two dimensions, thus facilitating pedigree analysis. Avoid acquiring fields that are close to the edge of the agarose layer, however, as it might shrink and move slightly due to drying. For long-term acquisition, place wet paper pads or water drops near the FluoroDish edge.
5. Since cold is known to de-polymerize microtubules, wait for the dish surface to reach room temperature before loading the sample. Short-term storage of the FluoroDish-Agarose-SCD is possible at 4 °C, if sealed with Parafilm to prevent drying.
6. Immersion objectives are heat sinks for the sample in contact with the lens. Effective temperature under the objective can be 4 °C lower than the set temperature of the microscope environmental chamber. An objective heater should be placed at least 30 min before the acquisition in order to reach the expected temperature. Temperature sensors placed near the

sample (below the agar pad) should be used to set the correct temperature before starting the experiment.

7. In our setup, we found QUAD dichroic emission filters (440/40, 521/21, 607/34 and 700/45 nm) produced images with better signal to noise ratio.
8. By pressing autoscale during acquisition, it should be possible to see the nucleoplasmic Rad52-GFP and the mitotic spindle in the mCherry channel. If not, increase the laser intensity. Performing short acquisitions (2–3 frames), and immediately checking the final result with ImageJ (SUM slice or MAX projection) can also help setting the correct laser intensity. Change field of view after setting the parameters.

---

## Acknowledgments

We thank D. Tamborini and S. Piatti (CRBM, Montpellier) for the initial set-up of yeast live cell imaging, V. Georget (BioCampus MRI Imaging Facility) for endless support, as well as R. Rothstein (Columbia U., New York) and E. Schiebel (ZMBH, Heidelberg) for providing yeast strains and constructs. RM was a Master student from University Milan-Bicocca (Italy) holding an Erasmus+ stipend from the EU. ES thanks Association pour la Recherche contre le Cancer (ARC) and Agence Nationale de la Recherche (ANR-14-CE11-0007-03) for funding.

## References

1. di Pietro F, Echard A, Morin X (2016) Regulation of mitotic spindle orientation: an integrated view. *EMBO Rep* 17:1106–1130
2. Park S, Greco V, Cockburn K (2016) Live imaging of stem cells: answering old questions and raising new ones. *Curr Opin Cell Biol* 43:30–37
3. Miné-Hattab J, Rothstein R (2013) DNA in motion during double-strand break repair. *Trends Cell Biol* 23:529–536
4. Carlton PM, Boulanger J, Kervrann C et al (2010) Fast live simultaneous multiwavelength four-dimensional optical microscopy. *Proc Natl Acad Sci U S A* 107:16016–16022
5. Magidson V, Khodjakov A (2013) Circumventing photodamage in live-cell microscopy. *Methods Cell Biol* 114:545–560
6. Inoué S, Inoué T (2002) Direct-view high-speed confocal scanner: the CSU-10. *Methods Cell Biol* 70:87–127
7. Oreopoulos J, Berman R, Browne M (2014) Spinning-disk confocal microscopy: present technology and future trends. *Methods Cell Biol* 123:153–175
8. Ettinger A, Wittmann T (2014) Fluorescence live cell imaging. *Methods Cell Biol* 123:77–94
9. Frigault MM, Lacoste J, Swift JL, Brown CM (2009) Live-cell microscopy – tips and tools. *J Cell Sci* 122:753–767
10. Jena NR (2012) DNA damage by reactive species: mechanisms, mutation and repair. *J Biosci* 37:503–517
11. Davies MJ (2003) Singlet oxygen-mediated damage to proteins and its consequences. *Biochem Biophys Res Commun* 305:761–770
12. Lisby M, Rothstein R, Mortensen UH (2001) Rad52 forms DNA repair and recombination centers during S phase. *Proc Natl Acad Sci U S A* 98:8276–8282
13. Khmelinskii A, Lawrence C, Roostalu J, Schiebel E (2007) Cdc14-regulated midzone assembly controls anaphase B. *J Cell Biol* 177:981–993

14. Lisby M, Rothstein R (2004) DNA damage checkpoint and repair centers. *Curr Opin Cell Biol* 16:328–334
15. Horigome C, Dion V, Seeber A et al (2015) Visualizing the spatiotemporal dynamics of DNA damage in budding yeast. *Methods Mol Biol Clifton NJ* 1292:77–96
16. Schindelin J, Arganda-Carreras I, Frise E et al (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9:676–682
17. Alkaabi KM, Yafea A, Ashraf SS (2005) Effect of pH on thermal-and chemical-induced denaturation of GFP. *Appl Biochem Biotechnol* 126:149–156



## The CellClamper: A Convenient Microfluidic Device for Time-Lapse Imaging of Yeast

Gregor W. Schmidt, Olivier Frey, and Fabian Rudolf

### Abstract

Time-lapse fluorescence imaging of yeast cells allows the study of multiple fluorescent targets in single cells, but is often hampered by the tedious cultivation using agar pads or glass bottom wells. Here, we describe the fabrication and operation of a microfluidic device for long-term imaging of yeast cells under constant or changing media conditions. The device allows acquisition of high quality images as cells are fixed in a two-dimensional imaging plane. Four yeast strains can be analyzed simultaneously over several days while up to four different media can be flushed through the chip. The microfluidic device does not rely on specialized equipment for its operation. To illustrate the use of the chip in DNA damage research, we show how common readouts for DNA damage or genomic instability behave upon induction with genotoxic chemicals (MMS, HU) or induction of a single double-strand break using induced CRISPR-Cas9 expression.

**Key words** Microfluidic, Perfusion, Long-term imaging, Fluorescence microscopy, Single-cell analysis, *S. cerevisiae*, DNA damage, CRISPR-Cas9, MMS, Hydroxyurea, Genomic instability

---

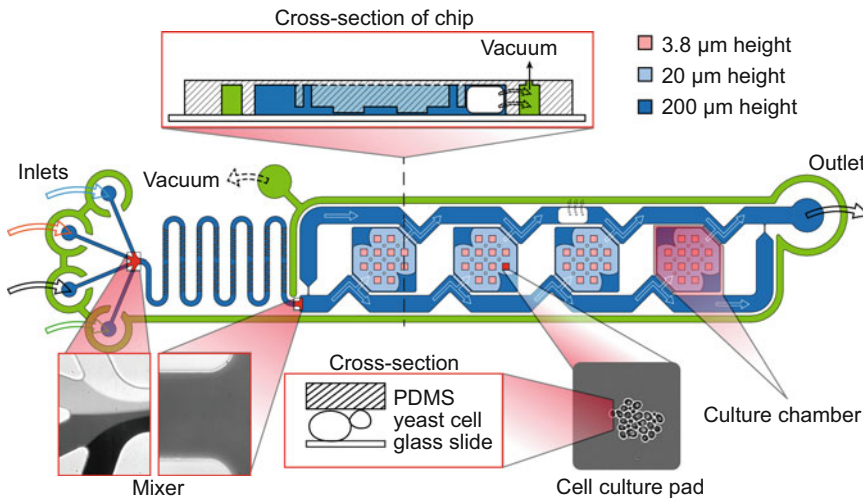
### 1 Introduction

Fluorescence time-lapse microscopy is instrumental for studying the spatial and temporal aspects of cell biology *in vivo*. Cellular processes are most commonly visualized by the use of genetically encoded reporter proteins. Fusing a fluorescent protein to a protein of interest provides a direct readout of its abundance and localization. Seminal time-lapse imaging studies used protein fusions to elucidate the spatio-temporal relationships between the budding yeast DNA damage checkpoint and repair proteins [1–3]. The timing and localization of these proteins remain useful reporters in live cell studies of the DNA repair process [4, 5]. Alternatively, fluorescent proteins can be used as transcriptional reporters by fusing their coding sequence to a promoter of interest [6, 7]. More elaborate reporter constructs allow for a monitoring of phosphorylation events upon DNA damage and therefore studies on the kinetics of the DNA damage response signaling became possible [8, 9].

The quality of time-lapse imaging data is determined largely by the experimental setup. Classically, time-lapse imaging of yeast is performed by embedding cells in agar pads. The agar keeps the cells in place and provides nutrients for the duration of the experiment. When studying DNA damage responses, induction of DNA damage can be achieved by chemical treatment of cells directly before embedding in the agar pad [3], or alternatively by exposing cells to UV-light using the microscope. While the agar pad has served researchers well, there are some drawbacks. The quality of the images is compromised by the autofluorescence of the media and the shape of the agar drop distorts the light path. Also, use of agar pads limits the duration of an experiment to <12 h and does not allow to dynamically administer genotoxic agents. One can partially bypass these nuisances by using glass bottom plates coated with concanavalin-A or poly-L-lysine for the cultivation of yeast cells [10]. While allowing for media exchange and better image quality, the obtained data is limited to the cells present at the beginning of the experiment as newborn cells usually do not adhere to the coating.

Microfluidic technology can overcome these limitations [11]. A recent study aimed at deciphering the spatial and temporal dynamics of the yeast proteome in response to DNA damage can illustrate the potential as well as the drawbacks of microfluidic devices [12]. In this study, DNA damage was repeatedly induced by pulsatile administration of genotoxic chemicals or UV-light allowing to study the initial relocalization kinetics of the proteins as well as their behavior when adapting to prolonged DNA damage. Continuous perfusion of microfluidic devices ensures rapid media exchange and enables long observation time through a constant supply of nutrients and removal of excess cells. Additionally, microfluidic devices allow better image quality as background fluorescence of the media can be excluded by trapping the cells between a glass cover slip and the microfluidic structure made of biocompatible and non-fluorescent polydimethylsiloxane (PDMS). In this case, high image quality can even be obtained when strongly autofluorescent media such as YPD are used.

Although many microfluidic devices for research have been developed and their advantages are evident, widespread use of microfluidics is hindered for several reasons. Commercial devices are either expensive or cannot be easily adapted to different experimental questions. Published devices have to be self-made and therefore require equipment and knowledge of the microfabrication process. Recently, custom-made microfabricated wafers or PDMS chips have become commercially available, allowing the experimentalist to bypass the fabrication process. However, establishing a working setup still requires dedicated equipment (e.g., plasma oven, microfluidic pump) and a substantial amount of time and knowledge which are not common in most biological labs.



**Fig. 36.1** Design of the microfluidic device for long-term culture and high-resolution imaging of yeast cells. Cells are trapped between a PDMS culture pad and a glass slide. The chip allows simultaneous culturing of four different yeast strains. A vacuum channel (*green*) surrounds the fluidic channel network (*blue*), and facilitates a tight seal of the PDMS to the glass and continuous removal of air bubbles from the system. Media can be switched or mixed from four separate inlets. A microfluidic herringbone mixer facilitates mixing, despite laminar flow, for flow rates up to 100  $\mu\text{L}/\text{min}$ . Design files are available upon request

These challenges can be overcome by engineering microfluidic devices which are simple to operate and do not rely on plasma bonding or actively driven pumping mechanisms.

Recently, we engineered a microfluidic device that can be operated with materials present in any biological lab [13] (Fig. 36.1). In the device, yeast cells are immobilized by clamping them between a glass slide and a PDMS culturing pad. The individual pads are surrounded by continuous media flow allowing the exchange of media and application of genotoxic chemicals. The media composition can be dynamically controlled and mixed from four different media reservoirs. In our microfluidic device, the PDMS and glass slide are kept together using a vacuum channel, obviating the need for plasma bonding used in conventional PDMS based devices. This allows for loading of cells by directly dispensing them on the microfluidic structure before assembling it with the coverglass. In each device, up to four different strains can be cultured simultaneously in four separate cell culture chambers. The vacuum channel, which is placed around the fluid filled channel system, continuously removes bubbles from the system ensuring long-term stability of the media flow. Perfusion of the device can be established using gravity flow or microfluidic pumps. Here, we describe the fabrication and operation of our microfluidic device in a step-by-step manual and show how DNA damage caused by genotoxic agents or induction of CRISPR-Cas9 can be measured inside the device. While the microfluidic device described in this

protocol was designed for culturing yeast (*S. cerevisiae* or *S. pombe*), it can be adapted to work with bacteria or mammalian cells by adjusting the height of the culturing pads.

---

## 2 Materials

Prepare all solutions using ultrapure water and use analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing waste materials. Wear protective clothing and glasses when handling solvents. Work with microfluidic devices requires a dust-free work environment.

### 2.1 Fabrication of Microfluidic Chip

- Particle free and sterile cleanroom gloves (*see Note 1*).
- Particle free cleanroom wipes.
- Silanized microfabricated wafer (*see Notes 2 and 3*).
- Scotch tape.
- Air spray gun with luer lock connector.
- 0.22 µm sterile filter with luer lock connector.
- Petri dishes (diameter: 150 mm).
- Polydimethylsiloxane (Sylgard 184).
- Plastic container and spatula.
- Desiccator.
- Level surface (optional: in convection oven at 80 °C).
- Razor blade or knife.
- Biopsy punch (outer diameter: 0.63 mm).

### 2.2 Preparation of Media Reservoirs and Microfluidic Tubings

- Microfluidic tubing (Tygon LMT-55, inner diameter: 0.51 mm, wall thickness: 0.85 mm).
- Blunt tip luer lock needle straight (inner diameter: 0.34 mm, outer diameter: 0.64 mm).
- Blunt tip luer lock needle 90° bent (inner diameter: 0.34 mm, outer diameter: 0.64 mm).
- Scissors.
- Flat pliers.
- Pincer pliers.
- Sterile disposable syringes with luer lock tip (10–60 mL).
- 70% ethanol in spray bottle.
- 70% ethanol in falcon tube.
- Growth media.
- 50 mL falcon tubes.

- Sterile disposable bottle top filters (pore size: 0.22  $\mu\text{m}$ ).
- Aluminum foil.
- Optional: Sterile hood.
- Optional: Microfluidic syringe pump and glass syringes with luer lock tip (10–60 mL, ILS Innovative Labor Systeme GmbH, Mittelstrasse 37, 98714 Stuetzerbach, Germany).

### **2.3 Setting Up a Time-Lapse Imaging Experiment Using the Microfluidic Device**

- Acetone, isopropanol, and ultrapure water in squeeze bottles.
- Tube rack which can accommodate syringe OR microfluidic pumps with matching syringe holders.
- Motorized inverted fluorescence microscope with high NA oil immersion objective (40–100 $\times$  magnification).
- Vacuum pump.
- Self-adhesive paper tape.
- Nail polish (clear).
- Cover slip to accommodate microfluidic chip (thickness: 0.13–0.16 mm). We use 60 mm  $\times$  24 mm (Thermo-Fischer Scientific, catalog number size: Q10143263NR1). Differently sized cover slips can be used but should measure at least 50 mm  $\times$  14 mm to fit microfluidic chip.
- Microscope slide holder fitting the cover slip with adjustable screws for levelling.
- Kimtech wipes.
- Fume hood.
- Yeast culture in exponential growth phase.
- Foldback clips.
- Small volume pipette (able to dispense  $\approx$  0.4  $\mu\text{L}$ ).
- 1.5 mL Eppendorf tubes.
- Microcentrifuge.

---

## **3 Methods**

### **3.1 Fabrication of Microfluidic Chip**

Work in a dust-free environment and use cleanroom gloves for all steps. PDMS molding of one wafer yields seven microfluidic chips which can be re-used as often as  $\approx$  30 times (*see Note 4*). Therefore these steps do not need to be carried out for each experiment.

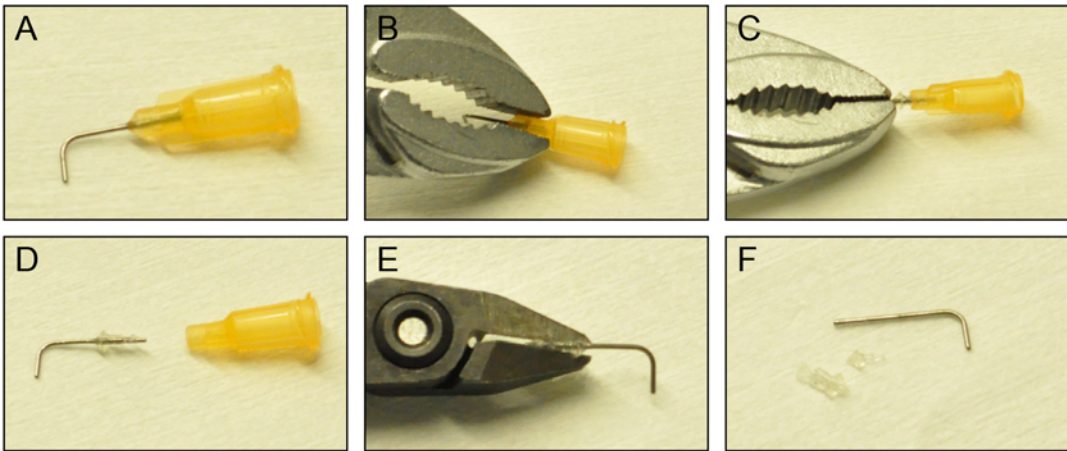
1. Using Scotch tape, build a barrier around the silicon wafer. Blow any dust off the wafer using the air spray gun (*see Note 5*). Place wafer into polystyrene Petri dish.

2. In a plastic beaker vigorously mix PDMS monomer and curing agent in a ratio of 10:1 (w:w) using a plastic spatula. Prepare 30 g of PDMS mixture per wafer (*see Note 6*).
3. Degass PDMS mixture for  $\approx 30$  min in a desiccator, until solution is free of bubbles (*see Note 7*).
4. Place wafer inside the Petri dish on a scale and pour 25 g of degassed PDMS onto the wafer.
5. Place Petri dish with wafer in a desiccator for  $\approx 30$  min to remove trapped air bubbles (*see Note 8*).
6. Place Petri dish with wafer on a level surface in a convection oven and incubate for at least 1 h at 80 °C to cure PDMS (*see Note 9*).
7. Remove Scotch tape from wafer. Carefully peel the PDMS from the wafer by slowly releasing it from each side.
8. Cut single microfluidic devices from PDMS slab using a razor blade or knife.
9. Punch inlet and outlet holes using the biopsy punch (*see Note 10*). The chip allows attachment of four different media reservoirs through separate inlets. Punch only inlet holes that you need, as they cannot be closed afterwards.
10. Store chips in petri dish at room temperature until further use.

### **3.2 Preparation of Microfluidic Tubings and Media Reservoirs for Gravity Driven Operation of the Microfluidic Device**

Handle all tubings and connectors using cleanroom gloves to prevent contamination with particles. Clean workspace using wet paper towels to create a dust-reduced environment or work in a sterile hood. Microfluidic tubing can be re-used. Therefore, **steps 1–5** do not need to be carried out for each experiment.

1. Loosen the plastic connector of a 90° bent blunt tip luer lock needle by gently squeezing the connector using flat pliers (Fig. 36.2b). Pull the metal tubing from the connector (Fig. 36.2c, d). Remove any residual glue from the metal tube using pincer pliers (Fig. 36.2e, f, *see Note 11*).
2. Cut microfluidic tubing for vacuum connection (10 cm), fluidic inlets (80 cm), and fluidic outlet (20 cm) using scissors.
3. For each microfluidic tubing insert a bend metal tube (from **step 1**) at one end.
4. For vacuum tubing and inlet tubings insert straight blunt tip luer lock needle into the other end (*see Note 12*).
5. Sterilize microfluidic tubing by flushing it with 1 mL of 70% ethanol using a syringe with a luer lock tip.
6. Prepare yeast growth media as usual and filter sterilize using disposable bottle top filters (*see Note 13*).



**Fig. 36.2** Fabrication of microfluidic connectors. (a) A 90° bent blunt tip luer lock needle. (b) Gently squeeze the seam where the metal tubing is glued into the connector using flat pliers. (c) Pull the metal tubing from the connector. (d) Extracted metal tubing. (e) Remove residual glue from the metal tubing using pincer pliers. (f) Ready to use metal tube

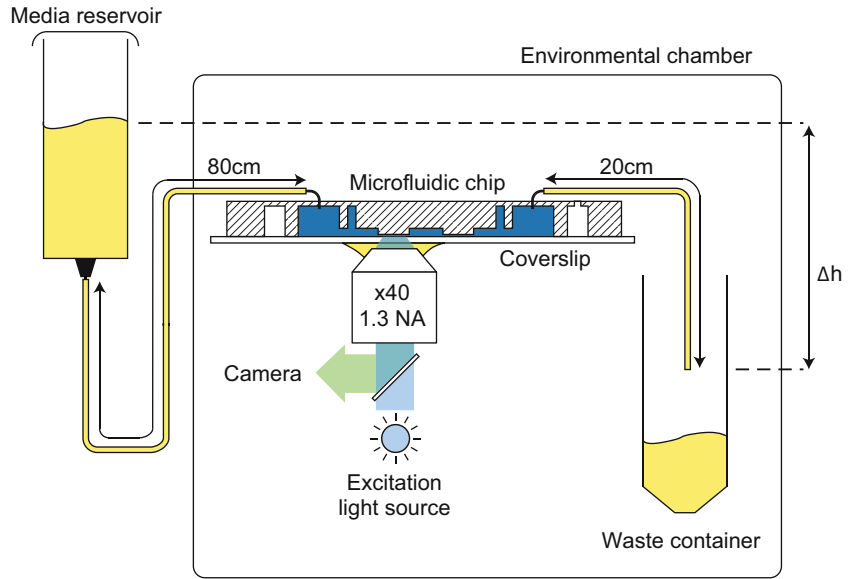
7. Disassemble the syringe you want to use as a media reservoir, by removing the piston from the syringe body. For setting up the experiment using syringe pumps (*see Note 14*).
8. Attach microfluidic tubing to the syringe using the luer lock connector.
9. Place syringe in a tube rack.
10. Fill media into syringe and close top of the syringe using aluminum foil.
11. Let media flow through the tubing until all residual ethanol in the tubing is removed. Close tubing using a foldback clip.
12. Store end of microfluidic tubing in a paper towel soaked with 70% ethanol.
13. Transfer rack with syringe(s) to microscope (Fig. 36.3).

### 3.3 Setting Up a Time-Lapse Imaging Experiment Using the Microfluidic Device

Wear cleanroom gloves, protective glasses, and labcoat.

1. In a fume hood, clean the PDMS chip from all sides using acetone from a squeeze bottle. Rub the structured side of the chip using your fingers to remove any PDMS particles from the fabrication process. Finally, hold the chip at its sides using thumb and forefinger and give it a final rinse (*see Note 15*).
2. Rinse the chip with isopropanol and water from squeeze bottles, while holding it between thumb and forefinger (*see Note 15*).
3. Dry the PDMS chip by blowing air from an air spray gun onto the chip surface. To avoid particle contamination from pressurized air supply, mount a sterile filter on the air spray gun. Hold



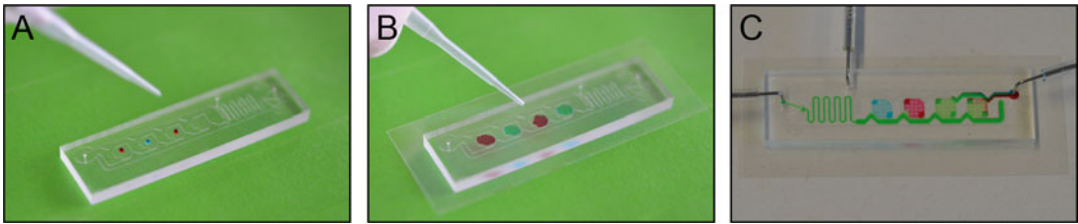


**Fig. 36.3** Overview of microfluidic chip setup. Media is delivered by gravity flow from a media reservoir. The flow rate is determined by the resistance of the microfluidic system (tubing and microfluidic chip) and the height difference ( $\Delta h$ ) between the media reservoir liquid level and the end of the outlet tubing (*see Note 27*). The media reservoir should be located outside the environmental chamber to prevent formation of bubbles and limit evaporation

the chip against a light source to check if the cleaning procedure was successful. If the chip shows traces of dust or particles, repeat previous steps (*see Note 16*).

4. Store chip with the structured side facing up on a cleanroom paper in a clean environment (e.g., cell culture hood).
5. Clean cover slip by rinsing with acetone, isopropanol, and water from spray bottles. Hold the coverglass between thumb and forefinger. Dry cover slip using air spray gun and check cleanliness against light source. Store clean coverglass on a cleanroom paper.
6. Pipette 0.4  $\mu\text{L}$  of cell solution ( $1 \times 10^7$  cells/mL) into the center of each culturing chamber of the microfluidic chip (Fig. 36.4). The PDMS chip is hydrophobic and the cell suspension will form a standing droplet on the chip surface (*see Notes 17 and 18*).
7. Position the coverglass  $\approx 3$  mm above the microfluidic chip surface. Bring the coverglass into contact with the PDMS on one end of the microfluidic chip. Suspend the coverglass on the other end using a pipette tip. Slowly lower the coverglass and let it adhere to the PDMS over the whole chip surface. **Steps 6–7** should be performed in  $<1$  min to prevent significant evaporation of cell media (*see Notes 19 and 21*).





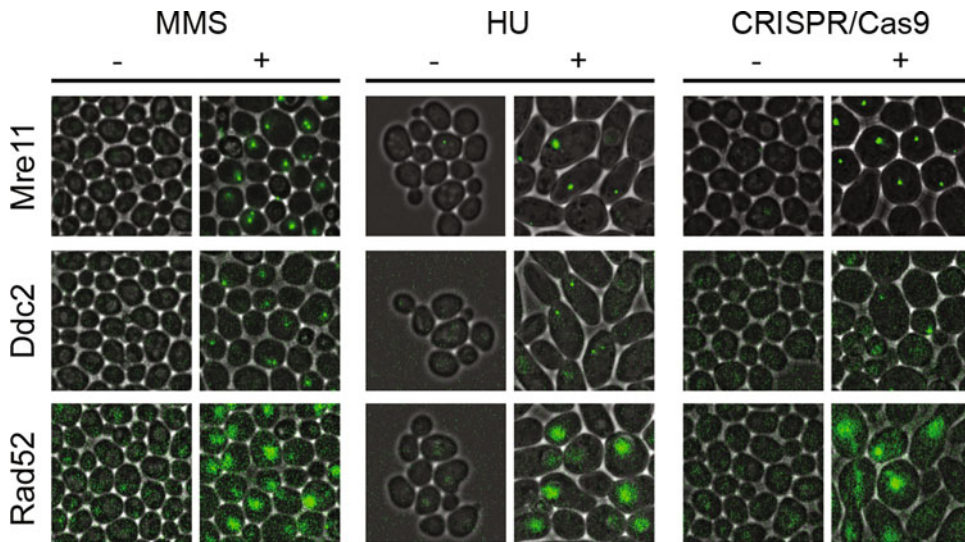
**Fig. 36.4** Loading of cells into the microfluidic device. (a) 0.4  $\mu\text{L}$  cell solution is pipetted into each culturing chamber (Step 6). (b) A coverglass is brought into contact with the PDMS chip (Step 7). The cell solution spreads in the culturing chamber. (c) The vacuum tubing (*top*), inlet (*left*) and outlet (*right*) tubing is connected to the chip (Step 13) and perfusion is started. The microfluidic device fills with media. Trapped air bubbles will be removed by diffusion of air through the PDMS to the vacuum channel

8. Pick up the PDMS-coverglass sandwich and place it glass-side down in a polystyrene Petri dish to transfer it to the microscope (*see Note 20*).
9. (Optional) If available preheat environmental chamber of microscope to desired temperature  $\approx 2$  h before starting the experiment, to allow thermal equilibration of the equipment.
10. Place vacuum pump in vicinity of the microscope. From the vacuum pump a tubing with a luer connector at its end should be laid into the microscope and fixed with self-adhesive paper tape.
11. The outlet tubing should be fixed on the microscope stage using self-adhesive paper tape. The end of the tubing should be placed into a waste container and fixed with self-adhesive paper tape (*see Note 22*).
12. Place the microfluidic chip on a flat part of the microscope stage.
13. Connect the microfluidic tubing by pressing the metal connectors into the access holes of the microfluidic chip. Connect tubings in the following order (*see Note 23*):
  - (a) vacuum tubing
  - (b) outlet tubing
  - (c) inlet tubing(s) (*see Note 24*).
14. Carefully place the microfluidic chip into the microscope slide holder (*see Note 25*).
15. Remove the foldback clip from the media reservoir with which you want to perfuse the cells (*see Notes 26–28*).
16. Place microfluidic chip in stage holder and fix microfluidic tubings on the microscope stage using self-adhesive paper tape. Try to fix tubing without creating tension on the microfluidic device (*see Note 29*).

17. Fix the stage holder on the microscope stage using self-adhesive paper tape (*see Note 30*).
18. (Optional) If using an immersion objective, raise the objective until it is located  $\approx 1$  mm below the coverglass. Move the microscope stage so that the objective is at the edge of the coverglass. Add immersion liquid to the gap between the coverglass and the objective (*see Note 31*).
19. Move the microscope stage so that the objective is located below a cell culture chamber. Bring sample into focus by slowly moving the objective upwards and observing the sample through the eyepiece or the camera live-stream. The cells should show a strong membrane diffraction pattern when they are slightly out-of-focus (*see Note 32*).
20. Fix the microfluidic device inside the stage holder by applying one drop of nail polish to each corner of the coverglass. Leave the doors of the environmental chamber open for  $\approx 30$  min to let nail polish solvent escape (*see Note 33*).
21. Use the adjustment screws of the stage holder to level the microfluidic chip with the imaging plane (*see Note 34*).
22. Leave microfluidic device to thermally equilibrate for at least 30 min (*see Note 35*).
23. Program multi-position time series measurement using the microscope software. If available enable hardware or software autofocussing routine, to keep sample in focus over the whole course of the experiment (*see Note 36*).
24. After the experiment, disassemble the microfluidic device by closing the media lines with foldback clips. Gently remove all paper tape and pull out all tubings from the microfluidic chip. Rinse microfluidic tubings with ultrapure water and 70% ethanol and store in a Petri dish until further use. Disassemble the microfluidic device by gently peeling the PDMS chip from the coverglass (*see Note 37*). Vigorously rinse the PDMS chip using ultrapure water to remove cells and blow dry using the air spray gun. Inspect cleanliness under a microscope. Store the PDMS chip in a Petri dish until further use.

### **3.4 Live-Cell Microscopy of DNA Damage Induced by Chemicals or CRISPR/ Cas9 Induction**

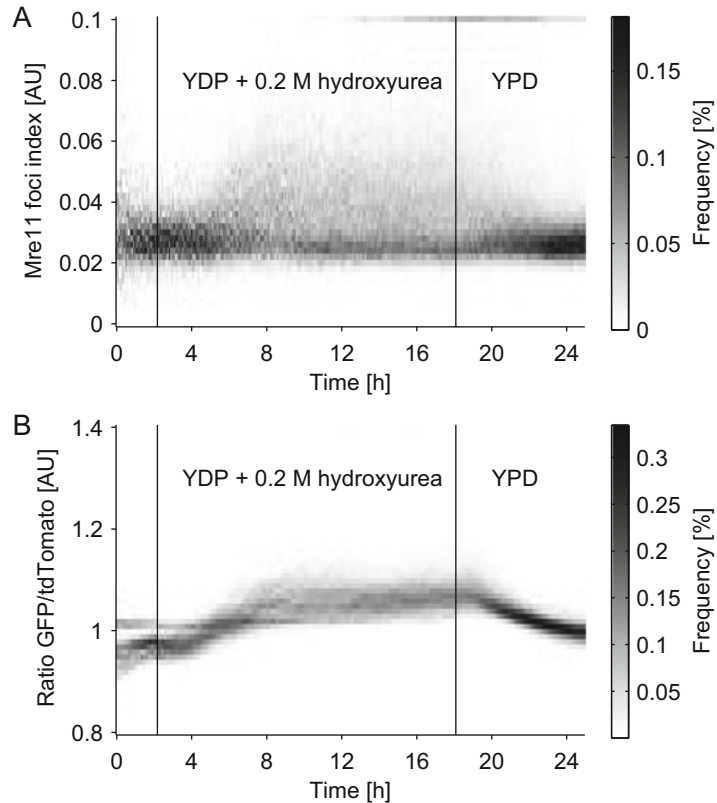
To illustrate the possibility of inducing DNA damage in the microfluidic device, we cultured strains containing a fluorescent protein fused to either Mre11, Ddc2, or Rad52. The cells were grown until they formed a small colony and DNA damage was induced by switching the media. First, we cultivated the cells using low fluorescent SD media and induced DNA damage using 0.03% MMS. The relocation pattern of all three proteins were as reported in the literature (Fig. 36.5). As the culture pads of the microfluidic chip reduce the background fluorescence by excluding media from the fluorescence excitation volume, the chip allows to capture images



**Fig. 36.5** Detection of DNA damage by imaging of fluorescent fusion proteins in the microfluidic chip. DNA damage was induced chemically by perfusing media containing either 0.03% methyl methanesulfonate (MMS) or 0.2 M hydroxyurea (HU). Site specific double-strand breaks were generated at the URA3 locus using an inducible CRIPRS-Cas9 construct (unpublished). DNA damage was detected by nuclear localization of fluorescent signals in strains carrying fluorescent protein fusions with Mre11, Ddc2, or Rad52. Images are overlays of brightfield and fluorescent channels. The dynamic range of the fluorescence channel has been equally adjusted for each pair of images (induced image and non-induced control image), to make the signal visible

with high signal-to-noise ratio even in strongly autofluorescent media. We cultured cells in YPD and induced DNA damage by either applying 0.3 M hydroxyurea or introducing a single double-strand break using an inducible CRISPR-Cas construct (unpublished) targeting the URA3 locus. In all cases, relocation of the fluorescent protein fusions was swift and well detectable (Fig. 36.5). This demonstrates that using our device allows microscopy assays to be performed independent of the autofluorescent properties of the media.

Next, we wanted to illustrate how dynamic exposure of cells to genotoxic agents can be monitored using the microfluidic device. We loaded cells carrying an Mre11 fluorescent protein fusion and cultured them in YPD for  $\approx 2$  h until they formed small colonies. We then exposed the cells to a pulse of YPD + 0.2 M hydroxyurea for 16 h, and finally released genotoxic stress by switching back to YPD. The formation of nuclear Mre11 foci was scored by calculating the change of the nuclear fluorescence interquartile range (Fig. 36.6a). Foci become visible in a small subset of cells within a short period of time and after 2 h the majority of the cells show nuclear foci and stop dividing. Within  $\approx 1$  h after hydroxyurea is removed, nuclear foci disappear and division restarts for a subpopulation of cells. Newborn cells exhibit no nuclear foci. However, for



**Fig. 36.6** Imaging DNA damage in an arrest-and-release experiment using hydroxyurea. Cells were loaded into the microfluidic device and cultured in YPD for 2 h. To induce DNA damage, the cell was exposed to a pulse of YPD + 0.2 M hydroxyurea for 16 h. After the pulse the media was switched back to YPD. Images were acquired in 5 min intervals. **(a)** A strain carrying an Mre11 fluorescent fusion protein was imaged. The Mre11 foci index was calculated for each single cell from the interquartile range of the nuclear fluorescent fusion protein signal ( $(nuc_{q75} - nuc_{q25})/nuc_{q75}$ ). At each timepoint the normalized histogram of the resulting foci indexes is plotted. **(b)** A strain carrying an Rnr3-GFP fusion protein and an Rpl39pr-tdTomato expression construct was imaged. For each cell, the median Rnr3-GFP signal was divided by the median signal of the constitutively expressed Rpl39pr-tdTomato construct. The normalized histogram of the resulting values is plotted for each timepoint

many old cells nuclear foci remain present (visible at the upper right corner of the plot). As the fraction of newborn cells in the field of view increases, the average Mre11 nuclear foci index returns to the same level as before exposure to hydroxyurea.

To illustrate the behavior of a protein abundance reporter, we repeated the hydroxyurea arrest-and-release experiment using the parent strain from a genome wide screen for genomic instability. In this strain, expression of an Rnr3-GFP fusion protein is induced in

response to DNA damage. To account for cell-to-cell variability, the strain also contains a constitutively expressed Rpl39pr-tdTomato reporter construct, which is used to normalize the Rnr3-GFP signal. The onset of Rnr3-GFP expression starts about 2 h after induction and reaches a plateau after 5.5 h. Upon removal of hydroxyurea, the GFP signal decays as a function of the division time. As restart of division is not instantaneous, there is a slight delay until the GFP-signal starts to vanish. Taken together, these experiments illustrate the usability of the device for imaging yeast cells under changing media conditions.

In all cases, the cells were cultivated in an environmental chamber at 30° and images were acquired using a Nikon Ti-E microscope equipped with a 40× objective (oil immersion, NA 1.3). The microscope was controlled using the software YouScope [14]. For accurate cell segmentation, brightfield images were taken above and below the focal plane ( $\pm 5$  AU, Nikon Perfect Focus System) and divided using ImageJ. Division of brightfield images corrects illumination and background aberrations, and enhances the diffraction pattern of the cell boundary. The divided images were segmented and quantified using CellX [15]. The extracted intensities were plotted using Matlab.

---

## 4 Notes

1. Always use cleanroom gloves when handling microfluidic components. Conventional gloves (even powder free ones) contain particles which can cause background fluorescence, blocking of the microfluidic channels and interfere with adhesion of PDMS to glass.
2. Design files for microfluidic devices are available upon request. If a cleanroom with SU-8 processing capabilities is available, wafers can be fabricated and silanized as described [16]. If microfabrication facilities are not available, silanized custom wafers can be ordered from a microfabrication company, e.g.,
  - MicruX Fluidic (Severo Ochoa Building, Julian Claveria s/n, 33006 Oviedo (Asturias), SPAIN)
  - SIMTech Microfluidics Foundry (Singapore Institute of Manufacturing Technology, 71 Nanyang Drive, Singapore 638075)
  - FlowJEM (<http://www.flowjem.com/>)
3. Store wafers in a cool, dry, and dark place. A humid atmosphere will compromise the silanization of the wafer. Exposure to heat and sunlight will be detrimental for adhesion of the SU-8 on the wafer.

4. The microfluidic device can be re-used as often as  $\approx 30$  times. However, hydrophobic chemicals can be absorbed by the PDMS of the microfluidic chip and the tygon tubing. Therefore, care has to be taken if parts of the fluidic setup are re-used to exclude cross-contamination from previous experiments.
5. Always handle the wafer by only touching it on its edges. Never touch or attempt to clean the wafer surface using solvents. Dust that cannot be blown off using the air spray gun will get embedded in the PDMS slab and removed when demolding, leaving a clean wafer surface.
6. Immediately clean spilled PDMS using a wipe soaked with isopropanol. Cured PDMS is not soluble, and has to be removed physically.
7. Take care to stop the vacuum pump before PDMS foam spills over the edge of the plastic beaker. Place a piece of aluminum foil in the desiccator to contain possible PDMS spills.
8. Take care to stop vacuum pump before PDMS foam spills over the edge of the Scotch tape.
9. If a convection oven is not available, place a Petri dish onto a level surface and leave for 24 h at room temperature to cure PDMS.
10. Take care to punch holes perpendicular to the chip surface. Otherwise it will be difficult to connect the microfluidic tubing later.
11. Try not to squeeze or bend the metal tubing itself with the pliers, as this might collapse the channel inside the tubing and block it.
12. Store tubing in a Petri dish if they are not to be immediately used.
13. Filtering is carried out to prevent clogging of the microfluidic device due to particles which might be present in conventionally prepared media.
14. If the microfluidic device is to be perfused using syringe pumps, follow these instructions. Syringe pumps allow precise control of the flow rate at the different inlets. Using syringe pumps media can be mixed on-chip and cells can be exposed to precisely timed trajectories of different media conditions. Glass syringes should be used with syringe pumps, as they can be more rigidly connected to the syringe pump holder than plastic syringes. Also, glass syringes have faster response times and show less hysteresis when the flow rates are changed.
  - (a) In a sterile hood, disassemble the glass syringes by pulling out the plunger.

- (b) Spray the syringe body and the plunger with generous amounts of 70% ethanol from a spray bottle, to sterilize the syringe. Leave the parts to dry in the sterile hood for  $\approx$  30 min.
  - (c) When the syringe is completely dry, reassemble it and push the plunger all the way into the syringe body. Connect a sterile filter (pore size: 0.22  $\mu\text{m}$ ) to the syringe.
  - (d) Through a microfluidic tubing, draw media through the filter into the syringe by pulling on the plunger until the syringe is full.
  - (e) Bubbles will be present at the plunger and wall inside the syringe. Hold the syringe with its tip is facing upwards while holding it on the plunger. Using your thumb and forefinger quickly rotate the syringe along its longitudinal axis, to release bubbles from the plunger and wall, which will move upwards towards the tip of the syringe.
  - (f) Once all bubbles have been collected at the tip of the syringe, remove the sterile filter and connect a freshly sterilized microfluidic tubing, while keeping the syringe tip facing upwards. Push the plunger to eject all bubbles from the syringe, until the microfluidic tubing is filled completely filled with growth media.
  - (g) Store end of microfluidic tubing in a paper towel soaked with 70% ethanol.
  - (h) Transfer syringe to microfluidic syringe pump located at the microscope.
  - (i) Syringe pumps should be located outside the environmental chamber of the microscope, since heating of the media inside the syringes will lead to out-gassing of dissolved air and formation of bubbles inside the syringes. Bubbles inside the syringes should be avoided to allow fast and precise control of the flow rates.
  - (j) The microfluidic device should be perfused with 5–100  $\mu\text{L}/\text{min}$ . Faster flow rates allow quicker media exchange and switching, and facilitate efficient removal of excess cells. At slower flow rates the limited syringe volume can be used for longer experiments. At flow rates  $>100 \mu\text{L}/\text{min}$  the cells underneath the culturing pads get flushed away by the media flow.
15. Avoid letting the solvents run from your gloves onto the chip, as it will leave traces which interfere with chip adhesion to the cover slide.
16. Try not to let water dry on the chip, but rather blow it off the chip, as it might leave traces.

17. The cell concentration determines the number of cells per culturing pad in the microfluidic chip. Higher cell concentrations (up to  $10\text{--}20 \times 10^7$  cells/mL) can be used if cells are to be cultured for a short time only (1–4 h) and more cells are needed for the analysis. For loading of single cells, cells should be in the exponential phase of growth. Determine cell concentration of yeast culture using a Coulter Counter, counting chamber, or calibrated photometer. To achieve the desired cell concentration transfer 1 mL of cell solution into a 1.5 mL Eppendorf tube. Sediment cells by centrifuging the tubes for 2 min at  $1000g$ . Remove enough supernatant to reach desired cell concentration after resuspending the cell solution. Take care not to remove cells when removing supernatant. Resuspend cells by carefully flicking the tube or by a short vortexing pulse at a low speed setting. Depending on the culture conditions and growth phase cells might sediment as a pellet, or might float at the top of the media solution. If cells are used at later cultivation stages they might form clumps and get loaded as cell colonies. Also, cells in later phases of culturing might stick to PDMS strongly.
18. Use low adhesion pipette tips to allow efficient transfer of the whole cell culture volume to the microfluidic chip surface. Depending on the pipette, pipette tip, and media composition the actual volume that gets transferred to the chip might vary. Adjust the setting on the pipette accordingly. The cell culture chambers should not be completely filled with cell suspension, as this can lead to cross-contamination of strains between chambers.
19. If culture chambers are filled  $<50\%$  the pipetted volume was too small or evaporation during chip loading was too high. Practice loading of the device.
20. Always handle the assembled chip by holding it between two fingers on the coverglass. Avoid any bending of the microfluidic device, as this can lead to adhesion of the culturing pads to the coverglass.
21. The chip structure should be clearly visible if the PDMS-glass sandwich is held against a light source. Particle or dust contamination is directly visible as defects in the glass-PDMS interface. If contamination is observed disassemble PDMS and coverglass. Discard the coverglass. Clean the PDMS chip and a new coverglass as described earlier.
22. Make sure the outlet tubing will not escape the waste container during stage movement, as this will cause spillage on the optical table. The outlet tubing should have enough slack, so it can be comfortably connected to the microfluidic chip without tension. The end of the outlet tubing should be  $\approx 5$  cm below the stage level.



23. Immediately before pressing the tubing into the access holes, sterilize them using a Kimtech wipe soaked with 70% ethanol. The ethanol will evaporate immediately and leave a thin film of water on the metal pin, which facilitates insertion of the tubing into the PDMS. Position the microfluidic tubing above the access hole with one hand. With the other hand, gently press on top of the bent metal tube to insert it into the PDMS. Try not to touch the PDMS chip itself, as the PDMS might adhere to the coverglass in the regions of the culturing pads.
24. If several media reservoirs are to be connected to the chip, connect the media reservoir which is to be used at the beginning of the experiment, first. Then remove the foldback clip from this media reservoir and let the media flow into the microfluidic device until it starts to flow out of the other inlet holes. Then close the media reservoir with the foldback clip again and connect other media reservoirs. When all reservoirs are connected proceed with next step.
25. Make sure the cover slide fits into the slide holder without tension. Tension on the coverglass leads to bending of the microfluidic device, which impedes alignment of the sample with the imaging plane of the microscope.
26. Make sure the liquid level of the reservoir is 5 cm above the microfluidic chip, to allow media flowing into the chip. The flow rate during filling of the microfluidic device is higher because an empty system has a lower resistance. If the flow rate during filling is too high, delamination of the chip and leakage into the vacuum channel is possible.
27. The sum of the length of the microfluidic inlet and outlet tubing determines the flow rate through the microfluidic device. For a tubing length of 100 cm (80 cm inlet tubing + 20 cm outlet tubing) and a height difference of 10 cm between the reservoir liquid level and the end of the outlet tubing, the flow rate will be  $\approx 10 \mu\text{L}/\text{min}$ . The flow rate scales linearly with tubing length and height difference.
28. If you observe media leaking into the vacuum channel, close the inlet tubing again. In this case the chip was not cleaned thoroughly enough and a particle contamination has created a connection between the media channel and the vacuum channel. Disassemble the chip and repeat the loading procedure.
29. If there is tension on the tubing, the microfluidic chip might delaminate from the coverglass which will result in leakage of media into the vacuum channel. Also, tension on the tubing may cause bending of the microfluidic device (*see Note 25*)
30. Usually the slide holder will have some play when it is inserted in the microscope stage. To prevent moving of the slide holder during time-lapse imaging, it needs to be fixed with adhesive tape.

31. Be generous with the amount of immersion liquid. When performing multi-position imaging, the objective will move over extended areas of the microfluidic device (distances of 20–30 mm), and the immersion liquid might not be able to follow the objective if the volume is too low.
32. If cells do not show a strong diffraction pattern, the cell culture pads have collapsed onto the coverglass (probably while the tubings were inserted). In this case disassemble the microfluidic device and repeat loading procedure.
33. Use only clear nail polish, as colored nail polish can contain fluorescent molecules which can increase background fluorescence during the measurement.
34. Move the stage so that one culture pad with a high number of cells is in the field of view. Move the objective downwards to slightly de-focus the sample. Then, slowly move the objective upwards and observe on which part of the field-of-view the cells come into focus first. The adjustment screws on this side of the chip have to be screwed inwards. Repeat the procedure until all cells come into focus simultaneously.
35. In the first 30 min the microfluidic device and the stage holder might still move 50–200  $\mu\text{m}$ , which is detrimental for time-lapse imaging.
36. If immersion liquid or hardware autofocus cannot follow the stage movement during multi-position imaging, reduce the stage speed. Make sure that all positions can be imaged in the desired imaging interval.
37. Take care not to break the coverglass, as the shards might damage the PDMS structure. Discard the coverglass.

---

## Acknowledgements

This work was supported by the FP7 grant no 28995 ITN ISOLATE and the Ambizione grant 142440 of the Swiss National Science Foundation to Olivier Frey. We thank Andreas Hierlemann and Jörg Stelling for their support. We acknowledge Mjriam Bächler and Andreas Cuny for testing the microfluidic device and Jennifer Ewald and Tania Roberts for critical reading of the manuscript.

## References

1. Lisby M, Rothstein R, Mortensen UH (2001) Rad52 forms DNA repair and recombination centers during S phase. *Proc Natl Acad Sci USA* 98(15):8276–8282. doi:10.1073/pnas.121006298
2. Lisby M, Mortensen UH, Rothstein R (2003) Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat Cell Biol* 5(6):572–577. doi:10.1038/ncb997

3. Lisby M, Barlow JH, Burgess RC, Rothstein R (2004) Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* 118 (6):699–713. doi:10.1016/j.cell.2004.08.015
4. Silva S, Gallina I, Eckert-Boulet N, Lisby M (2012) Live cell microscopy of DNA damage response in *Saccharomyces cerevisiae*. *Methods Mol Biol* 920:433. doi:10.1006/cbir.1999.0447. arXiv:1011.1669v3
5. Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, Jaschob D, Hendry JA, Ou J, Moffat J, Boone C, Davis TN, Nislow C, Brown GW (2012) Dissecting [DNA] damage response pathways by analysing protein localization and abundance changes during [DNA] replication stress. *Nat Cell Biol* 14(9):966–976. doi:10.1038/ncb2549
6. Gasch AP, Huang M, Metzner S, Botstein D, Elledge SJ, Brown PO (2001) Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. *Mol Biol Cell* 12(10):2987–3003. doi:10.1091/mbc.12.10.2987
7. Birrell GW, Brown JA, Wu HI, Giaever G, Chu AM, Davis RW, Brown JM (2002) Transcriptional response of *Saccharomyces cerevisiae* to DNA-damaging agents does not identify the genes that protect against these agents. *Proc Natl Acad Sci USA* 99(13):8778–8783. doi:10.1073/pnas.132275199
8. Johnson SA, You Z, Hunter T (2007) Monitoring ATM kinase activity in living cells. *DNA Repair* 6(9):1277–1284. doi:10.1016/j.dnarep.2007.02.025
9. Burgess A, Lorca T, Castro A (2012) Quantitative live imaging of endogenous DNA replication in mammalian cells. *PLoS One* 7 (9). doi:10.1371/journal.pone.0045726
10. Bush A, Chernomoretz A, Yu R, Gordon A, Colman-Lerner A (2012) Using cell-ID 1.4 with R for microscope-based cytometry. *Curr Protoc Mol Biol* (Suppl 100):1–27. doi:10.1002/0471142727.mb1418s100
11. Duncombe TA, Tentori AM, Herr AE (2015) Microfluidics: reframing biological enquiry. *Nat Rev Mol Cell Biol* 16(9):554–567. doi:10.1038/nrm4041
12. Dénervaud N, Becker J, Delgado-Gonzalo R, Damay P, Rajkumar AS, Unser M, Shore D, Naef F, Maerkl SJ (2013) A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc Natl Acad Sci USA* 110 (39):15842–15847. doi:10.1073/pnas.1308265110
13. Frey O, Rudolf F, Schmidt GW, Hierlemann A (2015) Versatile, simple-to-use microfluidic cell-culturing chip for long-term, high-resolution, time-lapse imaging. *Anal Chem* 87 (8):4144–4151. doi:10.1021/ac504611t
14. Ferry MS, Razinkov IA, Hasty J (2011) Microfluidics for synthetic biology: from design to execution. In: *Methods in enzymology*, vol 497. Elsevier, Amsterdam, pp 295–372
15. Lang M, Rudolf F, Stelling J (2012) Use of YouScope to implement systematic microscopy protocols. In: *Current protocols in molecular biology*, vol 14. Wiley, New York, pp 1–23. doi:10.1002/0471142727.mb1421s98
16. Mayer C, Dimopoulos S, Rudolf F, Stelling J (2013) Using CellX to quantify intracellular events. In: *Current protocols in molecular biology*, vol 14, pp 1–20. doi:10.1002/0471142727.mb1422s101

## Characterization of Structural and Configurational Properties of DNA by Atomic Force Microscopy

Alice Meroni, Federico Lazzaro, Marco Muzi-Falconi,  
and Alessandro Podestà

### Abstract

We describe a method to extract quantitative information on DNA structural and configurational properties from high-resolution topographic maps recorded by atomic force microscopy (AFM). DNA molecules are deposited on mica surfaces from an aqueous solution, carefully dehydrated, and imaged in air in Tapping Mode. Upon extraction of the spatial coordinates of the DNA backbones from AFM images, several parameters characterizing DNA structure and configuration can be calculated. Here, we explain how to obtain the distribution of contour lengths, end-to-end distances, and gyration radii. This modular protocol can be also used to characterize other statistical parameters from AFM topographies.

**Key words** Atomic force microscope/microscopy (AFM), DNA, Mica, DNA conformation

---

### 1 Introduction

Across more than three decades, atomic force microscopy (AFM) has become a technique of choice for the quantitative investigation of biomolecules such as DNA, proteins, and their complexes (for an overview, *see* [1] and references therein). The success of AFM relies on its ability to provide nanometer spatial resolution in  $XY$ , sub-nanometer resolution in  $Z$ , as well as on the capability of imaging biological samples in their physiological conditions. Since the advent of AFM, DNA has been the privileged target of innumerable studies, due to its paramount biological relevance (for a review, *see* [2]). Being a semirigid, charged, strong polyelectrolyte, DNA concentrates in itself a wealth of interesting physics, and has a great potential for nanobiotechnological applications [3, 4]; for this reasons, DNA has been an ideal benchmark for biophysical studies [5–17]. Such studies mostly rely on the statistical characterization of structural and configurational properties of a population of double stranded DNA molecules with fixed length. Several

parameters describe the equilibrium configuration of DNA molecules: contour length, end-to-end distance, rise per residue, radius of gyration, bending angle distribution along the DNA backbone, persistence length, to cite the most important. The statistical mechanics of semirigid polymers describes the distribution of these parameters in equilibrium conditions [5]. Structural and configurational changes of DNA molecule can occur as a consequence of modification of the DNA environment (ionic strength, pH, surface charge density) [6, 7, 18], as well as a consequence of internal changes at the base pairs level (such as mispaired and damaged bases [19, 20]); it follows that high-resolution imaging of DNA molecules, as provided by AFM, can be a very valuable complement of biomolecular studies.

Here, we describe a general protocol for acquiring high-resolution images of DNA molecules, either linear fragments or plasmids, using an atomic force microscope operated in Tapping (or intermittent contact) Mode in air. In Tapping (or intermittent contact) Mode, a sharp oscillating probe periodically touches the surface under investigation [21]. During each gentle tap, lateral forces are minimized, and this provides high spatial resolution and overall noninvasiveness of the measurement. DNA must be well attached to a smooth, flat substrate, so to obtain well-contrasted, well-resolved topographic maps. At the same time, the sample preparation procedure must preserve as much as possible the native DNA characteristics, so to avoid trapping the molecules in out-of-equilibrium configurations. We present all the steps required for extracting quantitative information on the structural and configurational properties of a population of DNA molecules: the preparation of DNA samples on mica surface; the imaging in air by AFM in Tapping Mode; the preparation of images for the analysis and the digitization of the DNA traces; the calculation of selected statistical parameters describing the state of the system (contour length, radius of gyration, and end-to-end distance). The approach here described is rather general and modular, and can be easily implemented to add the calculation of other statistical descriptors of the DNA structural properties, such as the persistence length or the bending angle distribution. The protocol here presented can be easily adapted to other imaging modes and conditions.

---

## 2 Materials

1. Cyanoacrylate glue (such as Loctite 406 or similar).
2. Steel magnetic disks, diameter 11–15 mm, thickness 0.2–0.5 mm.

3. Mica disks of the highest quality grade (VI, ruby muscovite), with thickness 0.2–0.6 mm and diameter 6–12 mm. (*see Notes 1 and 2*).
4. Ethanol (laboratory grade,  $\geq 95\%$  v/v).
5. Eppendorf tubes, 1.5 mL.
6. Plastic tubes, 50 mL.
7. Pipettes with plastic tips (p20, p200, and p1000).
8. Scotch-like adhesive tape (such as the Magic Transparent Tape).
9. Ultrapure MilliQ (typical resistivity  $\rho = 18.2 \text{ M}\Omega \text{ cm}$  at  $25 \text{ }^\circ\text{C}$ ) to prepare samples and solution for AFM imaging.
10. Nitrogen from a reservoir (purity  $\geq 99.999\%$ ).
11. Blotting paper to remove drops of water after rinsing.
12. Tweezers.
13. AFM tips for dynamic (or Tapping) Mode (*see Note 3*).
14. Deposition buffer: 2–5 mM  $\text{MgCl}_2$ , 10 mM NaCl, 10 mM HEPES-Na pH 7.5 in MilliQ  $\text{H}_2\text{O}$  (*see Note 4*).
15. Stock DNA in MilliQ water (at least  $0.05 \text{ ng}/\mu\text{L}$ ).

---

### 3 Methods

#### 3.1 Sample Preparation

##### 3.1.1 Substrate Preparation

A mica disk must be glued to a rigid support, whose dimensions and geometry can vary according to the instrument specifications. In most cases this support can be a metal disk with diameter 11–15 mm (at least 2 mm larger than the mica disk) and thickness 0.5–1 mm (*see Note 5*).

1. Clean the mica and the metallic surfaces with ethanol and dry all surfaces using blotting paper.
2. Glue the mica disk to the steel disk using cyanoacrylate-based adhesive (*see Note 6*).
3. Prepare at least 3–4 similar substrates, to be able to prepare many samples in parallel for quickly testing different deposition conditions.

##### 3.1.2 DNA Deposition on Mica

1. Freshly cleave the mica surface to be used for DNA deposition (*see Note 7*).
2. Dilute DNA to a final concentration of  $0.05 \text{ ng}/\mu\text{L}$  in the deposition buffer (so to have typically a 1–2 nM DNA solution) (*see Note 8*).
3. Put a 15–20  $\mu\text{L}$  drop of diluted DNA solution on the freshly cleaved mica (*see Note 9*). Avoid touching the mica surface with the pipette tip.

4. Incubate at room temperature for 2–10 min (longer incubation times will provide higher surface densities of DNA molecules).
5. Gently rinse the sample dropwise several times, using each time up to 1 mL of MilliQ water to remove the exceeding and loosely bound molecules. During the rinsing procedure, keep the sample tilted to help water slipping away.
6. Blot the remaining water placing the corner of a piece of blot paper next to the lower mica disk edge. Make sure the paper is not in contact with the surface.
7. Dry the sample under a gentle stream of clean nitrogen from a reservoir. Set a mild stream flux and keep the nozzle a few centimeter away from the surface (*see Note 10*).
8. The DNA sample is ready to be imaged (*see Note 11* for troubleshooting).

### 3.2 AFM Imaging in Air in Tapping Mode

AFM imaging is aimed at collecting a statistically meaningful number of high-quality, well-contrasted, high-resolution images of DNA molecules, which must then be digitized so to produce a large collection of molecular traces to be further analyzed. The general requirements are:

1. to collect in a single scan a reasonable number of molecules, so to obtain a good statistical sample (several hundred molecules) within 5–10 images;
2. to have a good sampling resolution overall in each image, in order to calculate accurate values of conformational parameters from the molecular traces.

Additional information can be found in **Note 12**.

Typical surface density of molecules on good samples is 15–40 molecules/ $\mu\text{m}^2$ , depending on the DNA length. Imaging of DNA samples in air can be performed in dynamic mode (usually called Tapping Mode, intermittent-contact or oscillating mode). Recently, new imaging modes based on a vertical tip-sample approach have been developed, which provide accurate control of applied force, low-invasiveness in air as well as in liquid and high spatial resolution (described briefly in **Note 13**); these modes could represent an alternative to the Tapping Mode, that we will consider hereafter.

We suggest the following scanning parameters and conditions for imaging DNA in air in Tapping Mode with a well-calibrated instrument (*see Note 14* on AFM calibration):

1. Mount a rigid cantilever for dynamic or Tapping Mode.
2. Set a free oscillation amplitude (target amplitude) of 10 nm or less. Smaller amplitudes provide high-quality and less-invasive imaging conditions, but require very clean, nonadhesive surfaces.

3. Initially, after engaging the tip on the sample surface, adjust the minimal amplitude (force) setpoint to track the surface while keeping the scan size to  $<1$  nm.
4. Initially, set a small scan size (100–500 nm) and optimize the gains and the amplitude setpoint to achieve optimal tracking conditions at the lowest applied force (*see Note 17*). Good tracking is witnessed by a good overlap of topographic profiles in both scan directions. Check the tracking in correspondence of points where a sudden change in surface slope occurs (typically, at the mica–DNA border). Poor tracking results in a loss of contact when the tip is crossing the DNA molecule downhill. Increase gradually the scan size up to  $2\ \mu\text{m} \times 1\ \mu\text{m}$  (aspect-ratio 2:1, *see Note 15*).
5. Set the sampling resolution to  $2048 \times 512$  points (number of points per line  $\times$  number of lines). This choice provides a sampling resolution of 1 nm/pixel and 2 nm/pixel in the fast and in the slow scan directions, respectively.
6. Set the scan rate to 1–4 Hz (*see Note 16*).
7. We suggest acquiring up to five images in each location, according to the simple scheme described below, then to withdraw the tip and engage some 100–500  $\mu\text{m}$  away. Three–four different macroscopic locations will provide several hundred molecules for the statistical analysis. The image acquisition scheme is the following: in each location, acquire the first image with no offsets in  $X$  and  $Y$  directions, at  $(0,0)$ . Acquire the other images at points  $(X_0,0)$ ,  $(-X_0,0)$ ,  $(Y_0,0)$ ,  $(-Y_0,0)$ , set  $X_0 = 3\ \mu\text{m}$ ,  $Y_0 = 2\ \mu\text{m}$ , so to avoid overlap among the scan areas (*see Note 18*).
8. In order to maintain stable imaging conditions, in particular to minimize capillary adhesion at the tip–sample interface, the AFM head and the sample with the scanning stage can be hosted in a small chamber, inside which a dry  $\text{N}_2$  atmosphere is maintained, with relative humidity below 5%.

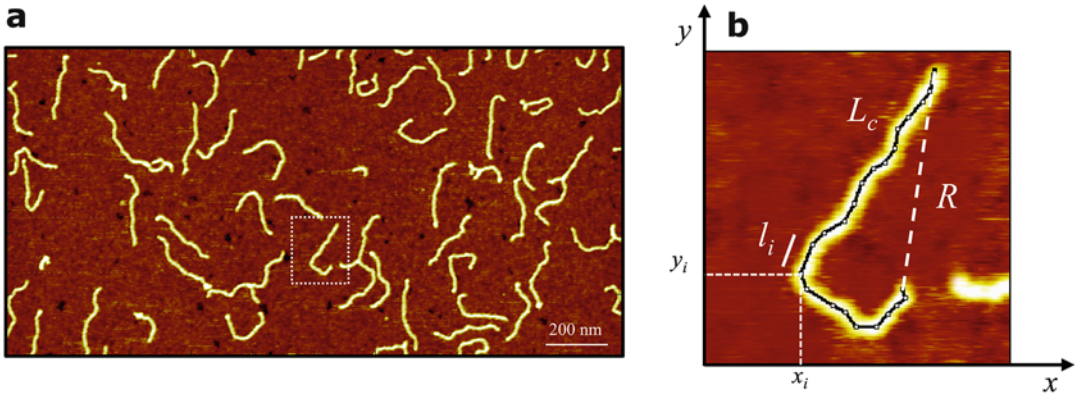
Figure 1a shows a typical topographic map of DNA molecules (727 bp) on mica, imaged in air in Tapping Mode, according to the described methodology.

### 3.3 Data Analysis (See Note 19)

#### 3.3.1 Image Preprocessing

Images must be prepared for the analysis, by removing standard artifacts related to the image formation process and by removing high-frequency noise, typically related to the feedback loop operation. Artifacts typically manifest themselves as baselines superimposed to the true topographic profiles. Identification and subtraction of the baseline from each topographic profile is essential for the accurate analysis of the AFM images. Details on the origin of artifacts and on the baseline subtraction procedures are provided in **Note 20**.





**Fig. 1** 727 bp DNA molecules on mica and details of the single-molecule analysis of AFM images. **(a)** Top-view AFM image showing molecules equilibrated on a mica surface, after gentle dehydration of the sample (the vertical range is 1 nm; heights increase from dark to bright colors). **(b)** A representative DNA molecule with highlighted the relevant parameters used for the characterization of configurational properties, as described in Subheading 3.3

1. Apply a global plane-fitting of the first order to the image, so to get rid of the global tilt of the sample. This operation will level the height values in the image, offering a complete overview of the molecules and of the overall image quality.
2. Apply a line-by-line flattening to the image by subtracting higher-order polynomials (usually up to the third order is enough) through suitable masks excluding features that do not belong to the reference substrate.
3. Check the quality of the flattening by looking at the height histogram of the image: a sharp Gaussian peak centered around the average height  $z_0$  of the substrate (typically  $z_0 = 0$  nm) with FWHM of a few Å should be present; DNA molecules will typically contribute a broader short tail in the height distribution.
4. Apply a median filter (with  $3 \times 3$ , maximum  $5 \times 5$  kernel) to the flattened image, to smooth high-frequency noise.

### 3.3.2 Tracing DNA Molecules

Once the images have been flattened and smoothed, the set of spatial coordinates  $\{x_i, y_i\}_{i=1:N}$  defining the backbone of each molecule for each particular experimental condition (i.e., for each particular length of the DNA molecules) must be determined. Here  $N$  is the total number of molecules in all the AFM images that can be used for the analysis. The tracking can be done manually or by means of (semi)automatic algorithms; some of them are freely available upon request to the developers (*see Note 21*). Here we describe how to manually trace the molecules using ImageJ/Fiji, an open-source software written in Java and supported by a broad community of scientists.

1. Export topographical maps in a format compatible with the ImageJ software (i.e., .tif). Resize the image to include the scanned area only.
2. Import the AFM image into ImageJ/Fiji and define image size and sampling resolution.
  - (a) Define the image size using the command: Image → Scale, and set width and height (in pixels).
  - (b) Define the scale using the command: Analyze → Set Scale, and set the scale (in nanometer).
3. Use the “Segmented Line” command to draw by hand the backbone of the molecule. Keep possibly a constant distance between each point that should not be too far or too close from each other. Consider to maintain a point-to-point distance from 1/2 to 2/3 of the molecule width.
4. Save the molecule trace using the command: “Save As *XY* Coordinates”. With this operation ImageJ/Fiji saves an ASCII (.txt) file containing the *XY* coordinates of the selected track, columnwise.

An example of the manual tracking of a DNA molecule is shown in Fig. 1b.

### 3.3.3 Basic Structural Analysis

Once the spatial coordinates of each molecule have been obtained, several statistical parameters, describing the structural and configurational properties of DNA in the studied conditions, can be determined (see among the others [5, 6, 8, 9, 11, 15, 22]). First, the contour length of DNA depends on the conformation adopted. Generally, DNA assumes the B-form in physiological conditions, but it can adopt other different conformations, like A- and Z-form. Moreover, DNA can display more open or compact configurations depending on its persistence length, which in turn is sensitively dependent on the ionic strength of the buffer and on the nature of the surrounding ions. Here we focus our attention on a selection of parameters that can be readily obtained by the *XY* coordinates: the contour length  $L_c$ , the end-to-end distance  $R$ , the gyration radius  $R_g$  [5, 8, 22, 23]. All these parameters depend on DNA form, base-pair composition, and persistence length; therefore, their accurate statistical determination can provide important information on the structural properties of a DNA population under study. Typically, the configurational parameters are evaluated from digitized traces, and then the average values with standard deviations are calculated from the distributions of these parameters.

1. Calculate  $L_c$ ,  $R$ , and  $R_g$  according to the following equations (refer to Fig. 1b; see also Note 22):

$$L_c = \sum_{i=1}^{N-1} l_i = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (1)$$

$$R = \sqrt{(x_N - x_1)^2 + (y_N - y_1)^2} \quad (2)$$

$$R_g = \sqrt{1/N \sum_{i=1}^N r_i^2} \quad (3)$$

where:

$$r_i^2 = (x_i - x_{CM})^2 + (y_i - y_{CM})^2 \quad (4)$$

$$(x_{CM}, y_{CM}) = 1/N \left( \sum_{i=1}^N x_i, \sum_{i=1}^N y_i \right) \quad (5)$$

In Eq. 5  $(x_{CM}, y_{CM})$  are the coordinates of the center of mass of the molecule.

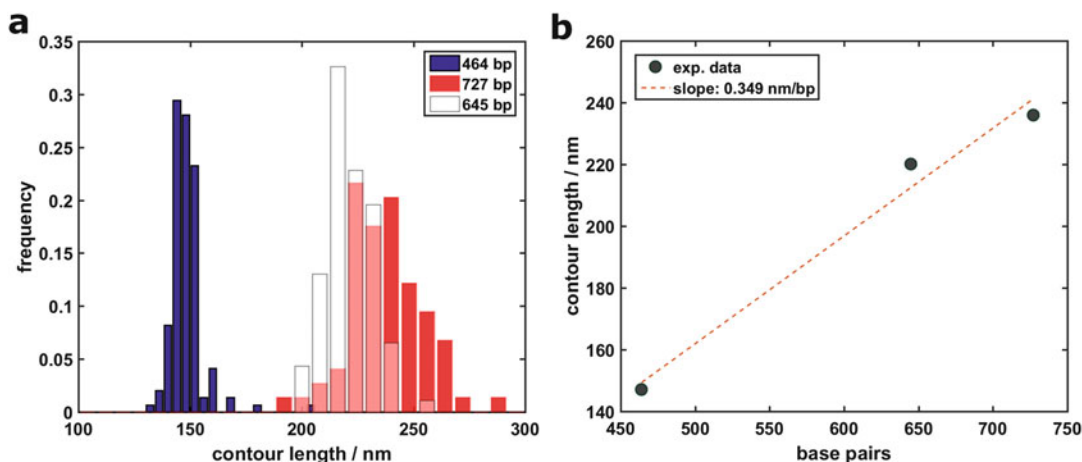
2. Calculate the histograms of the values of the above observables, in order to represent their statistical distributions. Always look first at distributions, and then calculate suitable estimators of the true values and their dispersions.
3. Calculate mean (median) values, standard deviations and standard deviations of the mean, or other suitable statistical estimators, depending on the particular statistical properties of the given observable [5].

An example of the analysis performed on DNA according to the presented protocol is shown in Fig. 2.

---

## 4 Notes

1. The typical substrate used for DNA immobilization is mica, a mineral belonging to the sheet silicate groups. There are many varieties of mica and Muscovite is the most used form. It is constituted by tetrahedral sheets of  $(\text{Si,Al})_2\text{O}_5$  ionically linked by a central layer of  $\text{Al}_2(\text{OH})_2$  [24]. The net negative charge of the basal oxygen between these double layers is balanced by a layer of hexagonally coordinates cations ( $\text{K}^+$  in Muscovite). This negatively charged layer becomes exposed after the standard cleavage procedure and the dissociation of  $\text{K}^+$  ions. The most prominent characteristic of mica is the nearly perfect cleavage, due to its intrinsic atomic structure [25]. Due to its chemical composition, the outmost layer of mica after cleavage



**Fig. 2** Distribution of contour length ( $L_c$ ) values and assessment of the form of DNA molecules upon dehydration on mica, according to the proposed protocol. Three populations of DNA molecules have been investigated, with lengths 464 bp, 645 bp, and 727 bp, respectively. (a) The distribution of contour length values is calculated according to Eq. 1. About 150, 80, and 80 molecules have been traced, respectively, for the 464 bp, 645 bp, and 727 bp populations. The measured average lengths agree with those expected for the B-form of DNA within 5–10%. (b) A linear fit of the curve  $L_c$  vs. bp provides a value of the rise per residue parameter of 0.349 nm/bp, confirming unambiguously that DNA molecules, despite the dehydration, are in the B form. The discrepancy between absolute observed values and the expected ones could be due to the partial transition towards the A-form that DNA faces when deposited on mica [22], with the A-form domains likely located at the DNA free ends [8]

is negatively charged in humid air and in particular in water. Freshly cleaved mica could hence provide negative, ultra-flat and clean surfaces, functional for high-quality AFM measurements [26]. By using divalent positive ions ( $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Mn^{2+}$ ,  $Ni^{2+}$ , and  $Zn^{2+}$ ) [7], or molecules carrying a positive charge, as in the case of natural or artificial polyamines (poly-L-lysine and poly-L-ornithine [27, 28]) it is possible to bind the negatively charged DNA backbone to the mica surface, to the purpose of imaging DNA by AFM.

2. A cheaper alternative to pre-cut mica (and Teflon) disks is represented by mica (and Teflon) sheets of the same quality and thickness, from which disks of the desired diameter can be obtained using a hole punch.
3. Rigid cantilevers for dynamic modes must be used. Typical parameters characterizing these cantilevers are: resonance frequency  $f \approx 300$  kHz; single-crystal silicon tips with radius of curvature  $R < 10$  nm; force constant  $k \approx 40$  N/m; optionally, a gold or aluminium reflective coating on the back of the cantilever.
4. All the stock saline solutions are prepared starting from the powder that is dissolved in MilliQ water. Filtering should

possibly be avoided, as in general the entities that are removed are very large compared to DNA and would not represent a serious issue. The sample has to be as clean as possible in order to achieve the best condition of imaging, and clean buffers are a fundamental prerequisite. For this reason all dilutions and samples must be prepared using ultrapure MilliQ water based buffers. Never use simple (bi)distilled water, and never sterilize buffers.

5. If imaging in liquid is envisaged, it can be useful to put in between the mica disk and the metal support a Teflon spacer, about 2 mm larger than the mica disk and up to 2 mm thick, aimed at blocking the spread of the liquid used as imaging buffer, so that a stable droplet is obtained for in-liquid imaging. In this case, it is better to use different adhesive depending on the surfaces to be bound together. Mica on Teflon: two-component epoxy glue; Teflon on metal disk: cyanoacrylate glue. This metal–Teflon–mica substrate can be used of course also for imaging DNA in air.
6. A tiny amount of glue must be used, initially placed in the middle of the disk. Apply a gentle pressure so that the glue is distributed uniformly in between the mica disk and the support. Carefully dose the amount of glue so to avoid it spreading outside the mica disk area; this will likely cement the mica layers together from the side and will disturb the stripping procedure.
7. Mica disks should be always freshly cleaved immediately before the deposition of DNA, using soft adhesive tape to peel the topmost layers away. In this way it is possible to create a flat atomically smooth clean surface free of contaminants. To this purpose, firmly attach the scotch tape over the supported mica surface and remove it so to peel the topmost layers away. Repeat the operation using a clean portion of tape until a uniform thin circular layer remains on it (repeat in any case 2–3 times). To achieve a homogeneous stripping no air bubbles must be present below the adhesive tape. Apply a constant tension to the adhesive tape while peeling the mica. Remove the tape in one continuous movement.
8. DNA must be used freshly prepared, resuspended in MilliQ water. For best results, it can be stored at 4 °C for possibly no more than 1–2 weeks. Repeated thawing and freezing damages the DNA backbone [29].
9. The layer used for the imaging has to be cleaved just before sample deposition. It is a good practice to protect the surface from the contact with air, if it is not used immediately.
10. When significant amount of water remains on the surface upon drying, in the form of water islands, it may be the case that something went wrong either during the deposition of DNA

or the cleavage of the mica disk. Sticky/dirty/overcoated surfaces retain water, indeed. Such surfaces may result from the deposition of contaminated/degraded solutions or from remnants of the scotch tape adhesive.

11. In principle, the incubation time of the DNA solution on mica determines the number and density of molecules on the mica surface [5, 10, 12], and typically a few minutes are enough to obtain optimal imaging and analysis conditions. It may happen however that poor reproducibility and deviations from the expected behavior during sample preparation are observed. This problem typically occurs when the DNA concentration in solution and the incubation time are changed in the effort of obtaining the desired density of molecules. Usually, this anomalous behavior is also accompanied by the poor quality of the deposited molecules (condensed in blobs, totally or partially, or with small blobs at the free ends; aggregated or associated in complex two-dimensional structures or networks; etc.). All these can be symptoms that either the molecules in solution, or the buffer, or both, have some problems. For instance, a bad PCR reaction can produce weak DNA molecules with open, and therefore, sticky ends; because of the intrinsic DNA complementarity, several molecule ends will anneal between each other assembling networks, which will be mostly washed away, but also remain on the mica surface to some extent. A dirty buffer containing nanoscopic contaminants (*see Notes 4 and 8*) can promote DNA denaturation or aggregation, with similar effects. Large complexes will be typically washed away, as mentioned; therefore, only the minoritarian fraction of small objects will remain on the surface, with a density largely independent on both initial DNA concentration and incubation time. If similar issues are faced, it is usually wise to first prepare fresh clean buffers, then if needed fresh DNA stocks. Although only rarely observed in our experience, similar problems can also be due to poor quality of the mica surface (only use mica of the highest grade) or to issues in mica cleavage (including those due to the poor quality of the scotch tape, which can leave residues of glue on the mica surface).
12. AFM measurements can be carried out in air, as described in the present manuscript, as well as in a suitable saline buffer. Imaging DNA in liquid does not necessarily provides more accurate information, as long as one focuses on structural data, because the latter mostly rely on the accurate characterization of lengths along the DNA backbone, rather than on the measurement of heights. Tip convolution affects only at minor extent such measurements, at least for relatively long molecules (>100 bp). Moreover, there is evidence that dehydration required for sample preparation has little impact on the DNA

properties, probably due to the fact that the truly DNA–mica interface remains always partially hydrated (by a few monolayers of water) [5, 9, 28]. Nevertheless, the fine conformational changes induced by dehydration, as well as by the different sample preparation methods and choices of the bridging cations are still matter of discussion [8, 22]. The general methodology presented here applies irrespective to the imaging method adopted (air, liquid, Tapping Mode, Peak Force Tapping, etc.). For some specific indications in the case of imaging in liquid, *see* **Note 5**.

13. Imaging methods based on the vertical approach of the tip towards the sample have recently been introduced [30], and represent valid alternatives to dynamic (tapping) modes and contact modes, especially in liquid. The general idea, besides specificities related to the different implementations, is to record a set of force curves on a grid spanning the scan area, with a carefully controlled maximum force setpoint; during imaging in fluid, the maximum force can be kept on the 10–100 pN level, while lateral forces are minimized thanks to the vertical approach mode, similarly to Tapping or other dynamic modes. In Tapping Mode, however, peak forces are significantly higher.
14. Calibration of the piezo-scanners should be checked periodically (every 6–12 months) by imaging the surface of a calibration grating, with repeated morphological features of appropriate dimensions (in the present case, the  $XY$  period should be 1  $\mu\text{m}$ , the depths 10–200 nm). Unless a certified grating is used, the  $XYZ$  accuracy after the calibration procedure can be reasonably assumed to be  $\leq 2\%$ .
15. The number of scanned lines impacts on the acquisition time of a single image. Setting an aspect-ratio 2:1 allows reducing the number of scanned lines, i.e., the acquisition time, without affecting dramatically the image resolution in the slow scan direction. Given a target sampling resolution in the image, the scanned area can be kept large, for the sake of a better statistics, by increasing the scan size and the aspect-ratio value, and setting suitable values of the points per line and number of lines parameters. The main limit to be considered is that the scanning speed cannot grow arbitrarily, otherwise the feedback loop will not be able to track accurately the surface.
16. Scanning speed (which changes when either scan rate or scan size are changed) should not be too small, otherwise drifts could produce significant distortions in the image. At the same time, too high scan speeds will challenge the feedback loop of the AFM and determine inaccurate tracking.
17. Typically the integral gain has the highest impact on image quality. Each instrument has its own gain optimal settings.



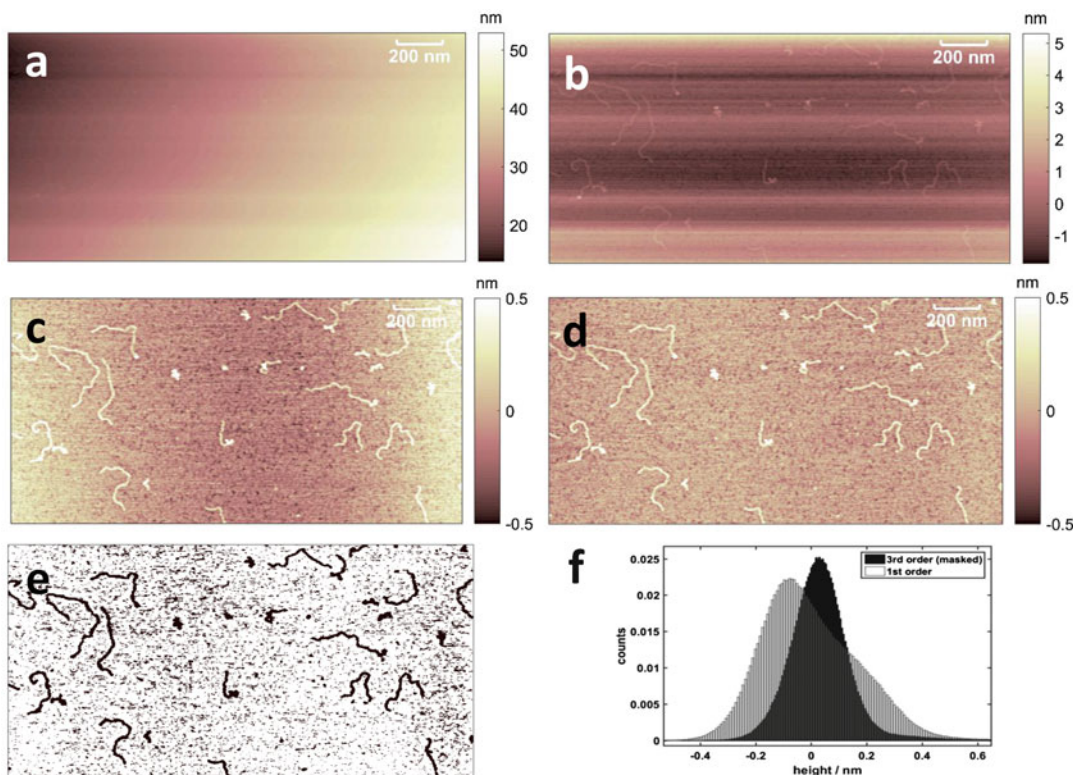
Increase the integral gain until oscillations appear in the height and error signals. Then decrease the integral gain just below the critical value. Increase the proportional gain until the quality of the image starts worsening. Setting high gains gives advantages in terms of tracking stability and scanning velocity that overcompensates the introduction of high-frequency noise in the image; the latter can be effectively removed a posteriori by applying a median filter with a  $3 \times 3$  kernel on the image.

18. Drifts must be minimized in order to obtain accurate topographic maps. Drifts can be due to thermal equilibration of the system components (sample, laser, cantilever, electronics as well as the scanner) or to mechanical hysteresis of the piezo elements. When a different scan area is selected by applying offsets to the piezo, the latter will typically keep some memory of the previous static deformation, resulting in a constant drift across the new image. In order to minimize this effect and completely refresh the scanner motion, after setting the new offsets it is effective to reduce significantly the number of lines and complete a few low-resolution images moving quickly up and down, until the hysteresis is lost. Also reverting a few times the slow scan direction (up-bottom/bottom-up) helps removing the hysteresis of the piezo scanner.
19. The data analysis procedures described in general terms in this manuscript can be implemented by means of custom routines, as well as by means of commercial and open-source software. Basic image processing tools are typically included in the control software of the AFM. The authors have developed their own libraries of data analysis routines in the MATLAB environment (many research groups use their own libraries). Some open-source software packages are listed below (this list is neither meant to be complete, nor is it expected to remain up to date for a long time):
  - (a) ImageJ/Fiji, <http://rsb.info.nih.gov/ij/>
  - (b) Gwyddion, <http://gwyddion.net/>
  - (c) WSxM, <http://www.wsxmsolutions.com/>
  - (d) FiberApp, <http://www.fsm.ethz.ch/publications-list/software.html>
  - (e) Image SXM, <https://www.liverpool.ac.uk/%7Esdb/ImageSXM/>
20. Mica is atomically smooth and overall flat. Deviations from a flat baseline can therefore be attributed to scanning artifacts, with the obvious exception of those due to the presence of a DNA molecule. A global tilt of the sample adds a linear baseline to each profile, and globally a plane to the topographic map as a whole. Tubular scanners will add an approximately parabolic



baseline to each profile (a bow), because the displacement of the sample placed on top of the tube follows a curved trajectory rather than a linear one (ideally an arc of circumference). Generally, each scanner will add its own polynomial distortion to the AFM topographic map. Drifts of different nature can add additional shifts between adjacent profiles (approximately constant along each profile, i.e., along the fast-scan direction, but appreciable along the slow-scan direction); the presence of these line-by-line deviations usually requires the application of line-by-line polynomial subtraction (also known as flattening) rather than simply the subtraction of a two-dimensional plane, or paraboloid, or higher-order surface (also known as plane-fitting). It is essential, in order to accurately determine the image baseline, to mask (i.e., not to consider in the polynomial fit) all the surface features that do not belong to the flat reference substrate (the DNA molecules, surface defects, ...). Masks are typically built by thresholding (flooding) algorithms, determining image segmentation; only the substrate is considered for the flattening. The absence of masking in the fitting procedure will introduce artifacts in the topographic maps (the ubiquitous black stripes), because in the presence of bumps and/or depressions the fitted polynomial typically deviates from the baseline. Masking is typically a feature of the analysis software. The flattening process is described in Fig. 3.

21. A semiautomatic tracing algorithm is in principle preferable to the manual tracing of DNA molecules, as the latter is more prone to introduce bias from the operator. An automatic algorithm can of course introduce systematic errors, although these would be the same for all molecules, irrespective to the operator. Ivan Usov's FiberApp is a free comprehensive suite of (MATLAB) routines with a GUI for "Tracking and Analyzing Polymers, Filaments, Biomacromolecules, and Fibrous Objects" (available at <http://www.fsm.ethz.ch/publications-list/software.html> [31]).
22. The formulas reported in the text for the calculation of basic structural parameters provide in general accurate results as long as the digitization of the molecules in the image is not poor, which in the case of DNA means that the point-to-point distance should be of the order of 0.5–1 nm. Separations significantly smaller than 0.5–1 nm are unreasonable, considered that the nominal width of DNA is 2 nm; oversampling can introduce spurious high-frequency components to the trace, that will affect the angle distribution as well as the calculated contour length. The contour length  $L_c$  is the most sensitive parameter, indeed. Optimized estimators for  $L_c$  have been developed, and can be used instead of Eq. 1 [32–34].



**Fig. 3** Overview of the image pre-processing procedure. The process starts with (a) a raw AFM image, where the sample tilt and line-by-line distortions hinder the target topographical features (the DNA molecules in this case); after (b) a global plane-fitting of the first order, and a series of line-by-line flattening of the (c) first, and (d) third order, the baseline is effectively removed and the molecules emerge, well-contrasted with respect to the smooth, flat substrate. In (e) the mask built to apply the third-order flattening is shown. This logical mask assigns a value of 1 to the points that must be considered for the fitting, i.e., those belonging to the substrate, and 0 elsewhere. In (f) the distributions of surface heights after the first and third order flattening are compared. A well-shaped, nearly symmetric dominant mode, representing the height values of the substrate, is typical of a properly flattened image

---

## Acknowledgments

We thank Francesca Borghi for support in AFM analysis. A.P. thanks the Dept. of Physics of the University of Milano for financial support under the project “Piano di Sviluppo dell’Ateneo per la Ricerca 2014. Linea B: Supporto per i giovani ricercatori”. Work in M.M-F lab is supported by AIRC (n.15631) and Telethon (GGP15227).

## References

- Alessandrini A, Facci P (2005) AFM: a versatile tool in biophysics. *Meas Sci Technol* 16: R65–R92. doi:10.1088/0957-0233/16/6/R01
- Kalle W, Strappe P (2012) Atomic force microscopy on chromosomes, chromatin and DNA: a review. *Micron* 43:1224–1231. doi:10.1016/j.micron.2012.04.004
- Podgornik Rudolf (2011) Physics of DNA. <http://www.fmf.uni-lj.si/~podgornik/download/physics-of-DNA-1.1.pdf>
- Pinheiro AV, Han D, Shih WM, Yan H (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat Nanotechnol* 6:763–772. doi:10.1038/nnano.2011.187
- Rivetti C, Guthold M, Bustamante C (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J Mol Biol* 264:919–932. doi:10.1006/jmbi.1996.0687
- Podestà A, Indrieri M, Brogioli D, Manning GS, Milani P, Guerra R, Finzi L, Dunlap D (2005) Positively charged surfaces increase the flexibility of DNA. *Biophys J* 89:2558–2563
- Pastré D, Piétrement O, Fusil S, Landousy F, Jeusset J, David M-O, Hamon L, Le Cam E, Zozime A (2003) Adsorption of DNA to mica mediated by divalent counterions: a theoretical and experimental study. *Biophys J* 85:2507–2518. doi:10.1016/S0006-3495(03)74673-6
- Japaridze A, Vobornik D, Lipiec E, Cerreta A, Szczerbinski J, Zenobi R, Dietler G (2016) Toward an effective control of DNA's submolecular conformation on a surface. *Macromolecules* 49:643–652. doi:10.1021/acs.macromol.5b01827
- Buzio R, Repetto L, Giacomelli F, Ravazzolo R, Valbusa U (2014) Symmetric curvature descriptors for label-free analysis of DNA. *Sci Rep* 4:6459. doi:10.1038/srep06459
- Bustamante C, Rivetti C (1996) Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscope. *Annu Rev Biophys Biomol Struct* 25:395–429. doi:10.1146/annurev.bb.25.060196.002143
- Valle F, Favre M, De Los RP, Rosa A, Dietler G (2005) Scaling exponents and probability distributions of DNA end-to-end distance. *Phys Rev Lett* 95:158105. doi:10.1103/PhysRevLett.95.158105
- Fang Y, Spisz TS, Wiltshire T, D'Costa NP, Bankman IN, Reeves RH, Hoh JH (1998) Solid-state DNA sizing by atomic force microscopy. *Anal Chem* 70:2123–2129. doi:10.1021/ac971187o
- Santos S, Stefancich M, Hernandez H, Chiesa M, Thomson NH (2012) Hydrophilicity of a single DNA molecule. *J Phys Chem C* 116:2807–2818. doi:10.1021/jp211326c
- Thomson NH, Santos S, Mitchenall LA, Stuchinskaya T, Taylor JA, Maxwell A (2014) DNA G-segment bending is not the sole determinant of topology simplification by type II DNA topoisomerases. *Sci Rep* 4:6158. doi:10.1038/srep06158
- Wiggins PA, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, Widom J, Dekker C, Nelson PC (2006) High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat Nanotechnol* 1:137–141. doi:10.1038/nnano.2006.63
- Savelyev A, Materese CK, Papoian GA (2011) Is DNA's rigidity dominated by electrostatic or nonelectrostatic interactions? *J Am Chem Soc* 133:19290–19293. doi:10.1021/ja207984z
- Lia G, Indrieri M, Owen-Hughes T, Finzi L, Podestà A, Milani P, Dunlap D (2008) ATP-dependent looping of DNA by ISWI. *J Biophotonics* 1:280–286
- Fang Y, Hoh JH, Spisz TS (1999) Ethanol-induced structural transitions of DNA on mica. *Nucleic Acids Res* 27:1943–1949. doi:10.1093/nar/27.8.1943
- Parker SCJ, Margulies EH, Tullius TD (2008) The relationship between fine scale DNA structure, GC content, and functional elements in 1% of the human genome. *Genome Inform* 20:199–211
- Knips A, Zacharias M (2015) Influence of a cis, syn-cyclobutane pyrimidine dimer damage on DNA conformation studied by molecular dynamics simulations. *Biopolymers* 103:215–222. doi:10.1002/bip.22586
- García R, Perez R (2002) Dynamic atomic force microscopy methods. *Surf Sci Rep* 47:197. doi:10.1016/S0167-5729(02)00077-8
- Rivetti C, Codeluppi S (2001) Accurate length determination of DNA molecules visualized by atomic force microscopy: evidence for a partial B- to A-form transition on mica. *Ultramicroscopy* 87:55–66. doi:10.1016/S0304-3991(00)00064-4
- Shi Y (1996) Statistical mechanics of the extensible and shearable elastic rod and of DNA. *J*

- Chem Phys 105:714–731. doi:[10.1063/1.471927](https://doi.org/10.1063/1.471927)
24. Kuwahara Y (1999) Muscovite surface structure imaged by fluid contact mode AFM. *Phys Chem Miner* 26:198–205. doi:[10.1007/s002690050177](https://doi.org/10.1007/s002690050177)
  25. Ostendorf F, Schmitz C, Hirth S, Kühnle A, Kolodziej JJ, Reichling M (2008) How flat is an air-cleaved mica surface? *Nanotechnology* 19:305705. doi:[10.1088/0957-4484/19/30/305705](https://doi.org/10.1088/0957-4484/19/30/305705)
  26. Thundat T, Allison DP, Warmack RJ, Brown GM, Jacobson KB, Schrick JJ, Ferrell TL (1992) Atomic force microscopy of DNA on mica and chemically modified mica. *Scanning Microsc* 6:911–918
  27. Bussiek M, Mücke N, Langowski J (2003) Polylysine-coated mica can be used to observe systematic changes in the supercoiled DNA conformation by scanning force microscopy in solution. *Nucleic Acids Res* 31:e137. doi:[10.1093/nar/gng137](https://doi.org/10.1093/nar/gng137)
  28. Podesta A, Imperadori L, Colnaghi W, Finzi L, Milani P, Dunlap D (2004) Atomic force microscopy study of DNA deposited on poly L-ornithine-coated mica. *J Microsc* 215:236–240. doi:[10.1111/j.0022-2720.2004.01372.x](https://doi.org/10.1111/j.0022-2720.2004.01372.x)
  29. Bustamante C, Vesenka J, Tang CL, Rees W, Guthold M, Keller R (1992) Circular DNA molecules imaged in air by scanning force microscopy. *Biochemistry* 31:22–26
  30. Adamcik J, Berquand A, Mezzenga R (2011) Single-step direct measurement of amyloid fibrils stiffness by peak force quantitative nano-mechanical atomic force microscopy. *Appl Phys Lett* 98:193701. doi:[10.1063/1.3589369](https://doi.org/10.1063/1.3589369)
  31. Usov I, Mezzenga R (2015) FiberApp: an open-source software for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects. *Macromolecules* 48:1269–1280. doi:[10.1021/ma502264c](https://doi.org/10.1021/ma502264c)
  32. Sanchez-Sevilla A, Thimonier J, Marilley M, Rocca-Serra J, Barbet J (2002) Accuracy of AFM measurements of the contour length of DNA fragments adsorbed on mica in air and in aqueous buffer. *Ultramicroscopy* 92:151–158. doi:[10.1016/S0304-3991\(02\)00128-6](https://doi.org/10.1016/S0304-3991(02)00128-6)
  33. Rivetti C (2009) A simple and optimized length estimator for digitized DNA contours. *Cytom Part A* 75:854–861. doi:[10.1002/cyto.a.20781](https://doi.org/10.1002/cyto.a.20781)
  34. Gomez AI, Cruz M, Cruz-Orive LM (2016) On the precision of curve length estimation in the plane. *Image Anal Stereol* 35:1. doi:[10.5566/ias.1412](https://doi.org/10.5566/ias.1412)

## Genome-Wide Quantitative Fitness Analysis (QFA) of Yeast Cultures

Eva-Maria Holstein, Conor Lawless, Peter Banks, and David Lydall

### Abstract

We provide a detailed protocol for robot-assisted, genome-wide measurement of fitness in the model yeast *Saccharomyces cerevisiae* using Quantitative Fitness Analysis (QFA). We first describe how we construct thousands of double or triple mutant yeast strains in parallel using Synthetic Genetic Array (SGA) procedures. Strains are inoculated onto solid agar surfaces by liquid spotting followed by repeated photography of agar plates. Growth curves are constructed and the fitness of each strain is estimated. Robot-assisted QFA, can be used to identify genetic interactions and chemical sensitivity/resistance in genome-wide experiments, but QFA can also be used in smaller scale, manual workflows.

**Key words** SGA, QFA, High-throughput, Yeast, Genetic interaction, Genome-wide, Fitness, Robotic, Quantitative, Population model, Image analysis, Growth curve

---

### 1 Introduction

Comparing how well specific strains grow in a carefully chosen environment or genetic background can reveal how different genes interact in living cells. For instance, deletion of *EXO1*, affecting a nuclease, or *RAD9*, affecting a checkpoint protein, each suppress telomere induced growth defects of budding yeast *cdc13-1* mutants [1]. Quantitative Fitness Analysis (QFA) allows genome-wide identification of similar suppressing and enhancing genetic interactions [2, 3]. For example, QFA revealed that nonsense mediated mRNA decay proteins affect the fitness of *cdc13-1* mutants in a way similar to Exo1 and Rad9 [2]. QFA can also be used for drug interaction screens [4].

QFA evolved from qualitative spot test assays that were manually performed and scores were not photographed [5]. The procedure we describe here is now high-throughput automated, documented, and quantitative. However, we and others, also apply some of the procedures on lower scale, manually performed experiments [3].

Typically, to perform a genome-wide QFA, we begin with Synthetic Genetic Array (SGA), crossing a query strain, containing one or more mutations, to a library of single gene deletions and DAmP (Decreased Abundance by mRNA Perturbation) alleles [2, 6]. Strains derived from SGA are grown to saturation in liquid medium, in 96-well plates. Liquid cultures are next spotted onto solid agar plates in 384-format arrays and repeatedly photographed during growth. The images are analyzed to generate quantitative cell density estimates which are used to create growth curves and ultimately to measure the fitness of each culture. This approach allows fitness of thousands of strains grown in parallel to be measured and compared. We describe detailed protocols for SGA and QFA using query strains containing a single gene deletion (*abcΔ*), a point mutation (*xyz-1*) or both (*abcΔ xyz-1*), any of which can grow on the same selection media. Since we often cross temperature sensitive mutations to libraries, our SGA protocols have been optimized to work at 20 °C.

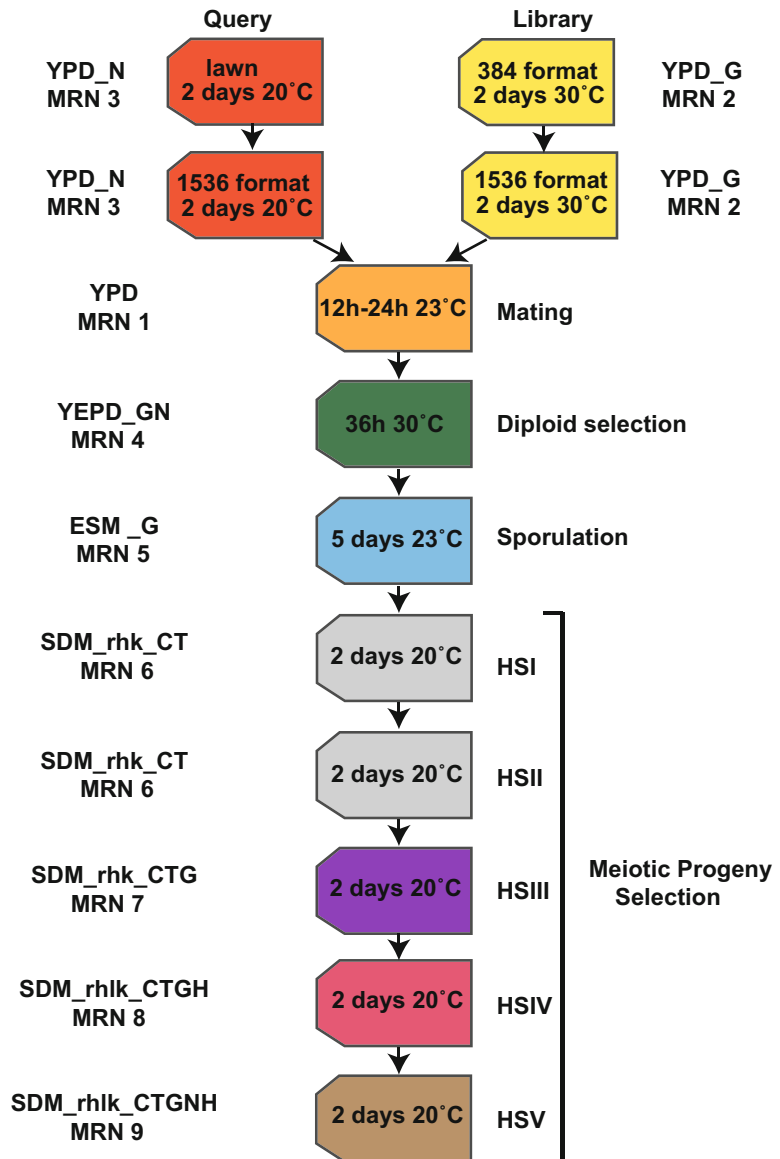
---

## 2 Materials

### 2.1 Stock Solutions and Media

Figure 1 gives an overview of the SGA procedure and in particular the media used at each stage, as indicated below (MRN = Media recipe number).

1. Ethanol: 70% in water.
2. Geneticin (G418): dissolve in water at 200 mg/mL, filter-sterilize and store in 1 mL aliquots at −20 °C.
3. Nourseothricin (clonNAT): dissolve in water at 100 mg/mL, filter-sterilize and store in 1 mL aliquots at −20 °C.
4. Hygromycin B: dissolve in water at 300 mg/mL, filter-sterilize and store in 1 mL aliquots at −20 °C.
5. Canavanine: dissolve in water at 100 mg/mL, filter-sterilize and store in 0.5 mL aliquots at −20 °C.
6. Thialysine: dissolve in water at 100 mg/mL, filter-sterilize and store in 0.5 mL aliquots at −20 °C.
7. Amino-acid supplement powder mixture for synthetic media (CSM DO): 3 g adenine, 2 g uracil, 2 g inositol, 0.2 g para-aminobenzoic acid, 2 g alanine, 2 g arginine, 2 g asparagine, 2 g aspartic acid, 2 g cysteine, 2 g glutamic acid, 2 g glutamine, 2 g glycine, 2 g histidine, 2 g isoleucine, 2 g leucine, 2 g lysine, 2 g methionine, 2 g phenylalanine, 2 g proline, 2 g serine, 2 g threonine, 2 g tryptophan, 2 g tyrosine, and 2 g valine.
8. Drop-out (DO) powder mixture: combination of the ingredients described in Subheading 2.1, item 7 minus the appropriate supplement: rhk DO: −Arg/His/Lys. rhlk DO: −Arg/



**Fig. 1** Flowchart illustrating the steps of the SGA procedure. Query strain and library strains are pinned in a 1536 format, followed by mating, diploid selection, and induction of sporulation. The meiotic progeny undergoes several selection stages to select for double mutants expressing the STE2pr-SP-his5 reporter, *LEU2* and resistance to canavanine and thialysine (*can1Δ* and *lyp1Δ*), G418, cloNAT and Hygromycin B

His/Leu/Lys. For 1 L of medium: 2 g of the DO powder mixture.

- Amino-acid supplement for sporulation medium (ESM DO): 2 g histidine, 10 g leucine, 2 g lysine, and 2 g uracil. For 1 L: 0.1 g of amino-acid supplement powder mixture.



10. Dextrose: 40% in water, autoclave and store at room temperature.
11. Adenine: 0.5% in water, autoclave and store at room temperature.
12. YEPD (YPD, liquid): For 1 L add 10 g yeast extract, 20 g peptone to 935 mL. Autoclave (121 °C for 15 min), cool down to ~65 °C and add 50 mL of 40% dextrose solution and 15 mL of 0.5% adenine solution, gently invert bottle until mixture is homogenous and store at room temperature. (**MRN 1**).
13. YEPD (YPD, solid): For 1 L add 10 g yeast extract, 20 g peptone, and 20 g agar to 935 mL. Autoclave (121 °C for 15 min), cool down to ~65 °C, add 50 mL of 40% dextrose solution and 15 mL of 0.5% adenine solution, gently invert bottle until mixture is homogenous and pour plates. (**MRN 1**).
14. YEPD + G418 (YPD\_G) for library strain selection: For 1 L add 1 mL of 200 mg/mL G418 stock solution (final concentration 200 mg/L) to YEPD medium (Subheading 2.1, **item 13**), gently invert bottle until the mixture is homogenous and pour plates. (**MRN 2**).
15. YEPD + clonNAT (YPD\_N, solid) for query strain selection: For 1 L add 1 mL of 100 mg/mL clonNAT stock solution (final concentration 100 mg/L) to YEPD medium (Subheading 2.1, **item 13**), gently invert bottle until the mixture is homogenous and pour plates. (**MRN 3**).
16. YEPD + G418/clonNAT (YPD\_GN, solid) for diploid selection: For 1 L add 1 mL of 200 mg/mL G418 stock solution (final concentration 200 mg/L) and 1 mL of 100 mg/mL clonNAT stock solution (final concentration 100 mg/L) to autoclaved YEPD medium (Subheading 2.1, **item 13**), gently invert bottle until the mixture is homogenous and pour plates. (**MRN 4**).
17. Enriched sporulation medium (ESM\_G, solid) for sporulation: For 1 L add 10 g potassium acetate, 1 g yeast extract, 0.5 g dextrose, 0.1 g amino-acid supplement powder mixture for sporulation to 500 mL of water. In a separate bottle add 20 g agar to 500 mL water. Autoclave both solutions at 121 °C for 15 min. Add the agar to the media, cool down to ~65 °C and add 250 µL of G418 stock solution (final concentration 50 mg/L), mix thoroughly and pour plates. (**MRN 5**).
18. (SD/MSG) –His/Arg/Lys + canavanine/thialysine (SDM\_rhk\_CT, solid) for haploid I and II (HS I and HS II) selection stages: For 1 L add 1.7 g yeast nitrogen base, 1 g MSG, 2 g amino-acid supplement powder mixture (rhk DO), 450 mL water and autoclave at 121 °C for 15 min. In a separate bottle add 20 g agar to 500 mL water and autoclave at 121 °C



for 15 min. Combine autoclaved solutions, add 50 mL 40% glucose, cool medium to ~65 °C, add 0.5 mL canavanine (100 mg/L) and 0.5 mL thialysine (100 mg/L) stock solutions, mix thoroughly and pour plates. (**MRN 6**) (*see Note 1*).

19. (SD/MSG) –His/Arg/Lys + canavanine/thialysine/G418 (SDM\_rhk\_CTG, solid) for HS III selection stage: same as MRN 6 (Subheading 2.1, **item 18**) with 1 mL G418 (200 mg/L) stock solution added before mixing and pouring plates (**MRN 7**).
20. (SD/MSG) –His/Arg/Lys/Leu + canavanine/thialysine/G418/Hygromycin B (SDM\_rhlc\_CTGH, solid) for HS IV selection stage: For 1 L add 1.7 g yeast nitrogen base, 1 g MSG, 2 g amino-acid supplement powder mixture (rhlc DO), 450 mL water and autoclave at 121 °C for 15 min. In a separate bottle add 20 g agar to 500 mL water and autoclave at 121 °C for 15 min. Combine autoclaved solutions, add 50 mL 40% glucose, cool medium to ~65 °C, add 0.5 mL canavanine (100 mg/L), 0.5 mL thialysine (100 mg/L), 1 mL Hygromycin B (300 mg/L) and 1 mL G418 (200 mg/L) stock solutions, mix thoroughly and pour plates. (**MRN 8**).
21. (SD/MSG) – His/Arg/Lys/Leu + canavanine/thialysine/G418/clonNAT/Hygromycin B (SDM\_rhlc\_CTGNH, solid) for HS IV selection stage and QFA spotting: same as MRN 8 (Subheading 2.1, **item 20**) with 1 mL clonNAT (100 mg/L) stock solution added before mixing and pouring plates (**MRN 9**).
22. (SD/MSG) –His/Arg/Lys/Leu + canavanine/thialysine/G418/clonNAT/Hygromycin B (SDM\_rhlc\_CTGNH, liquid) for QFA: For 1 L add 1.7 g yeast nitrogen base, 1 g MSG, 2 g amino-acid supplement powder mixture (rhlc DO), 900 mL water and autoclave at 121 °C for 15 min. Add 50 mL 40% glucose, cool medium to ~65 °C, add 0.5 mL canavanine (100 mg/L), 0.5 mL thialysine (100 mg/L), 1 mL clonNAT (100 mg/L), 1 mL Hygromycin B (300 mg/L), and 1 mL G418 (200 mg/L) stock solutions, mix thoroughly and store at room temperature (**MRN 9**).

## 2.2 Plates and Tubes

1. Rectangular polycarbonate Greenlab SBS permaplates (<https://sites.google.com/site/greenlabuk/permaplates>). These plates are stackable and autoclavable. After use, the agar and cultures are discarded, disinfected using Virkon (1%, at least 30 min), washed in water, and autoclaved at 121 °C for 15 min. Alternatively, disposable OmniTrays (Nunc) can be used.
2. 96-well, flat bottom culture plates and low evaporation lids.
3. 15 mL glass test tubes with Bacti-Caps.
4. Single deep well reservoir containing 96 pyramidal bottoms.

- 2.3 Robotic Systems**
1. S&P Robotics Inc BM3-SC with 96, 384, and 1536 pin robotic pin tools.
  2. Beckman Biomek FX with 96 pin robotic pin tool (V&P Scientific magnetic mounting pin tool, 2 mm pin diameter).
  3. S&P Robotics automated imager.

- 2.4 Software**
1. Colonyzer image analysis software: <http://research.ncl.ac.uk/colonyzer/>
  2. QFA R package: <http://qfa.r-forge.r-project.org/>

**2.5 Strains for SGA**

The SGA protocol is based on that of Tong and Boone [6]. We have adapted it in two major ways. First, because we often use temperature-sensitive mutants, many of the steps are performed at 20 °C. Secondly, we use extra selectable markers. All starter strains contain three markers *LEU2*, *natMX*, and *hphMX* as well as *can1Δ* and *lyp1Δ* to allow us to assess query strains with up to three gene deletions. For example, query strains containing recessive ts mutations are flanked by two selectable markers (*LEU2* and *hphMX*). For SGAs where we use query strains with a single gene replaced with the *natMX* cassette, we still ensure that *LEU2* and *hphMX* are present in the query strains. By completing all experiments on media containing the same selection agents, we can more closely match environments between QFA experiments. Matching environments is important for detecting genetic or chemical interactions. For example, in order to calculate genetic interaction strength (GIS) [2], a control *CDC13* strain and an experimental *cdc13-1* query strain, containing the same set of markers, are crossed to the same library, on the same media, before comparing fitnesses at the same temperatures [2].

- 2.5.1 Libraries**
1. Single gene deletion yeast library with each gene disruption being replaced with the antibiotic resistance cassette *kanMX* [6].
  2. Decreased Abundance by mRNA Perturbation (DAmP) yeast library with the 3'-UTR being disrupted with the antibiotic resistance cassette *kanMX* [7].

**2.5.2 Universal Control Strain**

A universal control strain, with *LEU2*, *hphMX* and *natMX*, is often crossed to the library to measure the effects of library mutations on fitness under different conditions (e.g., of temperature or chemicals). This allows for direct comparison of fitness between strains containing a query mutation and control strains lacking the query mutation on the same media.

DLY9326: *MATalpha can1Δ::STE2pr-Sp-his5 lyp1Δ::hphMX::LEU2::natMX his3Δ leu2Δ ura3Δ met15Δ*

2.5.3 *Single Gene Deletion Starter Strain*  
(e.g., *abcΔ::natMX*)

Single genes are deleted with *natMX*.

Source strain: DLY7325: *MATalpha can1Δ::STE2pr-Sp-his5 lyp1Δ::hphMX::LEU2 his3Δ1 leu2Δ ura3Δ met15Δ*

Example query strain: DLY10276: *MATalpha can1Δ::STE2pr-Sp\_his5 lyp1Δ::hphMX::LEU2 smt3Δ::natMX his3Δ1 leu2Δ ura3Δ met15Δ*

2.5.4 *Point Mutation Starter Strain (LEU2::xyz-1::hphMX)*

Point mutation, and if necessary control WT allele, is flanked with *LEU2* and *hphMX*. Source strain: DLY7329: *MATalpha can1Δ::STE2pr-Sp-his5 lyp1Δ::natMX his3Δ1 leu2Δ ura3Δ met15Δ*

Example query strain: DLY8205: *MATalpha can1Δ::STE2pr-Sp-his5 LEU2::cdc13-1::hphMX lyp1Δ::natMX his3Δ1 leu2Δ ura3Δ met15Δ*

2.5.5 *Point Mutation and Single Gene Deletion Starter Strain (LEU2::xyz-1::hphMX abcΔ::natMX)*

Point mutation, and if necessary control WT allele, is flanked with *LEU2* and *hphMX*. Single gene is replaced with *natMX*.

Source strain: DLY3798: *MATalpha lyp1Δ can1Δ::STE2pr-his5Δ his3Δ1 leu2Δ ura3Δ met15Δ*

Example query strain: DLY6722: *MATalpha LEU2::cdc13-1::hphMX exo1Δ::natMX lyp1Δ can1Δ::STE2pr-his5Δ his3Δ1 leu2Δ ura3Δ met15Δ*

---

## 3 Methods

### 3.1 Sterilization Procedure for Pin Tools Before Each Pinning or Spotting Step

#### 3.1.1 S&P Robotics Inc BM3-SC

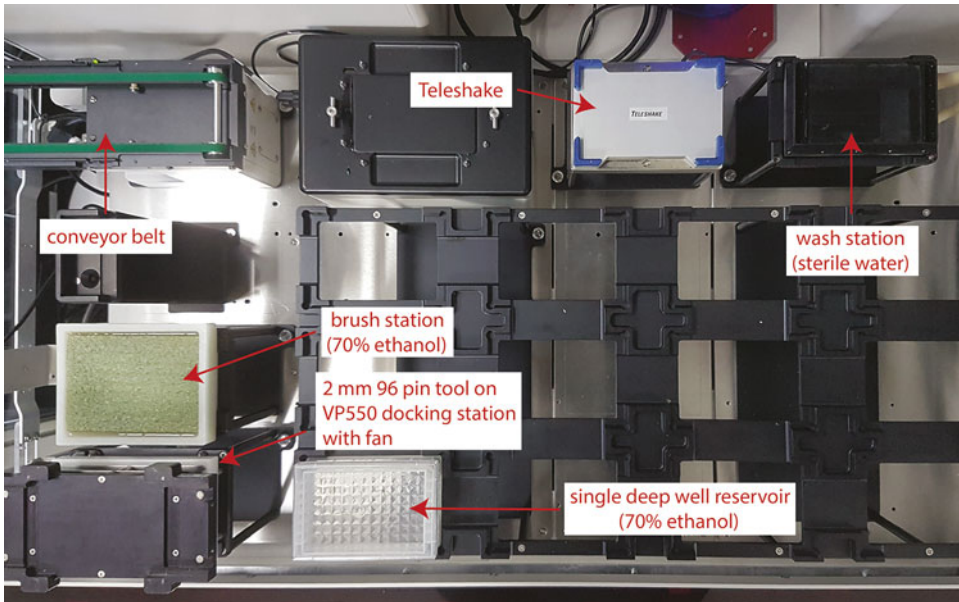
1. Rinse the pin tool for 10 s in revolving water reservoir filled with sterile water.
2. Move the pin tool to the revolving brush station containing 350 mL 70% ethanol for 20 s.
3. Sonicate the pin tool for 20 s in the sonicator containing 350 mL 70% ethanol.
4. Allow pin tool to dry over the fan for 40 s.

#### 3.1.2 Beckman Biomek FX

1. Rinse the pin tool in the wash station filled with sterile water (*see Fig. 2*).
2. Transfer the pin tool to the brush station, and rotate for two full circles in 70% ethanol.
3. Let the pin tool sit in the single deep well reservoir with 96-well pyramidal bottoms containing 70% ethanol.
4. Allow pin tool to dry over the fan on the VP550 docking station for 45 s.

### 3.2 SGA

Each 384 format library plate contains neutral strains (e.g., *his3Δ*) around the edges of the plate because edge cultures have a growth advantage and because image analysis is more difficult at plate edges. Therefore there are 308 experimental deletion mutants per



**Fig. 2** Layout of the FX robot deck prior to spotting

384-format plate. These strains are pinned to a 1536 format, resulting in four replicates for each genotype. Figure 1 gives an overview of the various stages of the SGA procedure to generate double mutants.

### 3.2.1 Revival of Libraries from Frozen Stocks

1. Let the plates thaw completely for 1 h at room temperature.
2. Wipe the 384-well library plates dry with disposable paper. Take the plastic lid off, wipe the inside with 70% EtOH, aseptically remove the covering film from the frozen 384-well plates and put the plastic lid back onto the plates. (*see Note 2*).
3. Replicate the glycerol stocks of yeast from the 384-well plates onto solid YEPD + G418 (YEPD\_G, Media recipe no. 2) agar plates using the 384 pin robotic pin tool on the BM3.
4. Reseal the 384-well plates with fresh sealing film using a rubber roller and return to  $-80^{\circ}\text{C}$ .
5. Incubate the transferred colonies on the YEPD + G418 plates at  $30^{\circ}\text{C}$  for 2 days.

### 3.2.2 Documenting SGA Procedures

1. Each SGA screen is assigned a unique screen number in a database. In addition, the database documents starter strain number and genotype, library used, screen description and the name of the person conducting the screen are stored. *See Fig. 3* for details.
2. For each step of the SGA procedure the media used, date, media batch number as well as the incubation temperature is

Screen Number 71 Starter Strain Number 8767 Name of Client (Initials) MJG Date Created 16/10/2013  
 Name of PI DAL Name of User (Initials) APB  
 Starter Strain Genotype MAT alpha LEU2::stn1-13::hphMX lyp1::natMX can1delta::STE2pr-Sp\_his5 his3delta1 leu2delta0 ura3delta0 met15delta0  
 Screen Name stn1-13 Organism cerevisiae Starting Library SDLV4\_1536  
 Screen Description matt crossing the ts stn1-13

	Date Pinned	Media Batch No.	Batch Passed	Temperature
SGA Selection 1 YPD_G	2013/10/16	282	Yes	30
SGA Selection 2 YPD	2013/10/20	330	Yes	23
SGA Selection 3 ESM_G	2013/10/24	317	Yes	23
SGA Selection 4 SDM_rhk_CT	2013/11/29	262	Yes	20
SGA Selection 5 SDM_rhk_CT	2013/12/02	262	Yes	20
SGA Selection 6 SDM_rhk_CTG	2013/12/04	295	Yes	20
SGA Selection 7 SDM_rhIk_CTGH	2013/12/06	325	Yes	20
SGA Selection 8 SDM_rhIk_CTGNH	2013/12/10	326	Yes	20
SGA Selection 9				
SGA Selection 10				
Comments				

**Fig. 3** Representation of SGA metadata in a database. A unique screen number is assigned to each SGA screen (e.g., screen number 71). Starter strain number and genotype, library, screen description and initials of three people engaged with the screen are recorded. For each SGA stage, media used, date of the experiment conducted, media batch number, and incubation temperature are recorded. Any comments or problems arising during the SGA procedure can also be stored in the database

documented. Any problems or comments arising during any of the steps of the SGA procedure can be stored in the database (*see Note 3*). *See Fig. 3* for an example screenshot of the database we currently use.

- Each SGA stage (mating, diploid selection, sporulation, and haploid selection I–V) is imaged in an S&P Robotics spImager before the next pinning step takes place. Images are saved with systematic, fixed length filenames (e.g., K000134\_027\_011\_2012-05-05\_12-22-56.JPG) including incubator ID (K in the example filename), unique batch ID (000134), incubator temperature (027), plate number (011), date and time (2012-05-05\_12-22-56).

### 3.2.3 Cultivation of the SGA Query Strain and Library

- Pool several colonies of the query strain and inoculate them into two 15 mL glass tubes, each containing 5 mL of YEPD. (*see Note 4*).
- Allow the strains to grow on a wheel overnight at 20 °C.
- Remove 10 µL with a sterile tip and check by microscopy for contamination. Discard contaminated cultures.
- Pour the culture of one tube over a rectangular YEPD plate containing cloNAT (**MRN 3**) and ensure that the culture is spread evenly across the plate by gentle agitation.
- Transfer the liquid to a second plate by pouring and spread the culture evenly again. Repeat for a third, fourth and fifth plates. The lawns are allowed to grow for 2 days at 20 °C.

6. Repeat **steps 4** and **5** depending on how many lawns you require (*see Note 4*).
7. Pick the three most evenly grown lawns and robotically pin the lawn containing the query strain in a 1536 format onto 6 fresh YEPD + clonNAT plates per lawn (**MRN 3**) using the 1536 pin robotic pin tool and incubate at 20 °C for 2 days.
8. Pin the library from the 384 format to a 1536 format onto fresh YEPD + G418 plates using the 384 pin robotic pin tool (**MRN 2**) and incubate at 30 °C for 2 days.

### 3.2.4 *Mating the Query Strain with the Library*

1. Pin the query strains (1536 format) onto a YEPD plate using a 1536 pin robotic pin tool.
2. Pin the library (1536 format) on top of the pinned query strains on the YEPD plates from **step 1** using the 1536 pin robotic pin tool.
3. Incubate the YEPD plates containing the query strains and the DMA array strains at 23 °C for 12–24 h.

### 3.2.5 *MATa/alpha Diploid Selection and Sporulation*

1. To select for diploids, resulting *MATa/alpha* zygotes are pinned onto YEPD + G418/clonNAT (**MRN 4**) plates using the 1536 pin robotic pin tool.
2. Incubate the plates at 30 °C for 1–2 days (ideally 36 h) (*see Note 5*).
3. To induce sporulation, pin the resulting diploid cells onto enriched sporulation medium containing G418 (**MRN 5**) using the 1536 pin robotic pin tool.
4. Incubate the plates at 23 °C for 5 days (*see Note 6*).
5. Pick cells from the edge of a colony on one of the sporulation plates and distribute them in 10 µL of water on a cover slide. Examine for tetrads under a light microscope and determine the tetrad frequency. A frequency of >1% is sufficient.
6. Repeat **step 5** for 2–3 plates to ensure that all cultures sporulated.

### 3.2.6 *MATa Meiotic Progeny Selection (Haploid Selection I + II)* (*See Note 7*)

1. For the first round of haploid selection (HSI), pin sporulated cells onto (SD/MSG) – His/Arg/Lys + canavanine/thialysine plates (**MRN 6, HSI**) using the 1536 pin robotic pin tool.
2. Incubate the pinned plates at 20 °C for 2 days.
3. Repeat **steps 1** and **2** for a second round of haploid selection (HSII).

### 3.2.7 *MATa kanMX Meiotic Progeny Selection (Haploid Selection III)* (*See Note 8*)

1. Pin the resulting cells from Subheading **3.2.6** onto (SD/MSG) – His/Arg/Lys + canavanine/thialysine/kanMX plates (**MRN 7, HS III**) using the 1536 pin robotic pin tool.
2. Incubate the pinned plates at 20 °C for 2 days.



**3.2.8** *MATa kanMX/hphMX Meiotic Progeny Selection (Haploid Selection IV)* (See **Note 9**)

1. Pin the cells from Subheading **3.2.7** onto (SD/MSG) – His/Arg/Lys/Leu + canavanine/thialysine/kanMX/hphMX plates (**MRN 8, HS IV**) using the 1536 pin robotic pin tool.
2. Incubate the pinned plates at 20 °C for 2 days.

**3.2.9** *MATa kanMX/hphMX/natMX Meiotic Progeny Selection (Haploid Selection V)* (See **Note 10**)

1. Pin the cells from Subheading **3.2.8** onto (SD/MSG) – His/Arg/Lys/Leu + canavanine/thialysine/G418/clonNAT/Hygromycin B plates (**MRN 9, HS V**) using the 1536 pin robotic pin tool.
2. Incubate the pinned plates at 20 °C for 2 days.

**3.3 QFA**

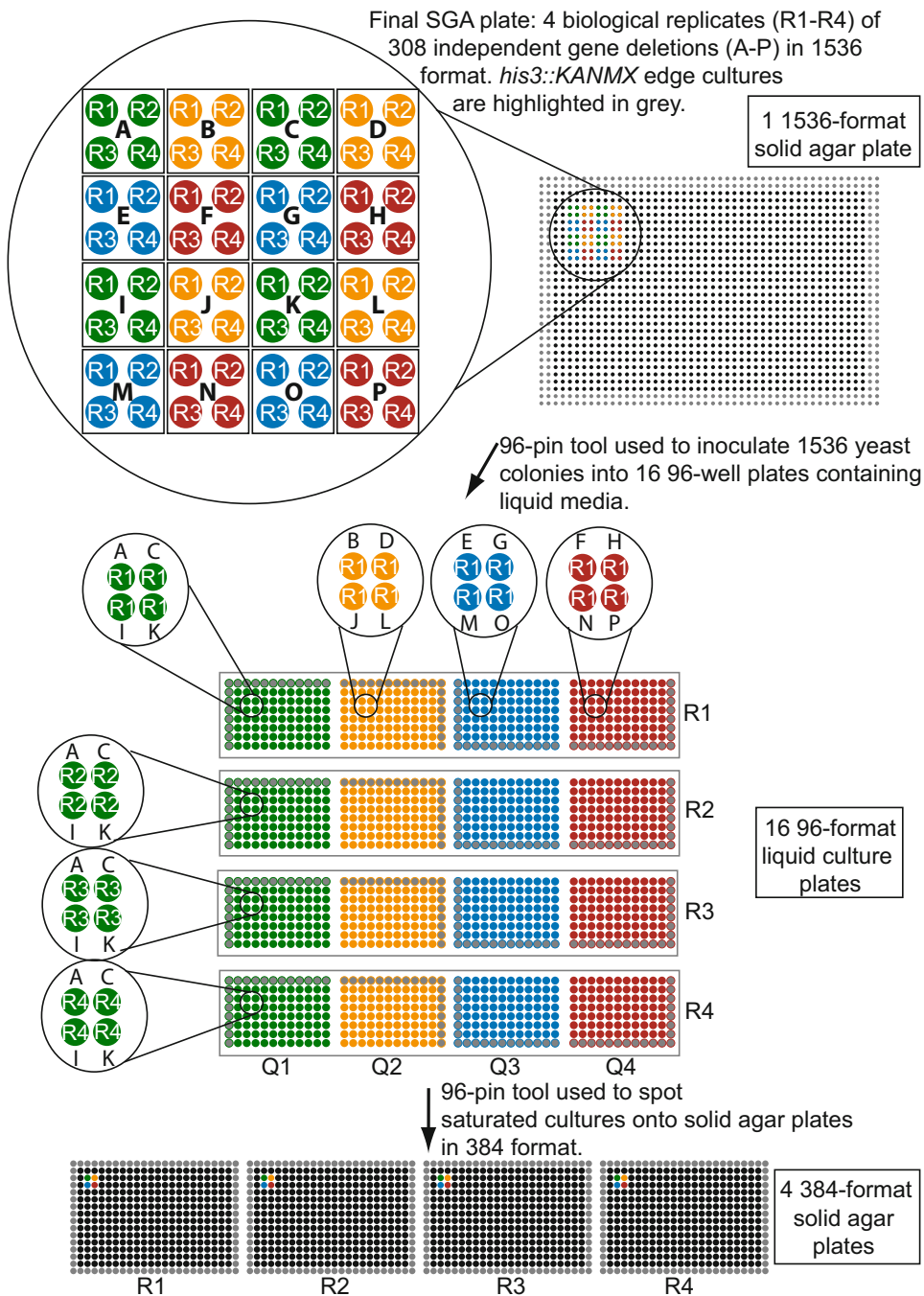
**3.3.1** *Cultivation of Yeast Strains in Liquid Media*

SGA plates in 1536 format contain four biological replicates (R1–R4) of 308 independent gene deletions per plate as well as *his3::kanMX* on the edges of the plate (Fig. 4). All strains are cultivated in 96-well culture plates in liquid and grown until saturation prior to spotting. Using a 96-pin tool, 1536 strains from a final SGA plate are transferred to 16 96-well culture plates using a BM3 robot. Figure 4 gives an overview of robotic inoculation from 1536 format to 96 format and the colony positions in the 96-well format. The sterilization procedure for the BM3 is described in Subheading **3.1.2**.

1. Fill each well of 96-well culture plates (16 96-well culture plates per 1536 format SGA plate) with 200 µL of SDM\_rh1k\_CTGNH media (**MRN 9**). (see **Notes 11–14**).
2. Load the source (solid agar plates in 1536 format) and destination plates (96-well culture plates) onto the stackers of the BM3 robot. A schematic overview of the order of plates in the stackers is shown in Fig. 5.
3. Using a 1 mm 96-pin tool, transfer 384 strains (R1) from each final 1536 format SGA solid agar plate into four 96-well culture plates (Fig. 4).
4. After the transfer, visually check each 96-well culture plate to confirm that each inoculated well contains a white dot (pooled cells from the pin) at the bottom of the plate. Missing empty wells in the 96-well culture plate should correspond to missing colonies on the SGA plate.
5. Place the inoculated 96-well culture plates in zip lock plastic bags. Seal bags and place in a temperature controlled incubator at 20 °C for 2 days (see **Note 15**).

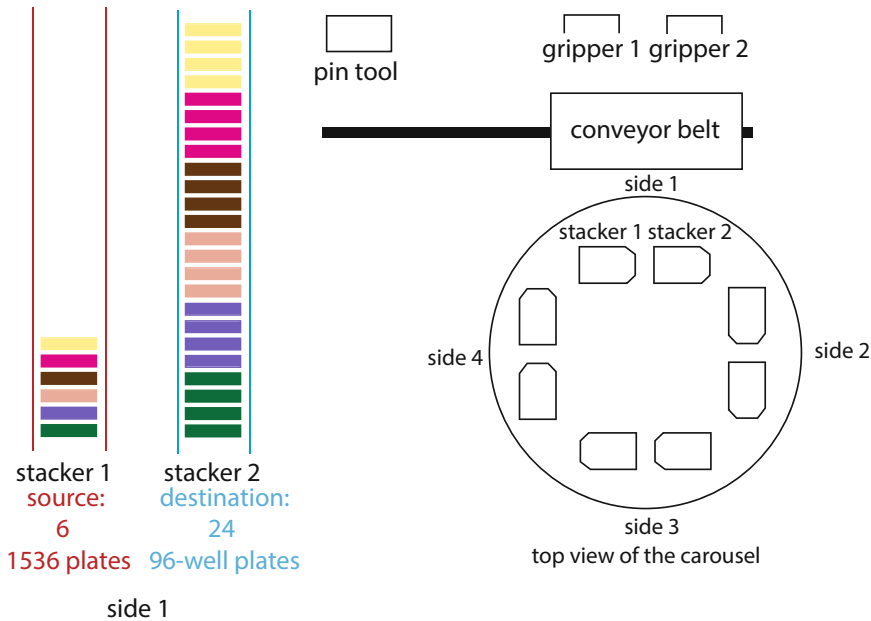
**3.3.2** *Spotting Yeast Cultures*

Cultures are spotted onto solid agar plates in 384 format using a 2 mm 96-pin tool on a Beckman Biomek FX with automated carousel (allows handling of many plates). 96-well culture plates and solid agar plates are placed in carousel stacker. Saturated cultures in 96-well plates are resuspended by orbital shaking at



**Fig. 4** Schematic overview of the QFA spotting procedure: inoculation from 1536 colony plate to 96-well liquid plates to spotting in 384 format. Strains grown in 1536 format on solid agar plates are inoculated into 96-well culture plates. Individual genotypes are indicated by letters (A–P). At least four independent replicates (R1–R4) of each genotype are present on each 1536 plate. Using a 96-pin tool, every fourth colony in a row and column (e.g., A(R1), C(R1), I(R1), K(R1)) is picked and transferred to a 96-well culture plate. Four 96-well culture plates (Q1–Q4) are required to inoculate one replicate of each genotype (e.g., all R1s). Each color (green, yellow, blue, and red) represent one of the four quadrants. To transfer all 1536 colonies of one final SGA plate, 16 96-well culture plates are required. Strains are grown to saturation (e.g., 2 days at 20 °C). All four quadrants (Q1–Q4) from one repeat (e.g., R1) are spotted onto the same solid agar plate in 384 format in the same pattern as the final 1536 format SGA plate





**Fig. 5** Overview of plate loading in the BM3 stacker. For each 1536 format source plate, four 96-well culture destination plates containing 200  $\mu$ L of selective media are placed on the stacker, allowing the transfer of one replicate of each genotype into liquid media. Each stacker permits the transfer from six 1536 format solid agar plates to 24 96-well culture plates. The 96-well culture plates corresponding to each 1536 source plates are indicated by the same color

1000 rpm for 20 s on a deck-mounted Variomag Teleshake. The pin tool is cleaned and sterilized prior to each spotting (see Subheading 3.1.2). A typical deck layout is shown in Fig. 2 (see Notes 16–18).

1. Place the 96-well culture plates from Subheading 3.3.1 in the stacker of the Biomek FX. Cut off corners of the 96-well culture plates face inward.
2. Place sterile rectangular solid agar plates (**MRN9**) in the stacker. Cut off corners of GreenLab plates (on short edge of one side of the plate) face inward and lids upward. Cut off corners of Nunc plates (on long edge of one side of the plate) face right and lids upward.
3. Use a 2 mm 96 pin robotic pin tool to spot four 96-well culture plates onto rectangular solid agar plates in 384 format (Fig. 4) (see Note 19).
4. Spotted rectangular plates are transferred to an S&P Robotics automated imager. (see Note 20).

### 3.3.3 Imaging

High-throughput imaging takes place in an S&P Robotics automated imager. The robot repeatedly takes the plates from a temperature-controlled incubator, removes the plate lid and

transfers the plate to an evenly illuminated space underneath a digital single-lens reflex camera (DSLR) with a 60 mm macro lens mounted on it. The imager computer controls the camera to capture an image at high resolution. To allow more direct comparison between images, images are captured in Manual mode (i.e., fixed aperture, shutter speed, ISO and color temperature) and with autofocus switched off. The image file is stored with a systematic filename. After image capture the plate is re-lidded and returned to the incubator until the next time point. We typically grow cells at fixed temperatures ranging from 20 to 37 °C and take images every 4 h. (*see* **Notes 21** and **22**).

1. Set the incubator to the desired temperature.
2. Using the Robosoft software controlling the S&P robotics automated imaging system, enter the cycle interval of image capture, temperature, length of the experiment, batch number, and number of plates. Assign an Image Data Directory for storing plate images from the experiment.
3. Assign a unique barcode for each experiment. Image filenames are constructed as during SGA (*see* Subheading **3.2.2**, **step 3**).
4. Take the first image immediately after placing the solid agar plates in the incubator to allow for a zero time point growth measurement.
5. Plates are incubated and imaged for up to 5 days until colony growth saturation. (*see* **Note 23**).

### 3.3.4 Experimental Description File

Metadata about each experiment are recorded in a tab-delimited text file, with a row for every plate. This can be created programmatically or built by hand with any text editor or spread sheet software. Alternatively a preexisting file can be altered. Using Excel, for example, set up a spreadsheet using the headers detailed in Fig. 6 and **Note 24** (*see* **Note 24**). The file can then be saved as a tab-delimited file (File → Save As → Text (Tab-delimited) (\*.txt)).

### 3.3.5 Library Description File

The position of each genotype in the library is recorded in a tab-delimited Library Description file. This metadata file is required for data analysis in Subheading **3.3.6** and can be reused for multiple experiments. An example Library Description file is shown in Fig. 7 (*see* **Note 25**).

### 3.3.6 Data Analysis

Following repeated imaging of yeast cultures as they grow on solid agar plates (Fig. 8), Colonyzer is used to estimate culture cell densities from photographs [8]. Colonyzer locates cultures on plates and tracks them over time, correcting for lighting effects on plate images and estimates culture cell densities from the sum of pixel intensities within each spot. It also estimates culture area as well as the shape and color of each culture. Colonyzer software is

Barcode	Start.Time	Treatment	Medium	Screen	Library	Plate	RepQuad	Client	ExptDate	User	PI	Inoc	Condition
J000231_030_001	2013-12-11_10-45-01	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	1	1	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_002	2013-12-11_10-45-02	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	2	1	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_003	2013-12-11_10-45-03	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	3	1	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_004	2013-12-11_10-45-04	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	4	1	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_005	2013-12-11_10-45-05	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	5	1	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_006	2013-12-11_10-45-06	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	1	2	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_007	2013-12-11_10-45-07	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	2	2	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_008	2013-12-11_10-45-08	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	3	2	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_009	2013-12-11_10-45-09	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	4	2	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_010	2013-12-11_10-45-10	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	5	2	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_011	2013-12-11_10-45-11	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	1	3	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_012	2013-12-11_10-45-12	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	2	3	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_013	2013-12-11_10-45-13	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	3	3	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_014	2013-12-11_10-45-14	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	4	3	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_015	2013-12-11_10-45-15	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	5	3	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_016	2013-12-11_10-45-16	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	1	4	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_017	2013-12-11_10-45-17	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	2	4	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_018	2013-12-11_10-45-18	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	3	4	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_019	2013-12-11_10-45-19	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	4	4	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH
J000231_030_020	2013-12-11_10-45-20	30	SDM_rhik_CTGNH	cdc13-1	SDLV4_384	5	4	EMH	'2013/12/10'	EMH	DAL	CONC	SDM_rhik_CTGNH

**Fig. 6** Representative Experimental Description file. Experimental metadata for each plate in a QFA experiment are recorded in separate rows in an Experiment Description file. A typical, genome-wide experiment with four replicates of each spot includes 60 plates. In this example the experiment included 20 plates: all four replicate cultures from five library plates. See Subheading 3.3.4 for interpreting column labels and details on file construction

written in the Python programming language and once both are installed, Colonyzer can also be used as a command-line tool. Installation instructions and documentation can be found at <http://research.ncl.ac.uk/colonyzer/>

We have developed a software package for QFA to read metadata from Library Description and Experiment Description files and to associate metadata with each culture. The software, developed in the statistical modeling language R, can also be used to construct growth curves from cell density estimates and to fit population models (e.g., the logistic model) to curves, giving a range of fitness estimates, including intrinsic growth rate. The QFA R package also produces a wide range of visual reports. These quality control reports: growth curve plots annotated with estimated culture fitnesses, plots showing correlation between fitness estimates for replicate cultures, plots examining spatial patterns within and between plates and HTML reports containing annotated previews of raw image data underlying fitness estimates. More importantly, there are also functions for generating text reports tabulating fitness estimates for each culture and fitness summaries for replicates of each genotype examined.

By matching appropriately paired QFA screens: one control screen of a library of strains crossed with the universal control strain (see Subheading 2.5.2) and a second query screen containing the same strains in the control crossed with a query mutation or grown under a different condition, we can measure genetic interaction strength (GIS). The QFA R package includes functions for estimating GIS as well as  $p$ -values for the statistical significance of genetic interaction for each genotype in an interaction screen. The package

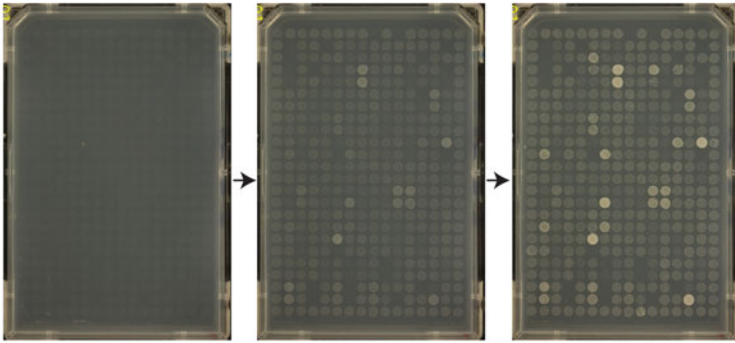
Library	ORF	Plate	Row	Column	Notes
SDLV4_1536	YBR184W	1	3	27	NA
SDLV4_1536	YBR184W	1	3	28	NA
SDLV4_1536	YAR029W	1	3	29	NA
SDLV4_1536	YAR029W	1	3	30	NA
SDLV4_1536	YBR217W	1	3	31	NA
SDLV4_1536	YBR217W	1	3	32	NA
SDLV4_1536	YBR174C	1	3	33	NA
SDLV4_1536	YBR174C	1	3	34	NA
SDLV4_1536	YBL043W	1	3	35	NA
SDLV4_1536	YBL043W	1	3	36	NA
SDLV4_1536	YAL059W	1	3	37	NA
SDLV4_1536	YAL059W	1	3	38	NA
SDLV4_1536	YBL011W	1	3	39	NA
SDLV4_1536	YBL011W	1	3	40	NA
SDLV4_1536	YBL001C	1	3	41	NA
SDLV4_1536	YBL001C	1	3	42	NA
SDLV4_1536	YBR036C	1	3	43	NA
SDLV4_1536	YBR036C	1	3	44	NA
SDLV4_1536	YBL102W	1	3	45	NA
SDLV4_1536	YBL102W	1	3	46	NA
SDLV4_1536	YOR202W	1	3	47	NA
SDLV4_1536	YOR202W	1	3	48	NA
SDLV4_1536	YOR202W	1	4	1	NA
SDLV4_1536	YOR202W	1	4	2	NA
SDLV4_1536	YAL018C	1	4	3	NA
SDLV4_1536	YAL018C	1	4	4	NA
SDLV4_1536	YBR210W	1	4	5	NA

**Fig. 7** Representative Library Description file. The physical location (Plate, Row and Column) of each strain genotype (library ORF) in a QFA experiment is recorded in the library description file. This file contains one row per library spot, typically ~6000 rows per library. Once constructed, this library file can be reused for every QFA experiment using the specified library. Optionally, it can be used to record the contents of multiple libraries, by specifying different sets of data annotated with different Library names. See Subheading 3.3.5 for details on file construction

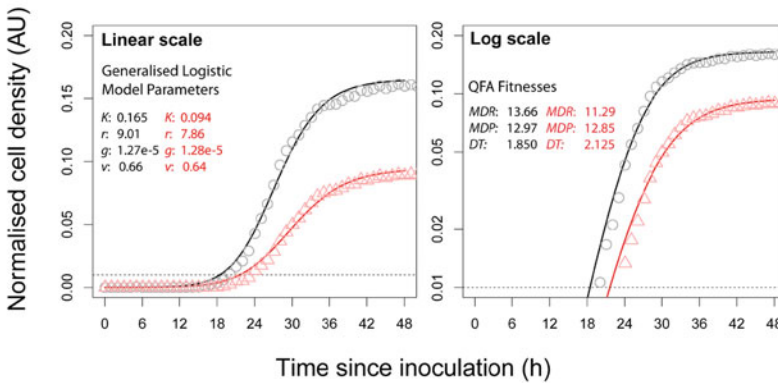
also includes functions for interactive visualization of evidence for genetic interaction in pairs of screens. Installation instructions and documentation can be found at: <http://qfa.r-forge.r-project.org/>

Steps to measure fitness and genetic interaction strengths and visualize results using the QFA R package are outlined below. Following programming convention, the names of functions (named sections of programs that carry out specific tasks) are written in fixed-width font (e.g., `colonyzer.read`) (*see* **Note 26**).

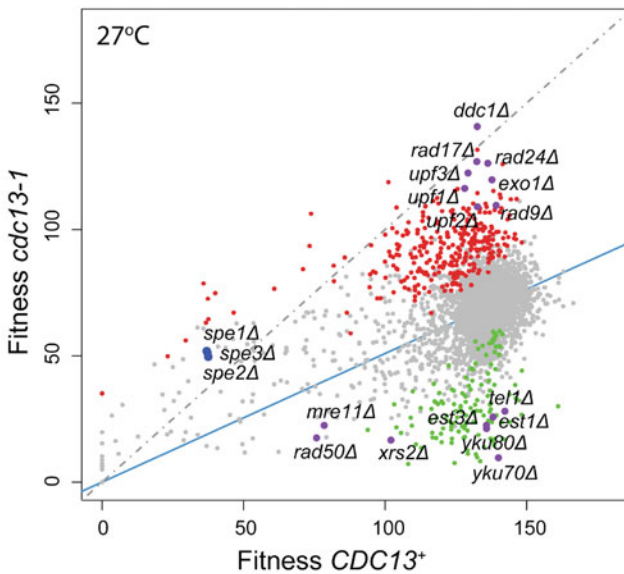
Timecourse photography



↓ Image analysis, building growth curves and fitting model to data



↓ Comparing two QFA screens to calculate strength and significance of genetic interactions



2D scatter plot comparing the quantitative fitnesses of deletion mutants

3.3.7 *Measuring Fitness*

1. Navigate to the Image Data Directory (*see* Subheading 3.3.3, **step 2**) containing the images captured during the experiment and run Colonyzer. This will generate two new sub-directories, one containing all analyzed images (Output\_Images) and one containing all data files (Output\_Data).
2. Go to the Output\_Images folder and confirm alignment of the spots within the rectangular Colonyzer segments. If automatic alignment is poor, consider using a Colonyzer calibration file to assist with spot location. Colonyzer calibration files can be built using the Parameteryzer software which is installed alongside Colonyzer.
3. Open the QFA package in R. Load the Experimental Description file (Subheading 3.3.4), Library Description file (Subheading 3.3.5) and Colonyzer output files (Subheading 3.3.6.1.2) using colonyzer.read. This function associates meta-data with each spot, grouping cell density data into growth curves.
4. Export the R data frames as QFA Raw data text files to an experimental results directory (write.table).
5. Use the QFA package to fit the logistic population model to observations (qfa.fit) and to generate .pdf report files including plots of growth curves and model fits together with fitness estimates for each spot (qfa.plot) (*see* **Note 27**).
6. Export the R data frames as QFA Fitness data text files to an experimental results directory (write.table).

3.3.8 *Measuring Genetic Interaction Strength*

1. Select an appropriately matched pair of query and control screens. For example, if using a temperature sensitive query mutation *xyz-1* and an array of strains from the single deletion library, *abcΔ*, then compare fitnesses from the double mutant screen *xyz-1 abcΔ* with the single mutant control screen *XYZ abcΔ*, where strains have been grown in the same environments (e.g., on the same medium and at the same temperature).
2. Load the QFA fitness data text files for both the query and control screens from the experimental results directory (read.delim).
3. Carry out genetic interaction (or epistasis) analysis (qfa.epi), as described in [2].



**Fig. 8** Schematic overview of imaging, growth curve analysis, GIS calculation, and visualization. Spotted colonies are photographed every 4 h for up to 5 days. Growth curves (cell density estimates over time) are generated for each spot by image analysis and a generalized logistic model (*black* or *red line*) is fitted to the cell density estimates (*circle* or *triangle*). Evidence for genetic interaction can be visualized as points on a 2D scatter plot comparing fitnesses of two matched QFA screens that deviate from a regression through the origin (model of genetic independence)

4. Generate plots for visualizing genetic interaction (`qfa.eplot`) and tab-delimited GIS text files summarizing genetic interaction strengths and significance of interaction for each array mutation (`report.epi`). These tabulated data can be sorted by GIS and filtered by  $q$ -value ( $p$ -value after correction for multiple testing). For example, users often classify “hits” as array mutations which have a  $q$ -value  $< 0.05$ . Hits can then be ranked by GIS for follow-up. Note that both strongly negative and strongly positive GIS are informative: check both the top and bottom of ranked lists for interesting hits.
5. Optionally, load GIS text files into iRVIS, an interactive visualization tool included in QFA R package to allow rapid browsing of screen data. Specific instructions for installing and using the iRVIS tool can be found at <http://qfa.r-forge.r-project.org/visTool/>

---

## 4 Notes

1. Monosodium glutamic acid (MSG) is used as a nitrogen source in synthetic media containing G418 or clonNAT as ammonium sulfate hinders the function of G418 and clonNAT. Ammonium sulfate is used as a nitrogen source in any synthetic media that does not contain G418 or clonNAT.
2. Libraries and query strains are stored at  $-80\text{ }^{\circ}\text{C}$  in 15% glycerol. Query strains used for any SGA are always revived from  $-80\text{ }^{\circ}\text{C}$ . A master plate stock of the libraries can be kept in a 384-format on YEPD + kanMX at  $4\text{ }^{\circ}\text{C}$ . The master plate stock is propagated once a month onto a new YEPD + kanMX plates. The library is revived from the frozen stock every 3 months. For an SGA, the library must be freshly re-pinned from the master plate stock onto a new YEPD + kanMX plate.
3. Examine the plates after each SGA stage for any of the following:
  - Fungal or bacterial contamination occurring at any stage of the SGA is dealt with by repeating the pinning from the previous uncontaminated stage of the SGA or by excluding the plate from the analysis.
  - Gaps in the plate due to strain inviability, or due to being intentionally left blank, are expected to be consistent in the subsequent SGA stages. Growth in previously empty spaces can be indicative of a potential mix up of plates or contamination.
  - Pinning artifacts such as large empty patches on part of the plate are indicative of alignment issues of the pin tool and need to be resolved before continuing the procedure.

- The single gene deletion strain *can1Δ::kanMX* in the library is expected to lose viability during the meiotic progeny selection stages as the haploid meiotic spore progeny harbor either *can1Δ::kanMX* of the library strain or *can1Δ::STE2pr-Sp\_his5* of the query strain. Viability is indicative of an incorrect genotype or of contamination.
  - Evidence for genetic linkage at the end of the SGA procedure. Query strains containing a mutated gene residing close to a gene that is deleted in the library strain should not produce double mutants when crossed due to genetic linkage. Therefore such double mutant colonies should not be visible on the final SGA stage (HS V).
4. One culture is used as a backup in case of contamination. Typically, one 5 mL overnight culture is used to create five lawns of which three lawns are taken forward to the next step (pinning to 1536 format). Each lawn is used to create six plates in 1536 format. Take into account how many library plates will be used in the experiment to determine how many cultures are required for the procedure.
  5. Timing of the diploid selection and sporulation is crucial. Ideally incubate diploid selection plates for 36 h at 30 °C prior sporulation. Diploid selection plates cannot be stored at 4 °C prior to sporulation. If necessary (e.g., weekend) store the plates after mating at 4 °C and then re-pin on diploid selection plate. At all other stages the plates can, if required, be kept at 4 °C for a couple of days before the next step.
  6. Temperature is crucial for sporulation and should be kept between 22 and 24 °C. Sporulated plates can be kept at 4 °C for up to 4 months without significantly affecting spore viability. >1% of spore frequency is required to continue with meiotic progeny selection.
  7. HSI and HSII allow germination of MATa haploids containing *can1Δ::STE2pr-Sp-his5* and *lyp1Δ* markers. Diploids containing a wild-type copy of *CAN1* and *LYP1* are not viable.
  8. HSIII allows growth of MATa haploids containing *can1Δ::STE2pr-Sp-his5*, *lyp1Δ* and kanMX markers.
  9. HSIV allows growth of MATa haploids containing *can1Δ::STE2pr-Sp-his5*, *lyp1Δ*, kanMX and hphmx markers.
  10. HSV allows growth of MATa haploids containing *can1Δ::STE2pr-Sp-his5*, *lyp1Δ*, kanMX, hphmx and clonNAT markers.
  11. If QFA cannot be performed following SGA, the final SGA plates can be stored up to a month at 4 °C. Stored plates need to be repinned onto a fresh final SGA plate (SDM\_rh1k\_CTGNH, media no. 9) and incubated for 2 days at 20 °C prior the QFA procedure.



12. Using this protocol, 1 L of liquid media can fill around 48 96-well culture plates. 16 96-well culture plates are required to culture every colony from a 1536 format SGA plate. For automated plate filling, a Biomek WellMate Dispenser can be used to consistently dispense 200  $\mu$ L into each well.
13. The number of solid agar plates required for the QFA will depend on the number of technical replicates, conditions (temperatures and drug concentrations) and the number of final SGA plates in 1536 format. Four solid agar plates are required for each SGA plate using one condition and no technical replicates.
14. Before starting the procedure, bear in mind how many plates can be run at once on the BM3 robot. Each 1536 format plate requires four 96-well plates to transfer one biological repeat of each strain on the plate. Each BM3 stacker can hold a maximum of 24 96-well culture plates. Therefore, up to six final SGA plates in 1536 format can be used on each of the four sides of the robot. For details of setting up the plates, *see* Fig. 5. When setting up the robot, always inoculate a complete repeat at once, do not split within repeats.
15. Liquid cultures can be kept at 4  $^{\circ}$ C and reused for 1 month.
16. In addition to concentrated spotting, it is possible to spot diluted cultures. To do this, saturated cultures in the 96-well culture plates are diluted in 200  $\mu$ L sterile water in a second 96-well plate using a 2 mm pin tool (approximately 1:70 dilution). Diluted cultures are then spotted onto solid agar plates.
17. A useful workflow is to spot one repeat at a time.
18. Prior to spotting, take into consideration the maximum capacity of the imaging incubators.
19. When spotting overnight, the start can be delayed to ensure that spotting finishes in the morning to ensure that the solid agar plates are freshly spotted and can be transferred to the incubator at time point 0.
20. When using a humidified incubator, wiping the bottom of the plates with a one in ten dilution of washing up liquid prevents condensation forming on the bottom of the plates which shows up during imaging. Plates incubated in a nonhumidified incubator are generally not affected by condensation problems.
21. Before entering the cycle interval time, ensure that the robots are capable of imaging all the plates in the incubator within the set time frame. When imaging a huge amount of plates, a longer cycle interval might be required. It takes on average 2 min for one plate to be imaged and returned to the incubator.
22. If the experiment is split into several batches, ensure that all repeats are imaged in the same incubator as the intensity of the

light bulbs and the color of the diffuser changes over time, which might affect analysis. The same consideration applies to control QFAs that will be compared to another QFA.

23. Image plates for up to 5 days. Incubating the plates for too long increases the risk of fungal or bacterial contamination. Also, agar will eventually dry and crack, making spot location and tracking more difficult.

24. Column headers for the experimental description file:

**Barcode**—A unique identifier for each plate. Barcodes can be generated automatically by imaging robot software, for example, (*see* Subheading 3.3.3, step 3).

**Start.Time**—The starting time of the experiment, should be an accurate estimate of the inoculation time for that plate. The time format should be yyyy-mm-dd\_hh-mm-ss.

**Treatment**—Information about external conditions applied to plates, such as incubation temperature or radiation.

**Medium**—Contents of the agar used in the screen (e.g., nutrients or drugs added).

**Screen**—An easily interpretable screen name.

**Library**—Library name (should correspond to a Library name in the Library Description file, *see* Subheading 3.3.5).

**Plate**—Libraries can have thousands of yeast strains across multiple different plates. The Plate Number column is used to link the barcode of the plate to the correct library plate; this allows us to map strains to specific positions on specific plates. The number should correspond to a plate number in the Library Description file (*see* Subheading 3.3.5).

**RepQuad**—Genome-wide QFA experiments in 384 format are derived from one of four replicates on final SGA plates (1536 format). This number is a record of which of those four replicates the QFA plate was derived from.

**Client**—The individual who requested the screen.

**ExptDate**—A more legible date for use in reports and plots.

**User**—The operator who performs the screen.

**PI**—The group leader of the client.

**Inoc**—Shorthand label for the inoculation density of the experiments. This could be CONC (no dilution prior to spotting) or DIL (dilution prior to spotting), for example.

**Condition**—The most important environmental attribute. Depending on the design of the experiment, this could be the temperature applied to the plate or the composition of the media.

25. Column headers for the library description file:
- Library name—This must match the library name specified in the Experimental Description file in Subheading 3.3.4.
  - ORF—Systematic open reading frame name specifying genotype of spot on library plate.
  - Plate—Library plate number.
  - Row—Spot row number.
  - Column—Spot column number.
  - Notes—Any notes that might be relevant.
26. For more detailed information on how to use QFA package functions, including what information they require to work and what their output will be, please refer to the QFA R package documentation. For instance, once the QFA package has been loaded, to access information about a specific function, type “?” into the R terminal followed by the function name (e.g., ?colonyzer.read) and press Enter. Package documentation also includes a complete example of fitness and GIS estimation from metadata files and Colonyzer output.
27. The QFA package allows edge colonies to be discarded as the decrease in neighboring colonies leads to more available nutrients and image analysis difficulties near plate walls. The QFA package also allows for stripping of specific genotypes if required (e.g., failed SGA, linkage).

## References

1. Zubko MK, Guillard S, Lydall D (2004) Exo1 and Rad24 differentially regulate generation of ssDNA at telomeres of *Saccharomyces cerevisiae* cdc13-1 mutants. *Genetics* 168(1):103–115
2. Addinall SG et al (2011) Quantitative fitness analysis shows that NMD proteins and many other protein complexes suppress or enhance distinct telomere cap defects. *PLoS Genet* 7(4): e1001362
3. Banks AP, Lawless C, Lydall DA (2012) A quantitative fitness analysis workflow. *J Vis Exp* (66): e4018
4. Andrew EJ et al (2013) Pentose phosphate pathway function affects tolerance to the G-quadruplex binder TMPyP4. *PLoS One* 8(6): e66242
5. Downey M et al (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell* 124(6):1155–1168
6. Tong AH, Boone C (2006) Synthetic genetic array analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol* 313:171–192
7. Breslow DK et al (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* 5(8):711–718
8. Lawless C et al (2010) Colonyzer: automated quantification of micro-organism growth characteristics on solid agar. *BMC Bioinformatics* 11:287

## Rewiring the Budding Yeast Proteome using Synthetic Physical Interactions

Guðjón Ólafsson and Peter H. Thorpe

### Abstract

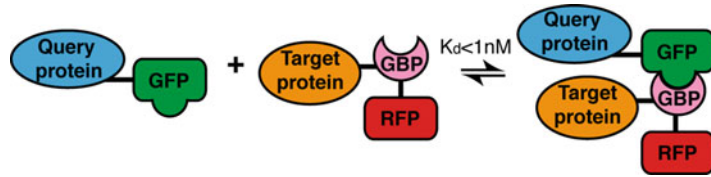
Artificially tethering two proteins or protein fragments together is a powerful method to query molecular mechanisms. However, this approach typically relies upon a prior understanding of which two proteins, when fused, are most likely to provide a specific function and is therefore not readily amenable to large-scale screening. Here, we describe the *Synthetic Physical Interaction* (SPI) method to create proteome-wide forced protein associations in the budding yeast *Saccharomyces cerevisiae*. This method allows thousands of protein–protein associations to be screened for those that affect either normal growth or sensitivity to drugs or specific conditions. The method is amenable to proteins, protein domains, or any genetically encoded peptide sequence.

**Key words** Green fluorescent protein (GFP), GFP-binding protein (GBP), Nanobody, Chromobody, Selective ploidy ablation, Protein-protein interactions (PPI), High-throughput screen

---

### 1 Introduction

Fusing proteins together by linking their open reading frames is a powerful method to query the role of specific protein–protein interactions. For example, fusing Cdc13 with Cdc2 in fission yeast is sufficient to drive cell cycle progression [1] and fusing spindle assembly checkpoint components arrests cell cycle progression in budding yeast [2]. These studies underlie the notion that a protein’s function often depends upon its interacting partners and more broadly that many cellular functions are controlled by the spatial organization of proteins [3], a concept familiar to the cell signaling field [4]. However, such studies rely upon making a separate genetic construct for each fusion, which limits the ability to test large numbers of possible fusions. To address this issue we have made use of two resources. First, a GFP binding protein (GBP, also called nanobody, nanotrap, or chromobody—the latter when linked to a fluorophore) derived from an alpaca antibody heavy chain [5–11]. Genetically encoded GBPs are typically 11–13 kDa



**Fig. 1** The cartoon illustrates binding between a query protein (blue) tagged with GFP (green) and a target protein (orange), tagged with GBP (pink) and RFP (red)

and bind to GFP with high-affinity in vitro ( $K_d < 1$  nM). Importantly for this approach, proteins tagged with GBP bind to GFP-tagged proteins in vivo (for examples *see* [12] or [13] (Fig. 1)). Thus fusing a protein of interest—a target protein—with the GBP allows it to associate with any GFP-tagged query protein. Second, in the budding yeast, *Saccharomyces cerevisiae*, a genome-wide library of strains was constructed each encoding a separate endogenously tagged GFP protein [14]. Our *Synthetic Physical Interaction* (SPI) method introduces a target protein tagged with GBP into the GFP library, each of which contains a separate GFP-tagged query protein, to create thousands of protein–protein associations. Each interaction is then assayed for growth or sensitivity to a particular condition.

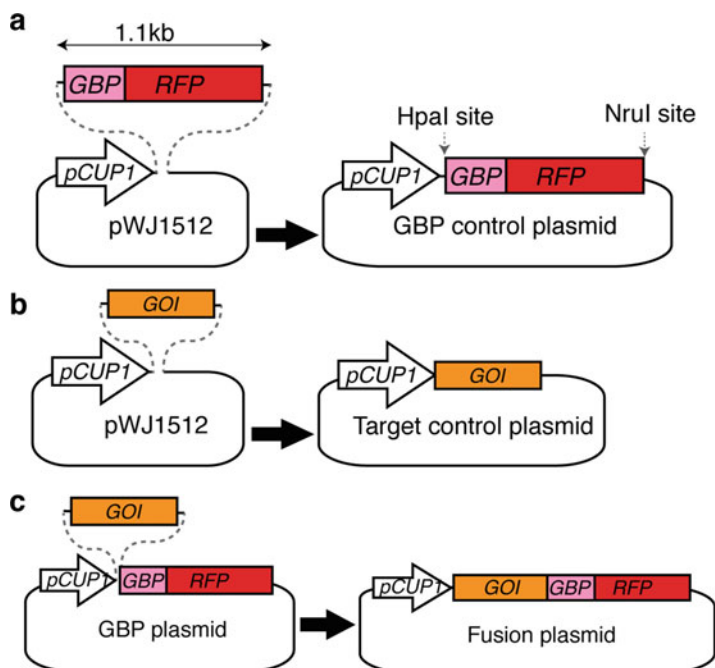
## 2 Materials

### 2.1 GFP-Binding Protein (GBP)

There are various peptide sequences derived from camelid antibodies that bind to GFP. The first GBP clone was isolated by Uhlrich Rothbauer, Heinrich Leonhardt and colleagues [5, 6], but more recently a large repertoire of GBPs have been isolated [9]. This new array of GBPs theoretically enable the SPI method to employ two GBPs that bind to different epitopes of GFP and so recruit either two copies of the same protein or two different proteins to GFP. The sequence encoding GBP is sufficiently short (<400 bp) that it can be introduced at either the 5' or 3' end of open reading frames to create amino- or carboxy-tags of GBP target proteins with ease and usually does not perturb their function. We use a plasmid to drive expression of the GBP-tagged target protein.

### 2.2 Plasmids for Synthetic Physical Interactions

In principle any suitable plasmid could be used to introduce the GBP-tagged target protein into the GFP strains. We have used a single copy *CEN* plasmid based upon pWJ1512 [15] (Fig. 2). This plasmid contains a *LEU2* marker gene for selection and a *CUPI* promoter to drive expression of the GBP construct. The *CUPI* promoter is active without copper, but its expression can be elevated with added copper [16]. The strains of the GFP collection are auxotrophic for leucine, uracil, and methionine. However, the selective ploidy ablation method to transfer plasmids uses selection



**Fig. 2** An example of the methodology used to create plasmids for the SPI screen is shown. (a) GBP–RFP is inserted into a suitable plasmid to create a GBP–RFP control plasmid. (b) A gene of interest (GOI) is separately inserted into the same plasmid to create a target protein control. (c) The gene of interest (GOI) is then inserted into the GBP–RFP plasmid (a) to create a construct encoding a fusion between a target protein and GBP–RFP

against *URA*<sup>+</sup>; consequently plasmids encoding *LEU2* or *MET15* could be used in addition to any drug selection markers (such as kanamycin, hygromycin, or nourseothricin). In addition to a fusion plasmid (Fig. 2c), we typically use a GBP plasmid (Fig. 2a) and “Gene of Interest” plasmid (Fig. 2b) as controls in the SPI screens.

### 2.3 Yeast Strains/ GFP Collection

The GFP collection of strains is currently distributed by ThermoFisher Scientific (MA, USA) and contains 4159 strains. All strains are derivatives of BY4741 (*MAT $\alpha$  his3 $\Delta$ 1 leu2 $\Delta$ 0 met15 $\Delta$ 0 ura3 $\Delta$ 0*) [14, 17]. The library is delivered in frozen glycerol stocks in 96-well plates and after thawing can be readily transferred onto rectangular agar plates using pinning tools.

### 2.4 Universal Donor Strain for Selective Ploidy Ablation

The *MAT $\alpha$*  Universal Donor Strain (W8164-2B) is available from the Rothstein lab [15], *MAT $\alpha$  CEN1-16<sup>GCS</sup> can1-100 his3-11,15 leu2-3,112 LYS2 met17 trp1-1 ura3-1 RAD5* (where *CEN1-16<sup>GCS</sup>* indicates that all the centromeres contain a counterselectable *Kluyveromyces lactis* *URA3::GAL* promoter cassette [18]). This strain can be grown with glucose as a carbon source and is readily transformed with plasmids for the SPI screen.

## 2.5 Media and Stock Solutions

1. YPD liquid media: 20 g/L Bacto peptone, 10 g/L Yeast extract and 2% (weight by volume) glucose. Dissolve these ingredients in water and sterilize by autoclaving.
2. YPD solid agar media: 20 g/L Bacto peptone, 10 g/L Yeast extract, 2% (weight by volume) glucose and 20 g/L Bacto agar. Dissolve ingredients in water (agar will not dissolve) then sterilize and melt agar by autoclaving. Allow to cool to ~55 °C prior to pouring plates. Each rectangular plate should take ~50 ml of media (*see Note 1*).
3. –Leucine amino acid mix: Amino acid supplement mixes, including drop-outs, such as the “minus leucine” media described here are available commercially. We use a mix containing the following: 20 mg/L adenine sulfate, arginine sulfate, histidine–HCl, methionine, tryptophan, and uracil; 30 mg/L isoleucine, lysine–HCl, and tyrosine; 50 mg/L phenylalanine and 150 mg/L valine.
4. GAL –Leu solid agar media: Prepare a solution containing 1.7 g/L yeast nitrogen base without amino acids, 5 g/L ammonium sulfate (*see Note 2*), 41 mg/L –Leucine amino acid mix (described in **item 3**), 20 g/L Bacto agar. Dissolve ingredients in water (agar will not dissolve) then sterilize and melt agar by autoclaving. After the media has cooled to 55–60 °C, add 100 ml of 20% (weight by volume) sterile galactose solution to each L of media, mix and pour as for YPD plates (*see Note 6* for diploid assay).
5. GAL –Leu 5-FOA solid agar media: Mix 20 g Bacto agar with 500 ml of water in a 1-L bottle, melt using an autoclave. Prepare a 500 ml solution in water containing 1.7 g yeast nitrogen base without amino acids, 5 g ammonium sulfate (*see Note 2*), 41 mg –Leucine amino acid mix (from **item 3**), 30 mg uracil and 750 mg 5-fluoro-orotic acid (5-FOA) and 2% (weight by volume) galactose. Allow this solution to dissolve with stirring for approximately 1 h at room temperature, then filter-sterilize through a 0.22 µm filter and place in a 60 °C incubator or water bath. When the autoclaved agar has cooled to 55–60 °C, add in the warmed solution, mix well by gentle stirring and pour the plates immediately.

## 2.6 Plates for SPA

All plates described in this method are the Singer Plus plates (Singer Instruments Ltd, Somerset, UK), which are designed for use with the ROTOR robotic pinning platform. However, when using other pinning tools, any 12 cm by 8 cm sterile rectangular plates are suitable; for example the Nunc™ OmniTray™ (Thermo Scientific, MA, USA).

## 2.7 Pinning Robot

The ROTOR robotic pinning platform (Singer Instruments Ltd, Somerset, UK) allows 96, 384, and 1536 arrays of yeast colonies on

rectangular agar plates to be copied from plate to plate using sterile pins. It is also simple to convert between 96, 384, and 1536 arrays. If hand pinning tools are used, it is very important to ensure that an even pinning pressure is applied in each copying step.

## 2.8 Plate Imaging

In principle, most desktop scanners that function in transmission mode are suitable for collecting images of plates. Images should be at least 300 dpi. It is possible to use specialist hardware such as the Phenobooth™ (Singer Instruments Ltd, Somerset, UK) to capture images of the plates.

---

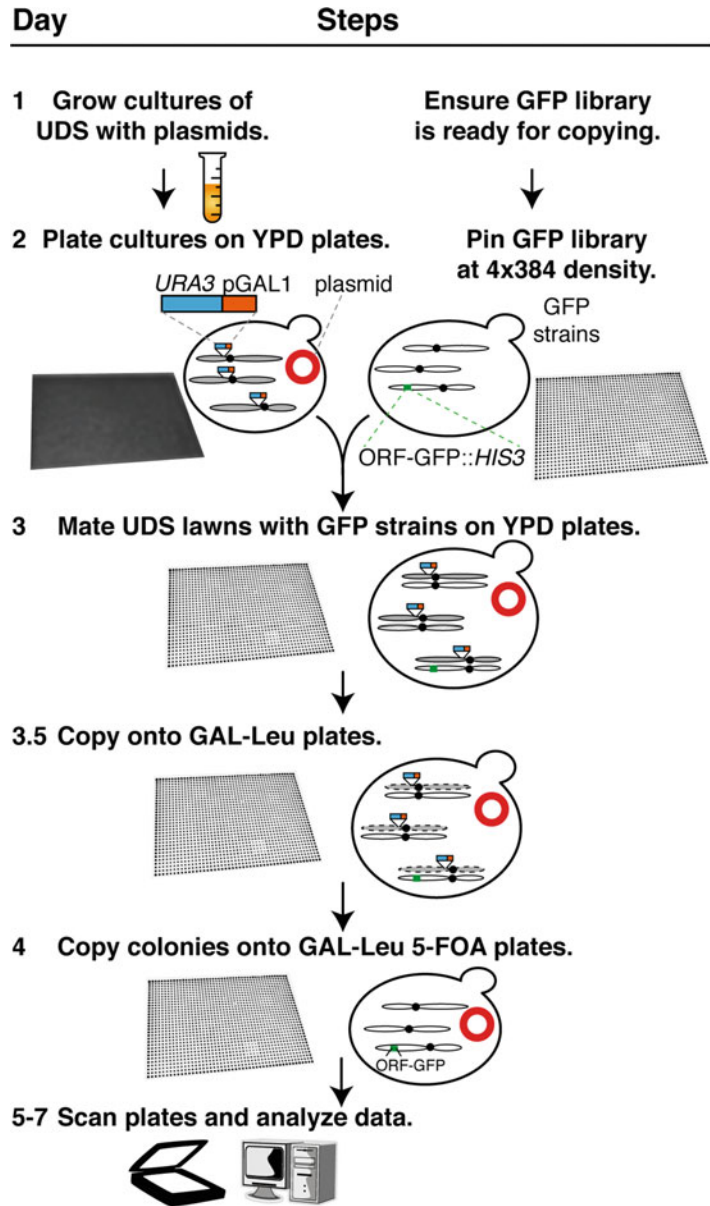
## 3 Methods

### 3.1 Selective Ploidy Ablation

We use the following general method for SPI screens (Fig. 3)

1. The arrayed GFP strains should be copied onto rectangular YPD agar plates. We typically have a copy of the GFP library stored at 4 °C on YPD plates, with 384 strains on each plate. Prior to the screen we copy this library onto fresh YPD plates at 1536 colonies/plate density, i.e., four copies of each of the 384 strains on each plate.
2. Grow up 5–10 ml cultures of the UDS strain (W8164-2B) each containing a separate plasmid for the SPI assay (e.g., GBP alone, kinase alone, kinase-GBP, and kinase mutant-GBP) in selective media at 30 °C, shaking, overnight. Collect the cells by centrifugation at  $5000 \times g$  for 5 min. Each lawn provides sufficient UDS cells to mate with four arrays of GFP strains, if there are more than four plates of GFP strains then additional lawns will need to be created (*see Note 3*).
3. Resuspend the cell pellets in ~400 µl of growth media and plate all of this onto YPD plates. Using 4 mm sterile glass beads to spread the cells onto the rectangular agar plates gives a more even lawn of cells than a glass spreader tool.
4. Grow these lawns overnight in 30 °C incubator.
5. Use a pinning tool or robotic pinner (e.g., Singer ROTOR pinning robot) to copy the GFP strains onto fresh YPD plates (*see Note 4*).
6. Overlay these YPD plates with the cells from the lawns by pinning in the same way as **step 5**.
7. Grow these “mating” plates for 6–7 h at 30 °C (*see Note 5*).
8. Copy the colonies by pinning onto GAL-Leu rectangular agar plates (*see Note 6*).
9. Grow for 24 h in a 30 °C incubator.
10. Copy the colonies by pinning onto GAL-Leu 5-FOA rectangular agar plates.





**Fig. 3** The key steps involved in performing a SPI screen are illustrated with a daily timeline on the left side. Cultures of the UDS strain containing the SPI plasmids are grown and plated to create lawns. In parallel, the GFP library is prepared on agar plates. The GFP strains are mated separately with each of the UDS plasmid strains by copying them together on an YPD plate. These arrays are then copied sequentially onto galactose –leucine (GAL –LEU) and then galactose –leucine + 5FOA (GAL –LEU + 5FOA) medium. All incubations are done at 30 °C. The resulting arrays of strains are then imaged to compare controls with experiments

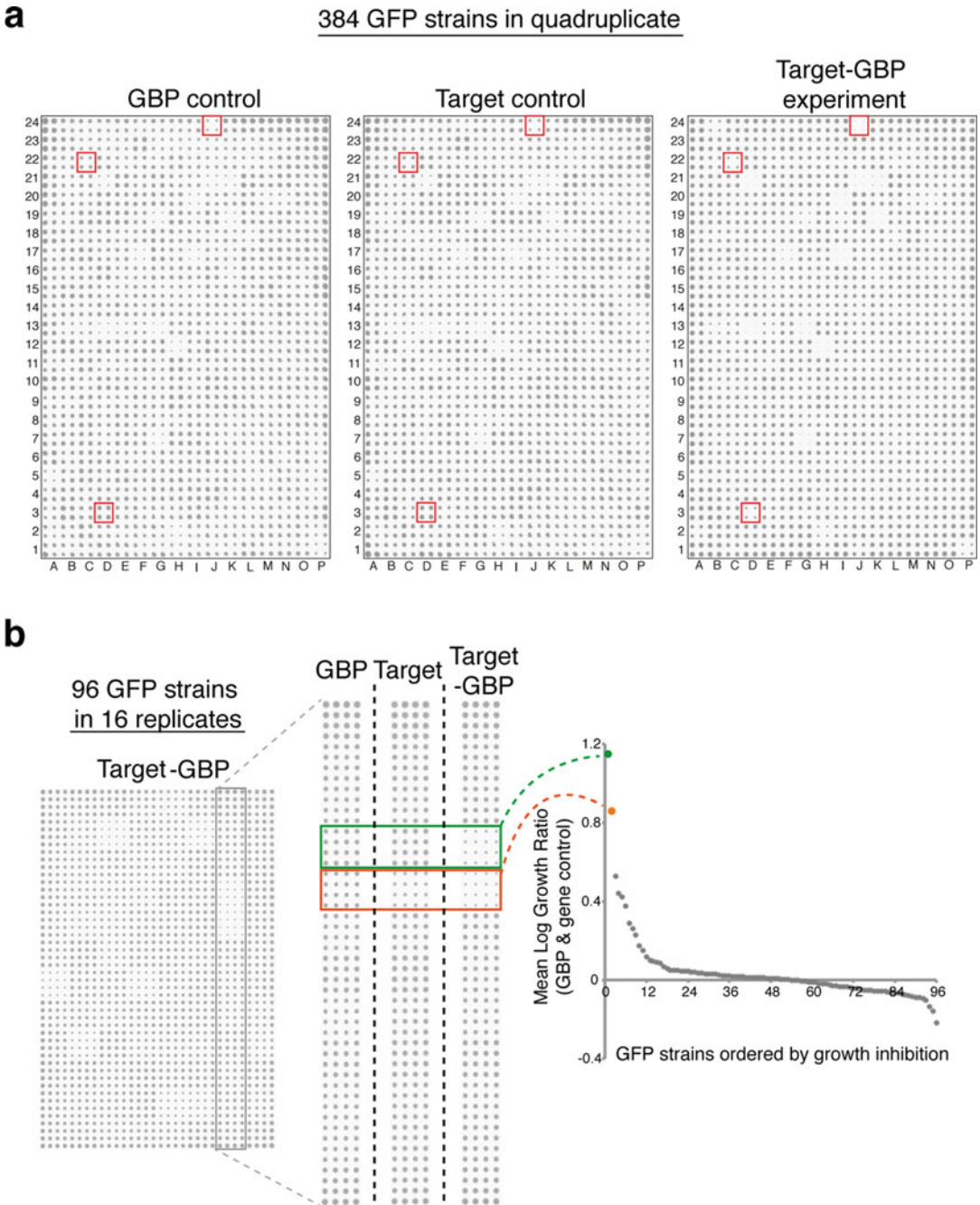
11. Grow for 24–72 h (depending on colony sizes) in a 30 °C incubator.
12. Capture images of the plates.

### 3.2 Data Analysis

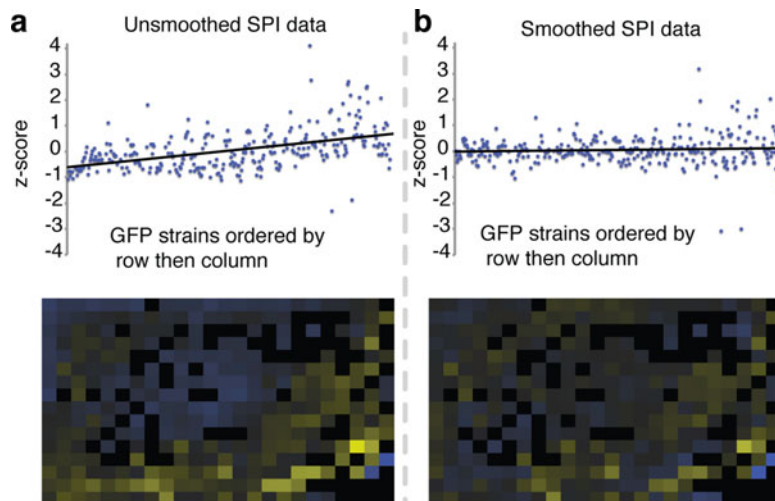
It is possible to assess the strongest interactions by eye (Fig. 4a), however, we recommend a more quantitative analysis using imaging. The plates from the SPI screen are imaged with a desktop scanner in transmission mode (colonies should appear dark on a light background) or plate imager (*see Note 7*). The resulting images are analyzed computationally to determine colony sizes:

1. We employ the “CM engine” (Colony Measurement engine), a freely available ImageJ plugin and part of the ScreenMill suite of software [19]. CM engine analyses images and reports the size of each colony on the plate in pixels, colony circularity and location coordinates. We use the default settings (*see Note 8*). The output file from the CM engine is a “Log file,” which lists the measurements for each colony.
2. The “Log file” is uploaded to the next part of ScreenMill: The Data Review engine (DR engine), which is available online: [http://www.rothsteinlab.com/tools/screen\\_mill/dr\\_engine\\_setup](http://www.rothsteinlab.com/tools/screen_mill/dr_engine_setup). Here the comparisons between the colonies on the control plates and the experiment plates are performed. For an easy readout of the data we recommend generating a “Key file.” The Key file is a list ascribing each position on each plate with a particular query GFP-tagged gene. We run the DR engine with the default settings (*see Note 9*).
3. Plate based screens often suffer from spatial anomalies, such as colonies growing larger at the top of the plate than the bottom [20, 21]. We use a simple smoothing algorithm to spatially adjust the relative growth rates on plates [13] (Fig. 5). The PERL script is freely available online: [http://www.sourceforge.net/projects/zspatialcorrect/files/spatial1\\_0.plx/download](http://www.sourceforge.net/projects/zspatialcorrect/files/spatial1_0.plx/download).

The resulting data output includes the Log Growth Ratio (LGR) or *z*-score for each query protein to assess specific interactions that affect growth (Fig. 4b). The LGR is the natural log of the average colony size on the controls divided by the average colony size of the experiment. Hence, higher LGRs indicate a stronger growth defect. To understand how a specific LGR equates with a growth difference, we calculate the theoretical difference in the number of yeast cells relative to colony size assuming that colonies are hemispheres (Fig. 6a). This works well in practice and shows that a typical LGR of 0.4 equates with ~80% more cells on the control than the experiment (Fig. 6b). The formula we derive for converting LGR to the ratio of the number of yeast on control and



**Fig. 4** An example of images of plates resulting from a SPI screen is shown. **(a)** Each plate has the same array of 384 GFP strains, each copied in quadruplicate (hence 1536 colonies in total). On each separate plate the GFP strains contain a different plasmid—the GBP control, the untagged target protein control, and the target–GBP fusion, from *left* to *right*. The three highlighted strains (*red boxes*) show three strains that are affected in the experiment but not controls. **(b)** An example of a SPI screen plate, now with 96 GFP strains each with 16 replicates. Two strains that result in a SPI are highlighted with *rectangles (orange and green)*. The Mean Log Growth Ratios (LGR) of the two controls versus experiment are then plotted to show the strongest SPIs (the *orange and green dots* represent the same SPIs indicated with the *rectangles* and are the strongest SPIs on this plate)



**Fig. 5** A spatial defect caused by cells growing differently across a plate is highlighted by plotting the LGR or z-scores by their row and column (**a**); a linear trend line is added to clarify the spatial defect. The bias is also visualized by a color-coded map of the plate, showing high z-scores in *yellow (light)* and low z-scores in *blue (dark)* (shown below). A simple smoothing algorithm can correct this defect by normalizing scores in each row and column (**b**)

experiment (YR) is  $YR = e^{1.5LGR}$ . We typically will retest all associations that have high LGRs with at least 16 replicates to confirm their reproducibility (for example *see* Fig. 4b).

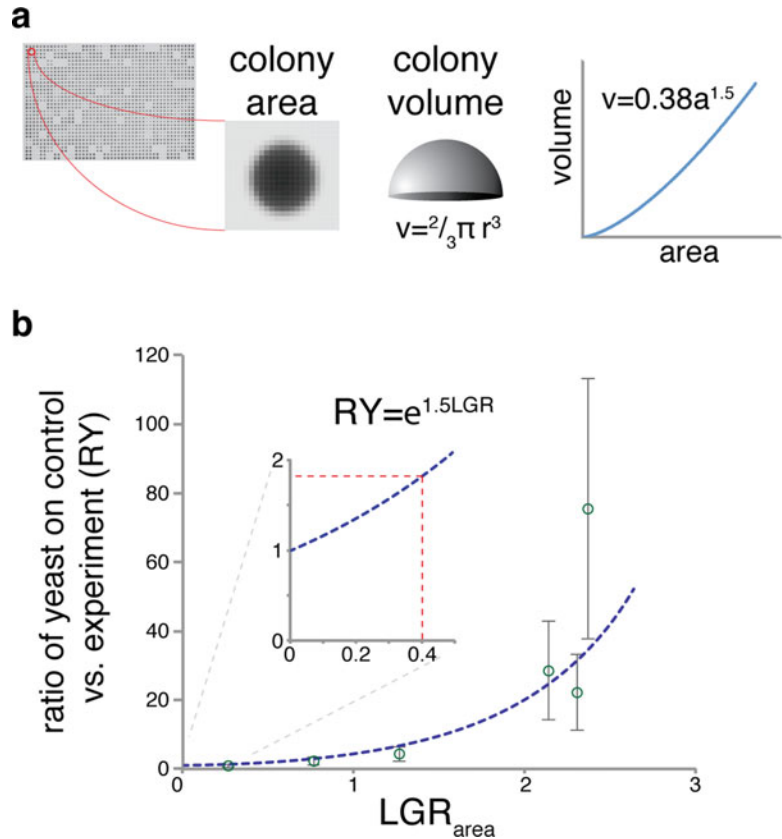
### 3.3 Microscopy

It is possible to validate that the GFP- and GBP-tagged proteins are interacting using fluorescence imaging, since the GBP tag also includes an RFP sequence (Fig. 7).

1. From the final GAL –Leu 5-FOA plate, colonies are picked and placed into 5 ml of GAL –Leu + Ade (*see* **Note 10**; *see* **Note 6** for diploid cells) and grown for ~24 h at 23 °C in a shaking incubator.
2. Collect the cells by centrifugation at  $5000 \times g$  for 5 min, resuspend in  $5 \times$  the pellet volume (GAL –Leu + Ade) (*see* **Note 11**; *see* **Note 6** for diploid cells).
3. Place 1.25  $\mu$ l of cells and 1.25  $\mu$ l agar (–Leu + Ade with 1.4% low melting point agarose—maintained as a liquid at 42 °C) onto a glass microscope slide (*see* **Note 12**) cover with a cover-slip and use fluorescence imaging to detect fluorophores.

## 4 Notes

1. It is very important that the plates are poured on a level surface to ensure that the agar surface is flat and level. This will ensure even colony growth across the plate. If the surface of the agar is

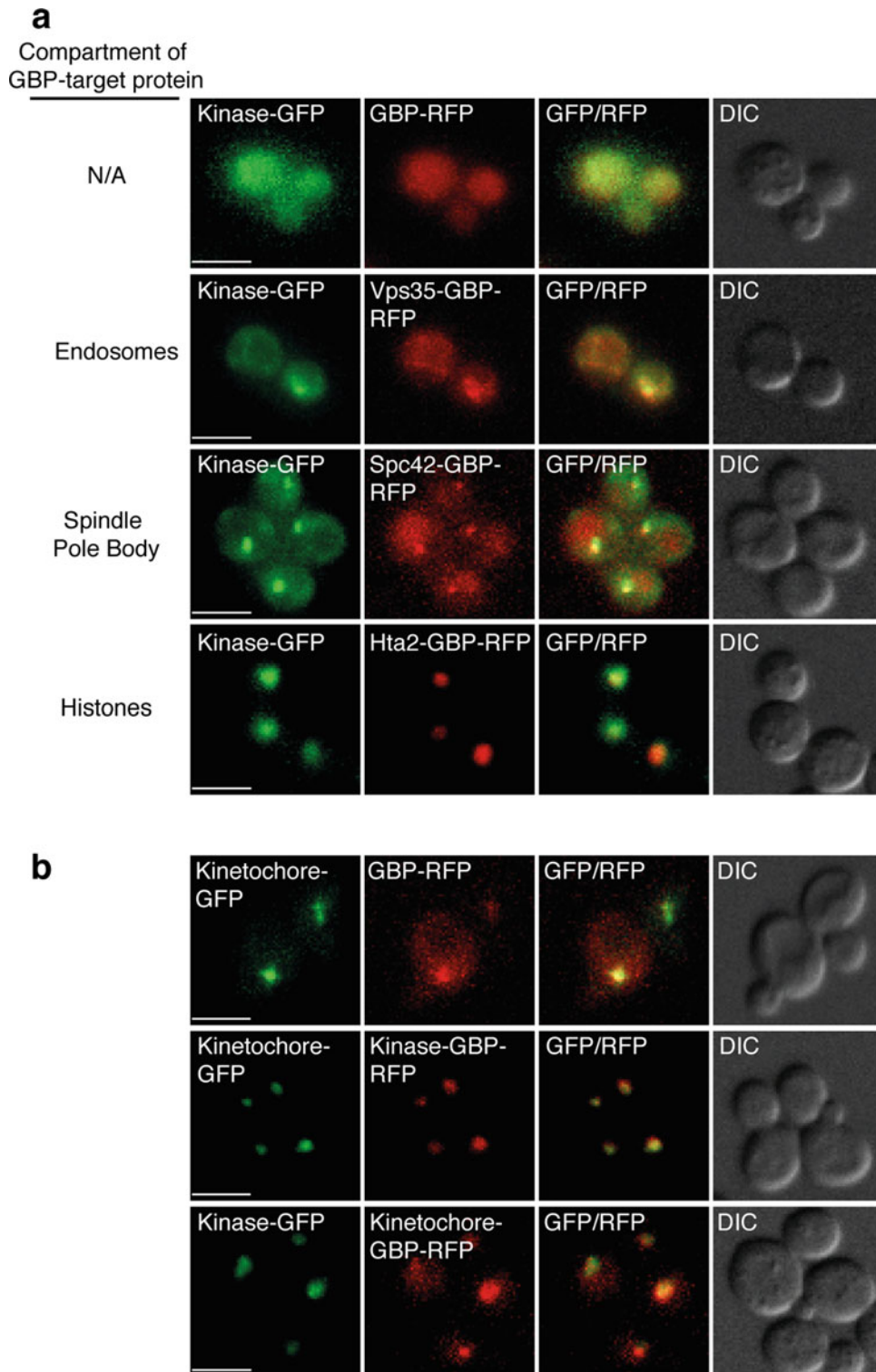


**Fig. 6** The assessment of growth relies upon measuring colony area as a surrogate for the number of yeast on the plate. **(a)** We theorize that colonies approximate to hemispheres and consequently the relationship between area ( $a$ ) and volume ( $v$ ) is  $v = 0.38a^{1.5}$ . **(b)** We collected and counted the number of cells in colonies to confirm the relationship between the ratio of yeast on control versus experiment (RY) and the LGR and find that broadly our data fit the model. We note that a cutoff of  $LGR = 0.4$ , the minimum LGR observable by eye, equates with growth ratio of 1.8, a 45% drop in the number of cells on the experiment versus the control (*inset*)

not level it can affect yeast colony pinning, i.e., if the agar level/height is low on one side compared to another it can cause the pins to not reach the low side. Additionally, if one side of the plate has higher/more agar then the colonies on that side will have more nutrients and thus could grow larger as a result.

2. Yeast nitrogen base is often sold as a mix with ammonium sulfate, in which case use 6.7 g/L of this mix for synthetic media. It is possible to buy commercially mixtures of amino acids and also “homemade” recipes vary slightly from lab to lab.





**Fig. 7** Examples fluorescence microscopy images of GBP-RFP being recruited to GFP-tagged proteins, the columns from *left to right* show GFP signal, RFP signal, GFP-RFP merge, and differential interference contrast (DIC) signal. The scale bar is 5  $\mu$ m in all images. **(a)** A GFP-tagged kinase can be recruited to different cellular compartments via interaction with GBP-target proteins. **(b)** A GFP-tagged kinase is recruited to a GBP-tagged kinetochore and a GBP-tagged kinase is recruited to a GFP-tagged kinetochore

3. Each plate of GFP strains will need to be copied onto several YPD plates (i.e., make multiple copies). The number of copies will depend upon the number of separate plasmids (or plasmid combinations) that you are using in the experiment. For example, in a simple experiment you may have three UDS lawns each containing a different plasmid (encoding GBP alone, protein of interest alone, and protein of interest fused with GBP). However, in more complex experiments you may have more than three conditions and consequently you will need more copies of the GFP strains.
4. The options for the Singer ROTOR pinning robot settings are quite extensive, for example the pinning pressure, diameter, and frequency can be customized. For 1536 colony density format we use the following settings for source and target plates: Pinning; Pin Pressure: 25%, Repeat Pin: two times. Dry Mix: Diameter: 0.1 mm, Cycles: two rotations. For 384 colony density format we use the following settings for source and target plates: Pinning; Pin Pressure: 10%, Repeat Pin: two times. Dry Mix: Diameter: 0.3 mm for source plates and 0.1 mm for target plates, Cycles: two rotations. For “lawn plates” we change The ROTOR robot pinning settings of the “Dry Mix: Diameter” to 1 mm for source plates. This is to ensure that cells are picked up in cases where there might be gaps or uneven cell density on lawn plates, thus pinning with an increased diameter further ensures cells are picked up.
5. The “mating” step can be done overnight without affecting results, but prolonged time on YPD can result in “overgrowth” of the colonies to the extent that separate colonies grow together.
6. Alternatively to, or in parallel with the haploid assay, the cells from the mating step can be kept as diploid cells by selecting on –Histidine –Leucine media with glucose as a carbon source. A diploid screen will determine dominant effects of the GBP–GFP association, since these diploids have an extra set of chromosomes encoding an untagged version of the GFP-tagged query protein. For a diploid screen, **steps 8–10** should be replicated using the following protocol: Copy the mated colonies by pinning onto glucose –His –Leu rectangular agar plates. Grow for ~24 h in a 30 °C incubator. Copy the colonies by pinning again onto fresh glucose –His –Leu rectangular agar plates. Grow for ~24 h in a 30 °C incubator before imaging/analyzing. Note that diploid cells in glucose media typically grow faster than haploid cells in galactose and 5-FOA, thus the time needed to grow cells for the diploid assay is reduced compared with the haploid assay. Media used for microscopy has to be adjusted when imaging diploid cells; Use glucose instead of galactose in the –Leu + Ade media when growing cultures overnight for imaging.

7. We scan the plates using a desktop flatbed scanner (Epson V750 Pro, Seiko Epson Corporation, Japan) at 300 dpi resolution in transmission mode. The plate lids are removed and the plates are inserted upside down in the scanner. We save the images in “tiff” format to use with the ScreenMill software.
8. We use the following default settings for the ScreenMill CM engine ImageJ plugin: Fine-crop Mode: Automatic; Colony measurement method: Standard; Plate Density: 384 or 1536; Running Mode: Standard.
9. We run the DR Engine in default mode and use the following Statistical Test settings: Normal Distribution; Do not Bonferroni correct *p*-values. We sometimes have non-GFP control colonies on the plates, for example when retesting phenotypes associated with specific interactions after a proteome-wide screen. In these circumstances we use the “Designated Controls” options under Data Normalization Method, however in cases when we do not use the non-GFP controls we use the “Plate Median” option. When using the DR Engine you can choose to exclude any colonies from the analysis, for example because of a pinning issue or contamination on the plates. ScreenMill has many other options and features that might be of use to some users, but are not covered in this protocol.
10. –Leu + Ade media used for microscopy contains 100 mg/L adenine, which helps to minimize the amount of autofluorescent purine precursors within cells.
11. The cells collected for imaging should be growing in log phase for best results.
12. A total volume of 2.5  $\mu$ l liquid is appropriate for 22 mm coverslips, assuming the liquid spreads out over the whole coverslip area (4.83 cm<sup>2</sup>). This allows for 5.1  $\mu$ m distance between slide and coverslip—bear in mind that haploid yeast cells are ~5  $\mu$ m in diameter. Use increased liquid volume for diploid cells, since they are larger.

---

## Acknowledgments

This work was supported by a Medical Research Council (MRC) UK centenary award and grant to P.T. (MC\_UP\_A252\_1027). The Francis Crick Institute is funded by the MRC UK, Cancer Research UK, the Wellcome Trust, Imperial College London, University College London and Kings College London. We thank Lisa Berry for comments on the chapter.



## References

- Coudreuse D, Nurse P (2010) Driving the cell cycle with a minimal CDK control network. *Nature* 468:1074–1079
- Lau DT, Murray AW (2012) Mad2 and Mad3 cooperate to arrest budding yeast in mitosis. *Curr Biol* 22:180–190
- Hagan IM, Grallert A (2013) Spatial control of mitotic commitment in fission yeast. *Biochem Soc Trans* 41:1766–1771
- Scott JD, Pawson T (2009) Cell signaling in space and time: where proteins come together and when they're apart. *Science* 326:1220–1224
- Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H (2008) A versatile nanotrapp for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics* 7:282–289
- Rothbauer U, Zolghadr K, Tillib S, Nowak D, Schermelleh L, Gahl A, Backmann N, Conrath K, Muyldermans S, Cardoso MC, Leonhardt H (2006) Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods* 3:887–889
- Kubala MH, Kovtun O, Alexandrov K, Collins BM (2010) Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein Sci* 19:2389–2401
- Helma J, Cardoso MC, Muyldermans S, Leonhardt H (2015) Nanobodies and recombinant binders in cell biology. *J Cell Biol* 209:633–644
- Fridy PC, Li Y, Keegan S, Thompson MK, Nudelman I, Scheid JF, Oeffinger M, Nussenzweig MC, Fenyo D, Chait BT, Rout MP (2014) A robust pipeline for rapid production of versatile nanobody repertoires. *Nat Methods* 11:1253–1260
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R (1993) Naturally occurring antibodies devoid of light chains. *Nature* 363:446–448
- Muyldermans S (2013) Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* 82:775–797
- Grallert A, Chan KY, Alonso-Nunez ML, Madrid M, Biswas A, Alvarez-Tabares I, Connolly Y, Tanaka K, Robertson A, Ortiz JM, Smith DL, Hagan IM (2013) Removal of centrosomal PPI by NIMA kinase unlocks the MPF feedback loop to promote mitotic commitment in *S. pombe*. *Curr Biol* 23:213–222
- Ólafsson G, Thorpe PH (2015) Synthetic physical interactions map kinetochore regulators and regions sensitive to constitutive Cdc14 localization. *Proc Natl Acad Sci U S A* 112:10413–10418
- Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. *Nature* 425:686–691
- Reid RJ, Gonzalez-Barrera S, Sunjevaric I, Alvaro D, Ciccone S, Wagner M, Rothstein R (2011) Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I-induced DNA damage. *Genome Res* 21:477–486
- Butt TR, Sternberg EJ, Gorman JA, Clark P, Hamer D, Rosenberg M, Crooke ST (1984) Copper metallothionein of yeast, structure of the gene, and regulation of expression. *Proc Natl Acad Sci U S A* 81:3332–3336
- Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14:115–132
- Hill A, Bloom K (1989) Acquisition and processing of a conditional dicentric chromosome in *Saccharomyces cerevisiae*. *Mol Cell Biol* 9:1368–1370
- Dittmar JC, Reid RJ, Rothstein R (2010) ScreenMill: a freely available software suite for growth measurement, analysis and visualization of high-throughput screen data. *BMC Bioinformatics* 11:353
- Baryshnikova A, Costanzo M, Kim Y, Ding H, Koh J, Toufighi K, Youn JY, Ou J, San Luis BJ, Bandyopadhyay S, Hibbs M, Hess D, Gingras AC, Bader GD, Troyanskaya OG, Brown GW, Andrews B, Boone C, Myers CL (2010) Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. *Nat Methods* 7:1017–1024
- Collins SR, Schuldiner M, Krogan NJ, Weissman JS (2006) A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol* 7:R63

## **Reporter-Based Synthetic Genetic Array Analysis: A Functional Genomics Approach for Investigating Transcript or Protein Abundance Using Fluorescent Proteins in *Saccharomyces cerevisiae***

**Hendrikje Göttert, Mojca Mattiazzi Usaj, Adam P. Rosebrock, and Brenda J. Andrews**

### **Abstract**

Fluorescent reporter genes have long been used to quantify various cell features such as transcript and protein abundance. Here, we describe a method, reporter synthetic genetic array (R-SGA) analysis, which allows for the simultaneous quantification of any fluorescent protein readout in thousands of yeast strains using an automated pipeline. R-SGA combines a fluorescent reporter system with standard SGA analysis and can be used to examine any array-based strain collection available to the yeast community. This protocol describes the R-SGA methodology for screening different arrays of yeast mutants including the deletion collection, a collection of temperature-sensitive strains for the assessment of essential yeast genes and a collection of inducible overexpression strains. We also present an alternative pipeline for the analysis of R-SGA output strains using flow cytometry of cells in liquid culture. Data normalization for both pipelines is discussed.

**Key words** Yeast, Synthetic genetic array, SGA, Fluorescent protein, Reporter gene

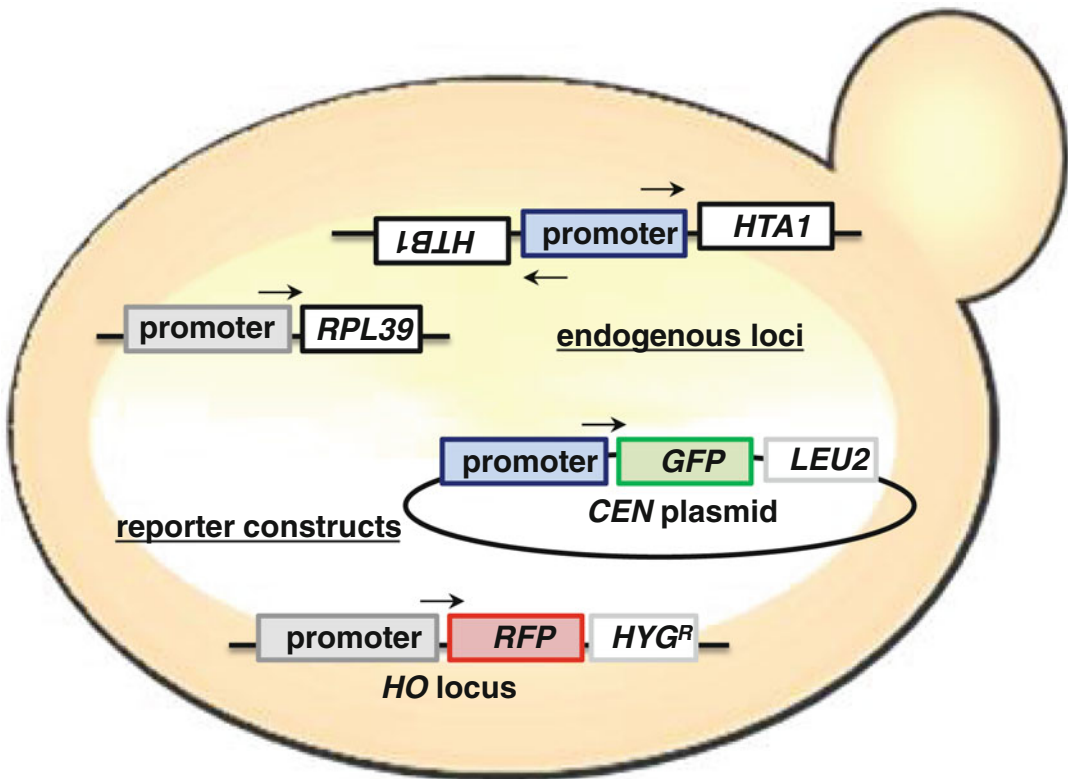
---

## **1 Introduction**

Reporter genes are useful tools for the screening of new regulators of transcript or protein abundance. For example, gene regulation can be indirectly monitored by fusing a promoter of interest to a fluorescent reporter gene such as Green Fluorescent Protein (GFP). Likewise, whole proteins can be tagged by fusing the GFP open reading frame to the gene encoding the protein of interest. Changes in gene expression or protein abundance can then be detected by assaying levels of fluorescence under various genetic and environmental conditions. A system called reporter synthetic genetic array analysis, or R-SGA, takes advantage of this concept [1]. R-SGA combines standard SGA technology [2] which allows

for the introduction of any kind of marked gene into any set of arrayed yeast strains with a dual-color reporter system. The R-SGA method presented here employs a wild-type query strain, which has the promoter of a gene of interest fused to a GFP gene carried on a plasmid, and a control promoter fused to an RFP gene integrated in the genome (Fig. 1). The query strain is crossed to an array of yeast mutants and, since all genes of interest are linked to selectable markers, a series of pinning steps on different selective media produces an output array in which each strain carries the fluorescent reporter genes and one mutation from the array strain. Fluorescence can then be imaged directly from colonies on a plate using a fluorescence plate scanner or strains can be transferred into liquid culture and assessed by flow cytometry to detect upregulation or downregulation of the fluorescent protein of interest in response to the tested perturbation.

The R-SGA method was first used to screen histone gene promoter–GFP fusion reporters in the KanMX deletion array [3], identifying a roster of loss-of-function mutants corresponding to



**Fig. 1** Reporter constructs to assess promoter activation. The query strain employed in the Reporter-Synthetic Genetic Array analysis (R-SGA) carries two reporter constructs with promoters driving fluorescent reporters. The example shows the reporter strain used to assess *HTA1* expression with *RPL39* promoter as control. Arrows indicate direction of transcription

putative activators and repressors of histone gene expression [4]. These screens revealed cell cycle-dependent regulation of active and repressive chromatin at histone gene promoters, controlled by a cell cycle oscillator that serves as a master regulator of histone gene activation [5]. A more recent study used the R-SGA methodology to assay for increased abundance of the strongly DNA damage-inducible protein Rnr3, to identify putative mutants exhibiting genome instability [6]. In addition to screening for genome instability mutants in the KanMX deletion array, which includes ~4200 nonessential mutant yeast genes, a set of ~1000 essential yeast alleles was assessed in this study, both in standard conditions and in the presence of the DNA-damaging agent methylmethane sulfonate [7, 8]. In theory the R-SGA approach is extensible to any colony-based assay for which the readout is abundance of a fluorescent protein, for example plasmid or chromosome loss readouts or assays for specific types of recombination.

---

## 2 Materials

### 2.1 Array and Query Strains and Plasmids

1. The R-SGA pipeline can easily be adapted for use with different yeast collections. Three examples are presented here: a haploid deletion collection [3], a collection of temperature sensitive (ts) mutants [7], and a collection of yeast strains with plasmids expressing genes with galactose inducible overexpression [9, 10]. Exemplar array and query strain and plasmid genotypes are summarized in Table 1 (*see Note 1*).
2. Promoters described here are defined as the intergenic sequence beginning at the first base pair (bp) upstream of a given open reading frame (ORF) and extending to the adjacent gene, up to a maximum of 1000 bp. This sequence is amplified by polymerase chain reaction (PCR) from a wild-type strain using custom-made primers and recombined into a CEN/ARS shuttle plasmid upstream of a promoterless GFP construct via transformation and homologous recombination in a wild-type yeast strain [11]. The plasmid is then extracted by plasmid rescue [12] and amplified in *E. coli* cells, then mini-prepped, sequenced, and transformed into a query strain that contains an integrated control promoter RFP reporter (*see Notes 2–4*).

### 2.2 SGA

1. Pinning robot or a hand pinning tool that can pin in 384 and 768 colony/plate format (*see Note 5*).
2. Rectangular media trays for solid media plates by Nunc or Singer (depending on the pinning robot/tool used).
3. SGA media

**Table 1**  
**Strains and plasmids used when screening the deletion, ts or FLEX array using R-SGA**

Plasmid	Genotype	Source
DELETION/TS query plasmid	<i>CEN ARS Amp<sup>R</sup> Promoter of interest-GFP LEU2</i>	Kainth et al. [1]
FLEX array plasmid pBY011	<i>CEN ARS Amp<sup>R</sup> URA3 GAL1pr-ORF</i>	Hu et al. [9]
FLEX query plasmid	<i>CEN ARS Amp<sup>R</sup> Promoter of interest-GFP NatMX</i>	Andrews Lab, University of Toronto
Strain	Genotype	Source
DELETION/TS array strain BY4741	<i>MAT<math>\alpha</math> his3<math>\Delta</math>1 leu2<math>\Delta</math>0 ura3<math>\Delta</math>0 met15<math>\Delta</math>0 ORF::KanMX</i>	Giaever et al. [3]
DELETION/TS query strain	<i>MAT<math>\alpha</math> ho:: Control promoter-tdTomato::HphMX can1<math>\Delta</math>:: STE2pr-S.p.His5 lyp1<math>\Delta</math> his3<math>\Delta</math>1 leu2<math>\Delta</math>0 ura3<math>\Delta</math>0 met15<math>\Delta</math>0</i>	Kainth et al. [1]
FLEX query strain	<i>MAT<math>\alpha</math> ho:: Control promoter -tdTomato::HphMX his3<math>\Delta</math>1 leu2<math>\Delta</math>0 ura3<math>\Delta</math>0 met15<math>\Delta</math>0</i>	Andrews Lab, University of Toronto

The deletion and the ts array strains are isogenic and thus the same query strains can be crossed to both arrays. The strain containing the FLEX overexpression collection has a different mating type and carries different selectable markers, therefore a different query strain needs to be employed. The control promoter of choice is integrated into the genome, while the promoter of choice to be tested is located on a plasmid.

	Deletion/TS array	FLEX array
Array plates	YPD + G418	SD –uracil
Query culture	SD –leucine	YPD + clonNat
Mating	YPD	YPD
Diploid selection (2 $\times$ )	SD –leucine + G418	SD –uracil + clonNat
Sporulation	Spo + ¼ G418	
Haploid selection (2 $\times$ )	SD –leucine –arginine –histidine –lysine + canavanine + thialysine + G418	
Final selection/ induction	SD –leucine –arginine –histidine –lysine + canavanine + thialysine + G418 + Hygromycin B	SGal –uracil + clonNat + Hygromycin B

YPD: 1% yeast extract, 2% peptone, 2% glucose with 40 mg/L adenine and 40 mg/L tryptophan.

*Synthetic dextrose medium (SD) without antibiotics:* 0.67% yeast nitrogen base without amino acids, 2% glucose plus complete amino acid mix (see below).

*SD medium containing G418, clonNat, or Hygromycin B:* 0.1% monosodium glutamic acid, 0.17% yeast nitrogen base without amino acids and without ammonium sulfate, 2% glucose plus complete amino acid mix (see below) (*see Note 6*).

*SGal and SRaf:* same as SD medium but with 2% galactose or raffinose as carbon source.

*Amino acid dropout mixes:* To make complete amino acid mix powder, grind or shake together for 15 min: 2 g each of alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, inositol, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tryptophan, uracil, valine plus 3 g adenine and 0.2 g para-aminobenzic acid. To make 10× complete amino acid mix for 1 L of SD, add 2 g of powder to 100 mL water and heat to dissolve. To make dropout, leave out the relevant amino acids or bases.

*Final antibiotic concentrations:* 200 µg/mL G418, 100 µg/mL clonNat, and 300 µg/mL Hygromycin B.

*Final amino acid analog concentrations:* 50 µg/mL canavanine and 50 µg/mL thialysine (S-(2-aminoethyl)-L-cysteine hydrochloride).

*Sporulation medium:* 1% potassium acetate, 0.1% yeast extract, 0.05% glucose, supplemented with 0.014 g/L uracil and histidine and 0.071 g/L leucine and antibiotic at one quarter standard concentration to retard the growth of mold.

Media composition is summarized in the table above and is the same for liquid culture as for solid plates but plates contain 2% agar (*see Note 7*).

### **2.3 Fluorescence Scanning**

1. Scanning fluorescence imager suitable for GFP and RFP image acquisition, e.g. Typhoon Trio Variable Mode Imager (GE Healthcare), Typhoon 7000 (GE Healthcare), or equivalent.
2. Software for image analysis such as GenePix Pro or SpotFinder.

### **2.4 Liquid Culture**

1. 96-well round bottom plates.
2. Breathable membrane covers, e.g. Breath-Easy by E&K Scientific.
3. HiGro shakers from Digilab (*see Note 8*).
4. Hand pinning tool, e.g. pin pad from Singer Instruments.

### **2.5 Flow Cytometry**

1. 384-well flat bottom polystyrene plates.
2. Multiwell plate mixer (Eppendorf MixMate).
3. Multiwell plate sonicator (QSonica MPH).
4. Flow cytometer suitable for GFP and RFP analysis, including high-throughput sampler and acquisition software, e.g. LSRII by BD Biosciences which is configured with a 488 nm and 561 nm laser for GFP and RFP analysis.

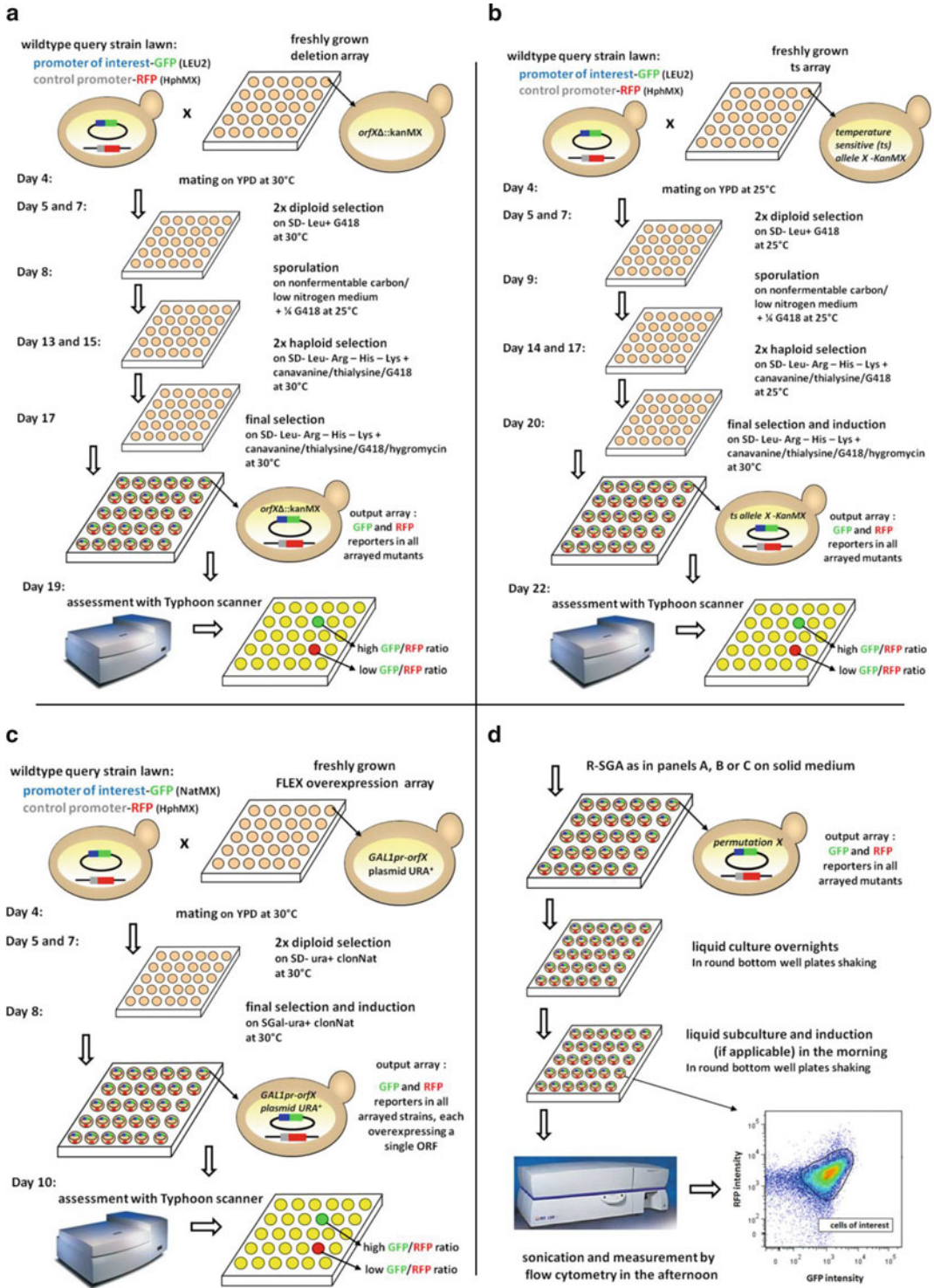
### 3 Methods

#### 3.1 R-SGA with a Haploid Deletion Array

A schematic diagram of the screening pipeline is depicted in Fig. 2a.

1. *Prepare query strain for lawns.* Set up an overnight culture of a single colony of the query strain containing the promoter of interest in 10 mL SD –leucine (Leu). Incubate at 30 °C for one night.
2. *Make lawns and assemble array in 768 format.* Spread 3 mL of the overnight culture on each of three YPD square agar plates. Incubate at 30 °C for two nights to grow confluent lawns of cells. Also re-pin the 14 plates of the deletion array, which are maintained in 384 colonies per plate format, into 768 format by pinning the same 384 colonies twice onto one plate diagonally to each other. Incubate at 30 °C for two nights.
3. *Mate.* Use a pinning robot (BioMatrix robot by S&P Robotics, Singer Rotor) or other pinning tool in 768 format to pin from the lawns onto fresh YPD square plates, up to six per lawn, without washing in between pinnings. Mate strains by pinning the 14 plates of the freshly grown deletion array in 768 format on top of the query strain, including a washing and sterilization step after each pinning action. Incubate at 30 °C overnight.
4. *Select for diploid cells that carry both the LEU2-marked promoter-GFP plasmid and a KanMX-marked gene deletion.* Pin from the mating plates onto SD-Leu + G418 and incubate at 30 °C for two nights.
5. *Repeat diploid selection.* Pin from diploid selection plates onto fresh SD-Leu + G418 and incubate at 30 °C for another night.
6. *Sporulate.* Pin from second diploid selection plates onto sporulation medium and incubate at 25 °C (room temperature) for five nights to allow for sufficient tetrad formation.
7. *Select haploid MATa strains carrying the GFP reporter plasmid and a gene deletion.* Pin from sporulation plates onto SD –Leu –arginine(Arg) –histidine(His) –lysine (Lys) + canavanine + thialysine + G418 and incubate at 30 °C for two nights.
8. *Repeat.* Pin from first haploid selection plates onto fresh SD –Leu –Arg –His –Lys + canavanine + thialysine + G418 and incubate at 30 °C for another two nights.
9. *Select for haploid MATa strains carrying the GFP reporter plasmid, the integrated control promoter (HphMX-marked) and a gene deletion.* Pin from second haploid selection plates onto SD –Leu –Arg –His –Lys + canavanine + thialysine + G418 + Hygromycin (Hyg) and incubate at 30 °C for two nights (*see Note 9*).





**Fig. 2** The R-SGA pipeline. Reporter-Synthetic Genetic Array analysis (R-SGA) uses multiple pinning and culture steps on selective plates to introduce two reporter cassettes into an array of yeast mutants. The pipeline can easily be adapted to screen reporter genes against different yeast arrays and with different means of assessing output strains. (a) The original protocol was designed to screen reporter genes against the



### 3.2 R-SGA with a Haploid Array Carrying Temperature-Sensitive Alleles of Essential Genes

A schematic diagram of the screening pipeline is depicted in Fig. 2b.

1. *Prepare query strain for lawn and assemble array in 768 format.* Set up an overnight culture of a single colony of the query strain containing the promoter of interest in 10 mL SD –Leu. Incubate at 30°C for one night. Re-pin the four plates of the ts array, which are maintained in 384 colonies per plate format, into 768 format by pinning the same 384 colonies twice onto one plate diagonally to each other. Incubate at 25°C for three nights.
2. *Make lawn.* Spread 3 mL of the overnight culture on a YPD square agar plate. Incubate at 30°C for two nights to grow a confluent lawn of cells.
3. *Mate.* Use a pinning robot in 768 format to pin from the lawn onto fresh YPD square plates, four per lawn, without washing in between pinnings. Then pin the four plates of the freshly grown ts array in 768 format on top of the pinned query strain, including a washing step after each pinning. Incubate at 25°C for one night to allow query and array strains to mate.
4. *Select for diploid cells that carry both the LEU2-marked promoter-GFP plasmid and a KanMX-marked ts allele.* Pin from the mating plates onto SD –Leu + G418 and incubate at 25°C for two nights.
5. *Repeat diploid selection.* Pin from first diploid selection plates onto fresh SD-Leu + G418 and incubate at 25°C for two nights.
6. *Sporulate.* Pin from second diploid selection plates onto sporulation medium and incubate at 25°C (room temperature) for five nights to allow for sufficient tetrad formation.
7. *Select for haploid MATa strains carrying the GFP reporter plasmid and a ts allele.* Pin from sporulation plates onto SD –Leu –Arg –His –Lys + canavanine + thialysine + G418 and incubate at 25°C for three nights.
8. *Repeat.* Pin from first haploid selection plates onto fresh SD –Leu –Arg –His –Lys + canavanine + thialysine + G418 and incubate at 25°C for three nights.

←

**Fig. 2** (continued) yeast KanMX deletion array. **(b)** Adaptation for use with the temperature sensitive (ts) collection involves growth at permissive temperature (25°C) and incubation at semipermissive temperature (30°C) prior to scanning. **(c)** A shorter pipeline applies when using an overexpression collection as it can be assessed in diploid strains. Prior to scanning, overexpression is induced by growth on galactose-containing medium. **(d)** To assess output strains by flow cytometry instead of colony scanning, cells are cultured in liquid medium, subcultured and placed at higher temperature or shifted to galactose-containing medium (as necessary for ts or overexpression) prior to scanning

9. *Select for haploid MATa strains carrying the GFP reporter plasmid, the integrated control promoter (HphMX-marked) and at the same time induce loss/reduction of function of ts alleles.* Pin from second haploid selection plates onto SD –Leu –Arg –His –Lys + canavanine + thialysine + G418 + Hyg and incubate at semipermissive temperature 30 °C for two nights (*see Note 9*).

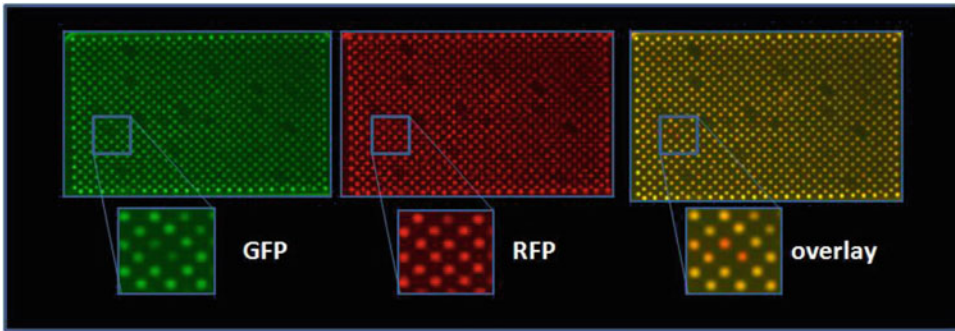
### 3.3 R-SGA with a Galactose-Inducible Overexpression Array

A schematic diagram of the screening pipeline is depicted in Fig. 2c.

1. *Prepare query strain for lawns.* Set up an overnight culture of a single colony of the query strain containing the promoter of interest in 10 mL YPD + clonNat. Incubate at 30°C for one night.
2. *Make lawns and assemble array in 768 format.* Spread 3 mL of the overnight culture on each of three YPD square agar plates. Incubate at 30°C for two nights to grow confluent lawns of cells. Also re-pin the 18 plates of the FLEX array, which are maintained in 384 colonies per plate format, into 768 format by pinning the same 384 colonies twice onto one plate diagonally to each other. Incubate at 30°C for two nights as well.
3. *Mate.* Use a pinning robot in 768 format to pin from the lawns onto fresh YPD square plates, six per lawn, without washing in between pinnings. Then pin the 18 plates of the freshly grown FLEX array in 768 format on top of the pinned query strain, including a washing step after each pinning. Incubate at 30°C for one night.
4. *Select for diploid cells which carry both the NatMX-marked promoter-GFP plasmid and the URA3-marked FLEX overexpression plasmid.* Pin from the mating plates onto SD –Ura + clonNat and incubate at 30°C for two nights.
5. *Repeat diploid selection and simultaneously induce overexpression by growth on galactose-containing medium.* Pin from diploid selection plates onto synthetic galactose medium (SGal) –Ura + clonNat + Hyg and incubate at 30°C for two nights (*see Notes 10 and 11*).

### 3.4 Analyzing Colony Fluorescence by Typhoon Scanning

1. Remove plates from incubator and cool to room temperature for at least 2 h before scanning to prevent condensation on the scanner surface (*see Note 12*).
2. Place plates upside down, without lids onto the cleaned scanning surface and tape edges to the Typhoon platen with clear tape to prevent movement due to vibration of the instrument (*see Notes 13 and 14*).
3. Use the Typhoon Scanner Control version 5.0 software with the following settings: acquisition mode = fluorescence, laser



**Fig. 3** Colony fluorescence detection using a Typhoon scanner. A plate with 768 yeast mutant colonies which carry a GFP and an RFP reporter were scanned using a Typhoon scanner. The GFP and the RFP channel are shown separately and as overlay. Two colonies of the same mutant strain display decreased GFP expression. The same colonies show unchanged expression levels of RFP and thus appear red in the overlay picture

1: 488 nm and emission filter 520 BP40, laser 2: 532 nm and emission filter: 580 BP30, pixel size: 100  $\mu\text{m}$ , focal plane: +3 mm. Photomultiplier tube (PMT) voltages are chosen empirically so that GFP and RFP intensities are below saturation with signals in a roughly similar range (*see Note 15*). An example of a scanned plate is displayed in Fig. 3.

### 3.5 Quantifying Colony Fluorescence with GenePix Software

1. Open both the GFP and RFP .gel Typhoon files in GenePix Pro version 6.0 software.
2. Assign appropriate wavelengths (488 nm for GFP and 532 nm for RFP).
3. To draw a template of circles for each colony on the plate, create “New Blocks” with the following features for 768 colonies/plate and replicate for all plates in the picture:
  - Number of columns: 24
  - Column spacing ( $\mu\text{m}$ ): 462
  - Number of rows: 32
  - Row spacing ( $\mu\text{m}$ ): 225
  - Feature diameter ( $\mu\text{m}$ ): 180
  - Feature Layout: Orange packing #3
4. Position grid by dragging it over colonies in image, aligning circles approximately with colonies. Click “Align Features in all Blocks” to automatically adjust circle size to actual colony size (*see Note 16*).
5. Click “Analyse” and find the results in the Results tab. Save results.

6. Use the median colony fluorescence intensity corrected for local background fluorescence for further normalization and analysis.

**3.6 Quantifying  
Colony Fluorescence  
Using SpotFinder  
Software (See Note 17)**

1. Upload in parallel .gel files of the GFP and RFP channel (ignore error messages).
2. To determine the location of colonies on each plate, a predefined grid is applied to the image and aligned manually for best fit. The two biological replicates present on the 768-colony R-SGA plate can be analyzed separately by applying a 384-spot grid.

For a typical R-SGA plate, grid parameters for one 384-colony replicate are:

Row number: 16

Column number: 24

$X$  spacing, pix: 45

$Y$  spacing, pix: 45

Meta row, Meta column, Pin  $X$ , and Pin  $Y$  values are adjusted depending on the number of plates present on the image and their positioning.

3. In the Gridding and Processing dialog, enter the Segmentation Method settings. In our hands, the Otsu Segmentation method produces good results. Min spot size and max spot size depend on the amount of colony growth and typically range between 5 and 40 pix. After adjusting the settings, the button "Process All" launches the segmentation and processing loop.
4. On the Data page, the user can review the segmentation output. The displayed data include information on the shape of the detected colonies and extracted numerical values.
5. After processing, save the data to a MEV file. Use the median colony fluorescence intensity corrected for local background fluorescence for further normalization and analysis.
6. Quality control filters (QC, QCA, and QCB) can be used to discard empty and badly processed colonies. Typically, a good colony will have a QC value of 0.4 or higher.

(See Note 18.)

**3.7 Normalization  
and Analysis of Solid  
Plate Data**

1. Assign GFP and RFP intensities obtained from GenePix or SpotFinder, as described above, to their proper gene or allele name according to their plate position, obtaining two values for each position when the screen was carried out in 768 format.

2. Remove border colonies from further analysis as these represent control strains to maintain similar colony sizes across the plate. Also remove negative intensity values indicating empty spots (*see Note 19*).
3. Divide each GFP value by the median GFP intensity of its plate to correct for plate specific variations (e.g., caused by different medium batches/thickness). Do the same for RFP.
4. Calculate  $\log_2(\text{GFP}/\text{RFP})$  ratios to correct for colony effects (e.g., size).
5. Subtract each value by the median  $\log_2(\text{GFP}/\text{RFP})$  ratio to center the whole screen at 0.
6. Average duplicate values for the two colonies next to each other on a plate.
7. Calculate *z*-scores by subtracting the screen median and dividing by the screen's median absolute deviation (MAD) from each averaged value (*see Notes 20 and 21*).

### **3.8 Liquid Culture and Assessment by Flow Cytometry**

To assess the final output array by flow cytometry, transfer the arrayed cells into liquid medium using a hand pinning tool. Do this before induction in case of overexpression and before shift to restrictive temperature in case of *ts* mutants. Proceed as follows (*see Fig. 2d*):

#### Day 1

1. Split each 768-format R-SGA output plate into two 384-format plates using a pinning robot to facilitate manual pinning into liquid medium.

#### Day 2 Evening

2. Transfer cells from freshly grown colonies on a solid plate into a 96-well round bottom plate with liquid medium, four per 384-format plate. Use a medium composition of SD –His –Leu –Arg –Lys + canavanine + thialysine + G418 for deletion and *ts* strains and synthetic raffinose medium (SR) –Ura + clonNat for overexpression strains, 200  $\mu\text{L}$  per well.
3. Cover plates with sterile breathable seals and incubate at 200 rpm overnight at 30 °C for deletion and overexpression strains and at 25 °C for *ts* strains.

#### Day 3 Morning

4. Subculture plates by transferring 5  $\mu\text{L}$  into 200  $\mu\text{L}$  fresh medium. Use the same medium as in the first step for deletion and *ts* strains. Induce overexpression strains in SGal –Ura + clonNat.

5. Culture all strains shaking at 30°C (semipermissive temperature for ts strains) for 5 h.

#### Day 3 Afternoon

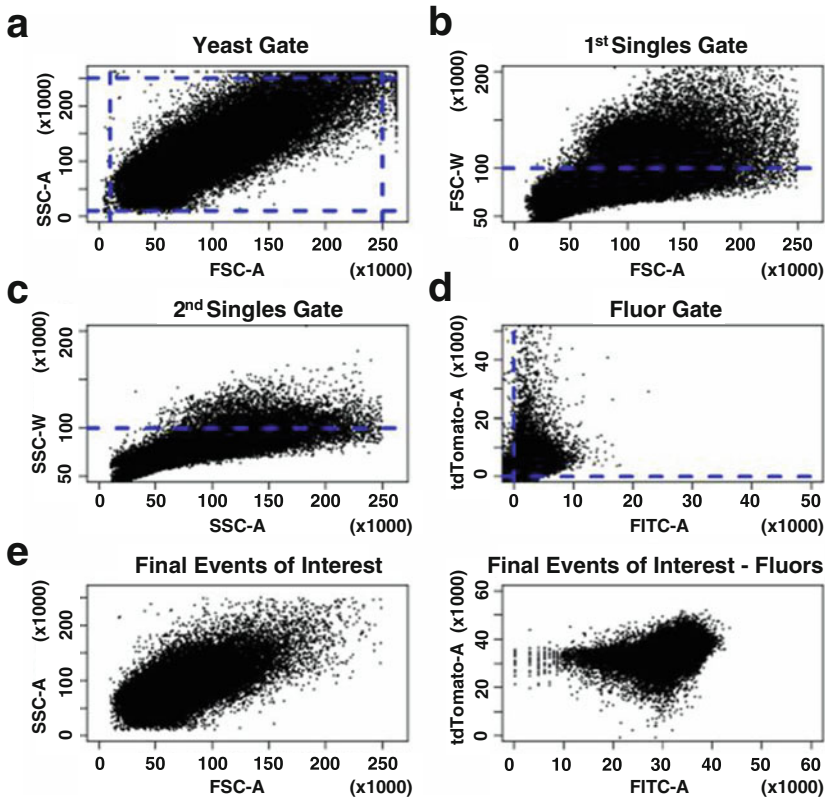
6. Shake the 96-well plates well using a plate mixer (Eppendorf MixMate or vortex).

Use a liquid handling robot to combine the four 96-well plates into one 384 flat bottom plate.

7. Sonicate 384 plates using a plate sonicator. We use 50% power for 60 seconds on a Qsonica MPH (This step should be empirically optimized by examining cells under a microscope to ensure that mother-daughter pairs have been separated).
8. Collect events from the 384 format plate using a flow cytometer equipped with an HTS by introducing 10 µL of culture per strain into the cytometer and collecting data for up to 50,000 cells per well:
  - Collect GFP, forward scatter (FSC), and side scatter (SSC) parameters from a 50 mW 488 nm laser through 510/20 nm and a pair of 488/10 nm filters, respectively.
  - Collect tdTomato levels from a 50 mW 561 nm laser through a 610/20 nm bandpass filter.
  - Collect both pulse area and width parameters for scatter parameters; collect pulse area only for fluorescence parameters.

### 3.9 Normalization and Analysis of Flow Cytometry Data

1. Open ungated cell population as forward scatter area (FSC-A, roughly indicative of cell size) vs. side scatter area (SSC-A, correlated with yeast size and budding status) scatterplot. Gate out debris and aggregates to select yeast-like events (Fig. 4a).
2. Open this gated cell population as an FSC-width (FSC-W) vs. FSC-A scatterplot. Gate out doublet events that are visible as a second mode on the FSC-W axis. These represent physically attached cells or cells that traversed the laser intercept nearly simultaneously (Fig. 4b).
3. Open gated cell population as SSC-W vs. SSC-A scatterplot. As in (step 2) above, gate out doublet events (Fig. 4c).
4. Open gated cell population as GFP-area (FITC-A) vs. RFP-area (tdTomato-A) scatterplot. Gate out events that do not express fluors properly (Fig. 4d) (*see Note 21*).
5. For the remaining events of interest (Fig. 4e) generate the median intensity for GFP and the median intensity for RFP for further analysis.
6. Remove wells with fewer than 2000 cells from further analysis.
7. Calculate  $\log_2(\text{GFP}/\text{RFP})$  ratios to correct for strain-specific effects (e.g. cell size).



**Fig. 4** Gating of flow cytometry data after acquisition. Gating allows for the exclusion of undesired data points or events prior to analysis. **(a)** Forward scatter area (FSC-A) is related to cell size. Side scatter area (SSC-A) reflects size and bud status. Excluding the extremes of FSC-A and SSC-A help to remove undesired events, e.g., debris, bacterial contamination, dead cells, or aggregates. **(b)** A gate using forward scatter width (FSC-W) excludes cases where two or more cells are stuck together or traverse the laser intercept simultaneously. **(c)** A gate using side scatter width (FSC-W) excludes additional cases where multiple cells were scanned together. **(d)** A gate excluding events with negative GFP (FITC-A) or tdTomato values removes cells which do not express the fluors properly. **(e)** Final events of interest after gating are used for further analysis

8. Determine the mean and standard deviation of all wild-type strain  $\log_2(\text{GFP}/\text{RFP})$  ratios.
9. We define two standard deviations above the mean wild-type  $\log_2(\text{GFP}/\text{RFP})$  as the cut-off to identify genes whose perturbation significantly upregulates expression of the tested promote and two standard deviations below the mean wild-type  $\log_2(\text{GFP}/\text{RFP})$  to identify genes whose perturbation significantly downregulates expression of the tested promoter (*see* **Notes 22** and **23**).

## 4 Notes

1. Colonies at the border of the plate grow faster than colonies inside the array due to decreased competition from neighbors.

Arrays should be built with a border (e.g. rows A and P, columns 1 and 24) consisting of an isogenic control strain that will be carried through all SGA steps. These colonies are excluded from analysis after data acquisition.

2. The two promoters should not be related. For example, if the goal is to explore cell cycle-regulated transcription, the promoter of interest would be cell cycle-regulated and the control promoter should not be cell cycle-regulated but rather constitutively expressed.
3. Having the GFP reporter on a plasmid facilitates query strain construction, especially when a large number of different promoters are to be screened. As the plasmid used here is a CEN plasmid, copy number variation is of negligible relevance when screening genome-wide. However, some mutations may cause increased plasmid loss, confounding results. To avoid this, the GFP reporter could be integrated into the genome.
4. The method can easily be adapted to assay readouts other than promoter activation.
5. Protocols are optimized for arrays pinned and grown in 768 colony format. Other formats can be used; however, incubation times need to be adapted as too small, too large or too variable colony sizes can make data normalization and analysis difficult. We have found, for example, that the colony size with 1534-colony format can be too small for accurate quantification.
6. Ammonium sulfate interferes with the sensitivity of cells to antibiotics [13]. We therefore use MSG as the nitrogen source in SD plates containing G418, clonNat, or hygromycin.
7. A plate pouring robot such as the Serial Filler from Singer, a calibrated pump or simply a graduated cylinder should be used to pour final selection plates for Typhoon scanning to ensure optimal and equal thickness of medium. Sixty milliliters is ideal when using rectangular plates from Nunc.
8. Other shakers can be used to culture cells in microwell plates but growth efficiencies might vary. Glass beads can be added to improve mixing and aeration but the additional stress on the cells has to be considered. We have found more dead cells are produced when glass beads are used.
9. An additional step could be added here, where colonies are subjected to different conditions, e.g. drug treatment.
10. Overexpression can be assessed in diploid cells as described here, saving time and materials. However, for certain studies it may be preferable to assess effects in haploid cells, which could be achieved by modifying the SGA pipeline.



11. Maximal induction of overexpression is achieved in <1 day; however, 2 days of growth are necessary to obtain colonies big enough to easily scan and analyze.
12. Preventing condensation is essential as imaging artifacts from moisture considerably affect results.
13. The Typhoon scanner can scan 12 plates at a time, thus scanning is done in two steps for large arrays such as the deletion and FLEX arrays. When scanning fewer than 12 plates, arrange them starting on the vertical (fast) axis on the left side of the scanner surface to minimize scanning time.
14. Remember that the scanner views the plates from below, thus the output picture will be inverted. When scanning multiple plates, it makes sense to place plate 1 in the top left corner and the last plate in the bottom right. Then choose the vertically mirrored R in the software to correctly display the output picture.
15. Fluorescence varies by strength of promoter, colony sizes, and even slight differences in media thickness; PMT voltages thus have to be optimized for each individual screen. To eliminate potential artefacts caused by photobleaching of the reporter proteins during the initial scan PMT should be determined on a “throwaway” plate.
16. Make sure all existing colonies are detected. If blocks and plates are poorly aligned, whole plates can be missed. Empty or undetected spots are indicated by  $\emptyset$ .
17. SpotFinder is an open-source multichannel image analysis tool which is part of the TM4 microarray software suite [14]. In our hands SpotFinder produces results comparable to GenePix.
18. All parameters and values are experiment-dependent and need to be adjusted accordingly for best performance. Detailed user instructions for grid positioning and other components of SpotFinder can be found in [15] and in the online manual (<http://www.tm4.org/spotfinder.html>).
19. In parallel colony sizes could be measured by photographing plates and running the pictures through an image analysis software like SGA Tools [16]. Very small colonies can then be eliminated from further analysis assuming that they give unreliable results.
20. For whole array screens,  $p$ -values can be calculated from  $z$ -scores and a statistical cut-off such as a 10% false discovery rate (FDR) can be used to determine hits. For small scale assays, for which the distribution is likely not normal,  $\log_2$  ratios should be compared between mutants and wild-type colonies on the same plate (not border colonies).
21. In any population of cells expressing fluorescent proteins, we find a small fraction of cells that have no GFP or RFP fluorescent signal.

22. We recommend confirming hits identified in a large scale screen by retesting them on mini-arrays using a strain carrying a different control promoter and comparing  $\log_2$  ratios between mutants and wild-type colonies on the same plate. We observe confirmation rates ranging from 40% to 70% depending on the promoter and yeast array used.
23. Assessment by flow cytometry can be used as a small scale orthogonal confirmation assay for hits identified on a large scale with the solid plate Typhoon method. We have found the two methods to agree with Pearson Correlation Coefficients around  $R = 0.7$ .

## References

1. Kainth P, Sassi HE, Pena-Castillo L, Chua G, Hughes TR, Andrews BJ (2009) Comprehensive genetic analysis of transcription factor pathways using a dual reporter gene system in budding yeast. *Methods* 48:258–264
2. Tong AH, Boone C (2006) Synthetic genetic array analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol* 313:171–192
3. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Vronneau S et al (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418:387–391
4. Fillingham J, Kainth P, Lambert JP, van Bakel H, Tsui K, Pena-Castillo L et al (2009) Two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. *Mol Cell* 35:340–351
5. Kurat CF, Lambert JP, van Dyk D, Tsui K, van Bakel H, Kaluarachchi S et al (2011) Restriction of histone gene transcription to S phase by phosphorylation of a chromatin boundary protein. *Genes Dev* 25:2489–2501
6. Hendry JA, Tan G, Ou J, Boone C, Brown GW (2015) Leveraging DNA damage response signaling to identify yeast genes controlling genome stability. *G3* 5:997–1006
7. Li Z, Vizeacoumar FJ, Bahr S, Warringer J, Li J, Vizeacoumar FS et al (2011) Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat Biotechnol* 29:361–367
8. Ben-Aroya S, Coombes C, Kwok T, O'Donnell KA, Boeke JD, Hieter P (2008) Toward a comprehensive temperature-sensitive mutant repository of the essential genes of *Saccharomyces cerevisiae*. *Mol Cell* 30:248–258
9. Hu Y, Rolfs A, Bhullar B, Murthy TV, Zhu C, Berger MF et al (2007) Approaching a complete repository of sequence-verified protein-encoding clones for *Saccharomyces cerevisiae*. *Genome Res* 17:536–543
10. Douglas AC, Smith AM, Sharifpoor S, Yan Z, Durbin T, Heisler LE et al (2012) Functional analysis with a barcoder yeast gene overexpression system. *G3* 2:1279–1289
11. van Leeuwen J, Andrews BJ, Boone C, Tan G (2015) Construction of multifragment plasmids by homologous recombination in yeast. *Cold Spring Harb Protoc* 9:pdb.top084111. prot085100
12. Robzyk K, Kassir Y (1992) A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Res* 20:3790
13. Cheng T-H, Chang C-R, Joy P, Yablok S, Gartenberg MR (2000) Controlling gene expression in yeast by inducible site-specific recombination. *Nucleic Acids Res* 28:e108
14. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N et al (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34:374–378
15. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA et al (2006) TM4 microarray software suite. *Methods Enzymol* 411:134–193
16. Wagih O, Usaj M, Baryshnikova A, VanderSluis B, Kuzmin E, Costanzo M et al (2013) SGA-tools: one-stop analysis and visualization of array-based genetic interaction screens. *Nucleic Acids Res* 41:W591–W596

# Chapter 41

## Statistical Analysis and Quality Assessment of ChIP-seq Data with DROMPA

Ryuichiro Nakato and Katsuhiko Shirahige

### Abstract

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) analysis can detect protein/DNA-binding and histone-modification sites across an entire genome. As there are various factors during sample preparation that affect the obtained results, multilateral quality assessments are essential. Here, we describe a step-by-step protocol using DROMPA, a program for user-friendly ChIP-seq pipelining. DROMPA can be used for quality assessment, data normalization, visualization, peak calling, and multiple statistical analyses.

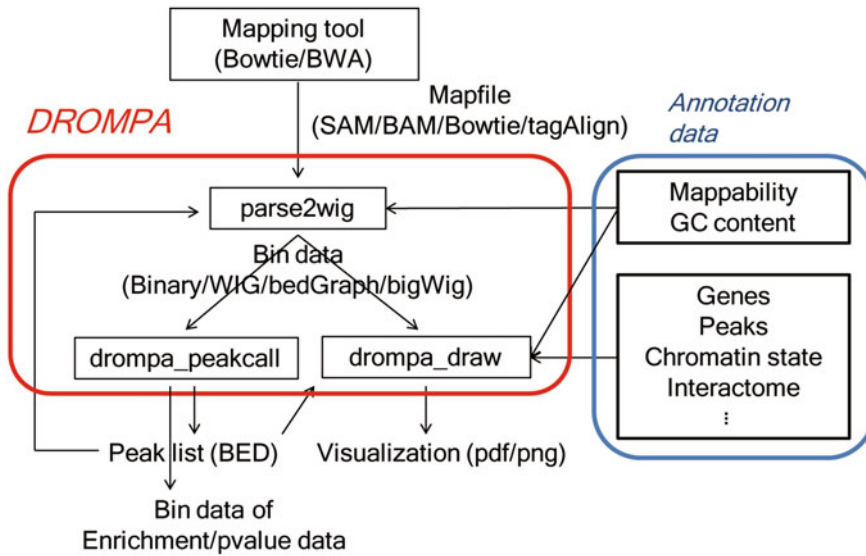
**Key words** Chromatin immunoprecipitation, ChIP-seq, High-throughput sequencing, Quality management, Visualization, Statistical analysis, Normalization

---

### 1 Introduction

Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) analysis was developed to identify protein/DNA-binding and histone-modification sites across an entire genome [1]. This assay is a mainstream method in genomics and epigenomics, and has led to important discoveries for various species [2–4]. A computational ChIP-seq analysis can be divided into several steps: (1) Mapping sequenced reads onto a reference genome; (2) Making data bins with normalized reads filtered for PCR biases; (3) Identifying significantly enriched regions (peaks); (4) Statistical analysis based on the obtained peak lists (e.g., aggregation plot). In this chapter, we describe a step-by-step protocol using DROMPA (DRow and Observe Multiple enrichment Profiles and Annotation), a program for user-friendly ChIP-seq pipelining [5]. The main features of DROMPA are:

- Any species whose genomic sequence is available can be used;
- Multiple input/output file formats (SAM, BAM, Bowtie, WIG, BED, TagAlign(.gz), bigWig, and bedGraph) are available;



**Fig. 1** Workflow of DROMPA pipeline. *Red box* indicates three internal programs of DROMPA and *blue one* does the reference annotations. The outputs of DROMPA include peak lists, pdf figures, and bin data of normalized read profiles

- Normalization using mappability and GC content biases that arise in ChIP-seq data;
- Output (in PDF or PNG format) of the read distribution, ChIP/input enrichment and *p*-values;
- In addition to typical peak calling, various types of ChIP-seq analysis are available.

Figure 1 shows the workflow of DROMPA (*see* Notes 1 in Section 4 for details on annotation data). From version 3.0.0 onward, DROMPA has three internal programs: *parse2wig*, *drompa\_peakcall*, and *drompa\_draw*. *parse2wig* preprocesses an input map file into bin data (the number of mapped reads per bin with fixed length). Generated bin data are used as input for both *drompa\_peakcall* and *drompa\_draw*. *drompa\_peakcall* calls peaks (peak calling) and *drompa\_draw* executes various types of visualization and quantitative analyses. DROMPA is available as an open-source C package with a detailed tutorial at <https://github.com/rnakato/DROMPA3>.

## 2 Materials

The examples are presented here as command lines to be used in UNIX shell prompts (e.g., bash). Commands are prefixed with “\$”. Comments are prefixed with “#” and can be ignored. Examples provided herein uses DROMPA version 3.2.3. Check the website for the latest version.

## 2.1 Installation

1. Installation of DROMPA requires several libraries and programs. On Ubuntu OS, these programs can be installed using the following commands:

```
$ sudo apt-get install git gcc libgtk2.0-dev libgs1-dev
samtools r-base
# download cpdf
$ wget http://github.com/coherentgraphics/cpdf-binaries/archive/master.zip
$ unzip master.zip
```

2. Install DROMPA through git:

```
$ git clone https://github.com/rnakato/DROMPA3.git
$ cd DROMPA3
$ make
```

If you get an installation error, make sure that all required libraries are installed.

3. Next, add the software directory to your PATH environment variable. For example, if you downloaded DROMPA and cpdf into the \$HOME/my\_chipseq\_exp directory, type:

```
$ export PATH = $PATH:$HOME/my_chipseq_exp/DROMPA3
$ export PATH = $PATH:$HOME/my_chipseq_exp/cpdf-binaries-master/Linux-Intel-**bit
```

## 2.2 Required Data: Genome Table File

DROMPA requires a genome table file, a tab-delimited file describing the name and length of each chromosome. Genome-table files can be generated by *makegenometable.pl* in the “scripts” directory as follows:

```
$ makegenometable.pl genome.fa > genometable.txt
```

Chromosome names in the genome table file and the reference genome should be identical. Hereafter, “genometable.txt” indicates this genome table file.

## 2.3 Read Mapping

ChIP-seq analysis with DROMPA starts with map files. For read mapping, we adopted the program Bowtie [6], which can distinguish unique and multiple mapped reads. To map reads (ChIP.fastq) onto the human genome build hg38 allowing uniquely mapped reads only, type:

```
$ bowtie index-UCSChg38 ChIP.fastq -m1 -S > ChIP.sam
```

To include multiple mapped reads, remove “-m1” option. See the manual for more information (<http://bowtie-bio.sourceforge.net/manual.shtml>). If the users want to treat insertions and deletions in reads, consider using other mapping tools such as Bowtie2 [7] and BWA [8].

---

## 3 Methods

### 3.1 Read Normalization and Bin Data Generation

*parse2wig* preprocesses an input map file into bin data (the number of mapped read per bin). The fragment length of each read is calculated automatically. For single-end mode (default), mapped reads are extended to the expected DNA-fragment length. In paired-end mode (with “-pair” option), each fragment length is obtained from the map file, and inter-chromosomal read-pairs and read-pairs longer than the maximum fragment length (specified by the “-maxins” option) are ignored.

The command:

```
$ parse2wig -f SAM -i ChIP.sam -o ChIP -gt genometable.txt -
binsize 100 -n GR -mp mappability/map_fragL150
```

creates a new directory “parse2wigdir”, and outputs bin files generated from the map file (ChIP.sam) with a bin size of 100 bp into it. A bin file is outputted for each chromosome (in the case of binary and WIG outputs) or for the whole-genome (in the case of bedGraph and bigWig outputs). In default, *parse2wig* filters “redundant reads” (reads starting exactly at the same 5' ends) as “PCR bias [1]”. When studying highly repetitive regions (e.g., rDNA regions in *Saccharomyces cerevisiae*), this filtering step should be omitted by supplying “-nofilter” option.

Read normalization is required for comparative ChIP-seq analyses that evaluate peak similarities and differences among samples [1]. Using “-n GR” option, *parse2wig* scales bin data so that the total number of mapped reads onto the whole genome is 20 million (by default, *see* Notes 2 for a point for read normalization). “-mp” option specifies the mappability files for normalizing each bin data based on the mappability. When “-mp” is not supplied, all bases are considered as mappable. These normalized bin data can also be visualized with other visualization tools such as IGV (<http://software.broadinstitute.org/software/igv/>) and UCSC genome browser (<https://genome.ucsc.edu/>).

### 3.2 Peak Calling

*drompa\_peakcall* identifies peaks by assuming a negative-binomial distribution as a background model and a binomial distribution for comparison between ChIP and input samples. Both ChIP and input samples should be specified with a single “-i” option using a comma as follows:

```
$ drompa_peakcall PC_SHARP -i parse2wigdir/ChIP,parse2wigdir/
Input -p ChIPpeak -gt genometable.txt
```

Since the bin files are chromosome-separated, use the prefix (before the underscore) to specify a sample file as an input file. The peak files “ChIPpeak.xls” and “ChIPpeak.bed” are outputted in a tab-delimited text file, which describe same peak sets in different formats. By default, chromosomes Y and mitochondria are ignored from peak calling. When mappability file is supplied (“-mp” option), the low mappable bins (<0.3 as default) are ignored. See Notes 3 for making bin data by `drompa_peakcall`.

### 3.3 Quality Assessment

In addition to the bin files, `parse2wig` also outputs the statistics of the input file into the output directory (the command in Subheading 3.1 produce the statistics file “ChIP.100.xls”). The statistics file describes various quality scores (e.g., read number, fragment length distribution, library complexity, GC contents, and FRiP score) [1]. The important issues are the number of “nonredundant reads” and “library complexity”. A low library complexity often occurs when a small amount of starting material is PCR over-amplified, resulting in a large amount of unreliable peaks. In other cases, if there are too few peaks obtained despite good quality scores, the chromatin immunoprecipitation (ChIP) step might not have worked. Check the antibody quality and the amount of ChIPed proteins.

### 3.4 Visualization of Read Profile

1. `drompa_draw` can execute various types of visualization including taking multiple ChIP-input pairs as input. Each pair should be specified with the option “-i”. For example, the command

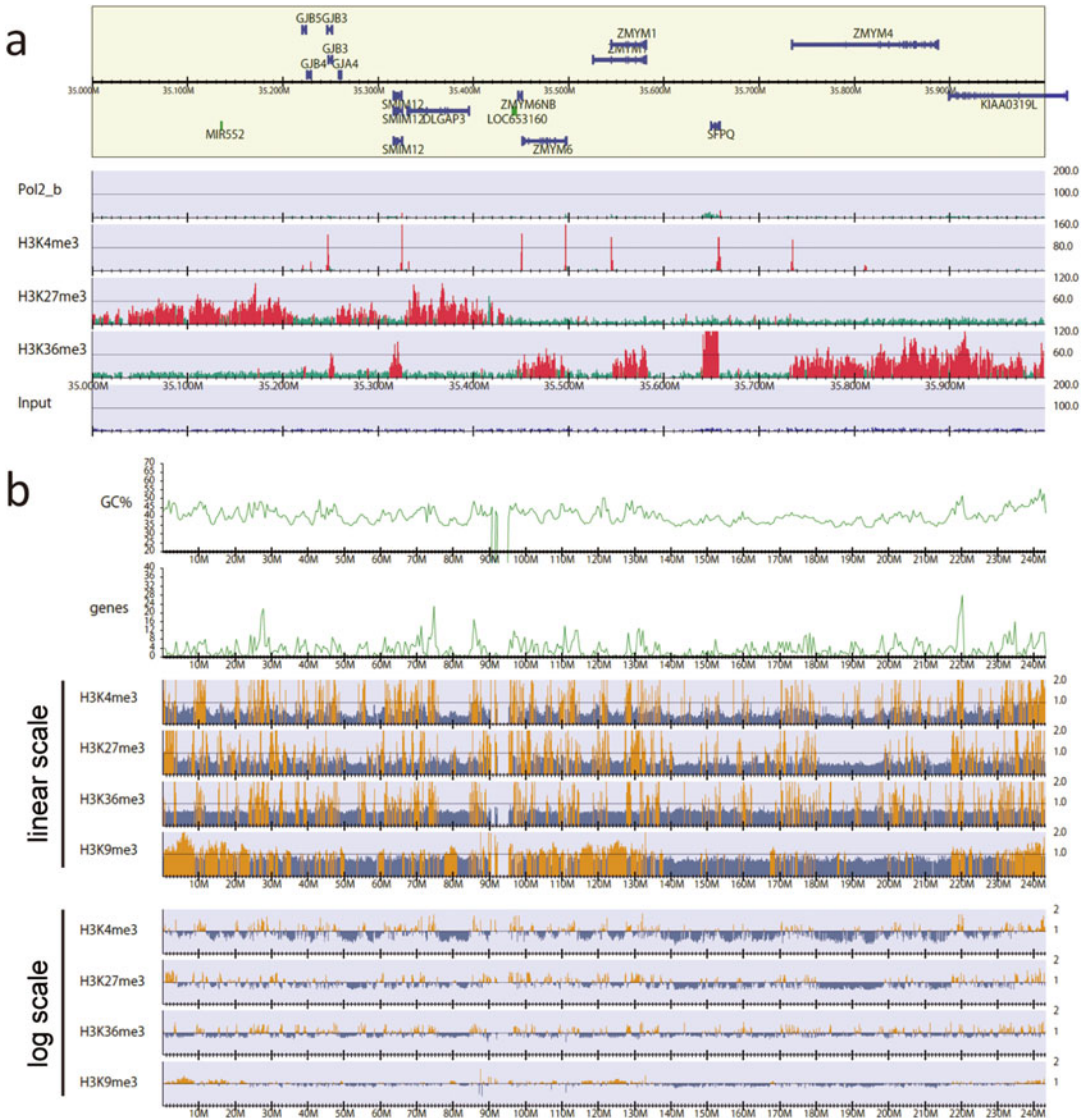
```
$ s1="-i parse2wigdir/Pol2_b,parse2wigdir/Control,Pol2_b"
$ s2="-i parse2wigdir/H3K4me3,parse2wigdir/Control,H3K4me3"
$ s3="-i parse2wigdir/H3K27me3,parse2wigdir/Control,H3K27me3"
$ s4="-i parse2wigdir/H3K36me3,parse2wigdir/Control,H3K36me3"
$ drompa_draw PC_SHARP -gt genometable.txt -gene refFlat.txt
$s1 $s2 $s3 $s4 -p HeLaS3 -ls 1000 -show_itag 2
```

generates the PDF files “HeLaS3.pdf” as shown in Fig. 2a (see Notes 4 for full details on “-i” option). “-ls 1000” indicates the length (kbp) per one line. By default, `drompa_draw` visualizes ChIP-read lines only. The “-show itag 2” option displays the input-read line at the bottom. Similarly, the users can display the lines of *p*-value (“-showpinter” and “-showpenrich” options) and ChIP/input enrichment line (“-showratio” option).

2. The GV mode visualizes a chromosome-wide overview (Fig. 2b) as follows:

```
$ drompa_draw GV -gt genometable.txt $s1 $s2 $s3 $s4 -p HeLaS3-
GV -GC GCcontents -gcsiz 500000 -GD genedensity -gdsiz 500000
```

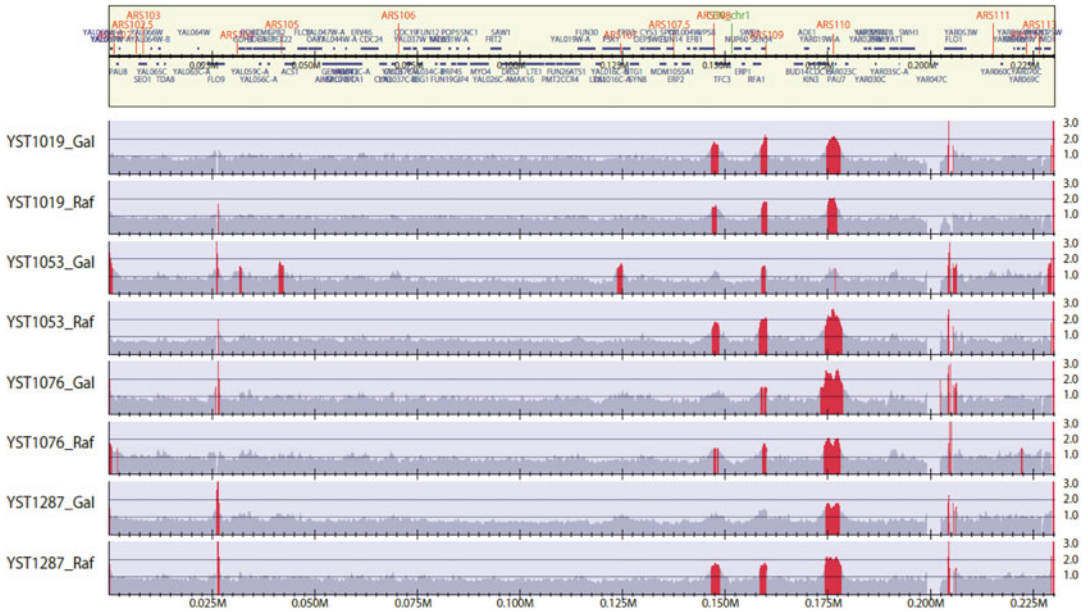




**Fig. 2** Visualization of ChIP-seq data for human HeLa-S3 cells. **(a)** PC\_SHARP mode: The normalized read distribution for four ChIP samples and one control sample for 100-bp bins, with a RefSeq gene annotation (chromosome 1, 35.0–36.0 Mb). Significantly enriched regions (*peaks*) are highlighted in *red*. **(b)** GV mode: Visualization of the ChIP/Control enrichment distribution for 100-kb bins (chromosome 2). The GC contents and gene numbers for 500 kb windows are also plotted

where “-GC GCcontents” and “-GD genedensity” specify the GC content files and the gene-density files, respectively. “-gcszise 500000” and “-gdszise 500000” specify the window sizes of them. The GV mode does not perform the significance test but simply highlights the bins containing ChIP/Input enrichments above the middle of the  $y$  axis (the value specified with the





**Fig. 3** Schematic representation of replication profile for *S. cerevisiae* (chromosome 1). Bins containing ChIP/Input enrichments above the threshold (defined by “-ethre” option) are highlighted in red

option “-scale ratio”) in orange. The bin size is 100k-bp and the ChIP/Input enrichment lines are shown by default (Fig. 2b upper). To make *y*-axis log scale, specify “-showratio 2” as follows:

```
$ drompa_draw GV -gt genometable.txt $s1 $s2 $s3 $s4 -p HeLaS3-
GV -GC GCcontents -gcsiz 500000 -GD genedensity -gdsiz 500000
-showratio 2
```

See Appendix 1 for the full commands to generate Fig. 2.

- For a small genome, such as the yeast’s, the sequencing depth is generally enough (>10-fold). In such cases, the genome-wide ChIP/Input enrichment distribution is informative, which minimizes the technical and biological bias in high throughput sequencing. To make a PDF file of the enrichment distribution for *S. cerevisiae* (Fig. 3), type:

```
$ drompa_draw PC_ENRICH -p SRP009385 -gt genometable.txt -ars
ARS-oriDB_scer.txt -gene SGD_features.tab -gftype 3 $s1 $s2
$s3 $s4 $s5 $s6 $s7 $s8 -ls 250 -lpp 2 -scale_ratio 1 -bn 3 -
ystep 14 -ethre 1.5
```

where “-lpp 2” specifies the number of lines for each page, and “-bn 3” and “-ystep 14” indicate the number and the height of

separations for  $y$ -axis. “-ethre 1.5” specifies the threshold for enrichment (red regions). When “-nosig” option is specified, all regions are colored in orange. See Appendix 2 for the full commands to generate Fig. 3.

### 3.5 Other Types of Visualization

*drompa\_draw* executes various types of visualization and quantitative analyses.

1. For example, the command PROFILE:

```
$ drompa_draw PROFILE -p aroundBED -gt genometable.txt -i
$IP,,ChIPname -i $IP2,,ChIPname2 -ptype 4 -bed peaklist.bed
```

outputs a PDF file (aroundgene.pdf) and an R script (aroundgene.R) of an aggregation plot around the specified bed regions (Fig. 4a).

2. The CI mode is for quantitative comparison of ChIP-seq samples.

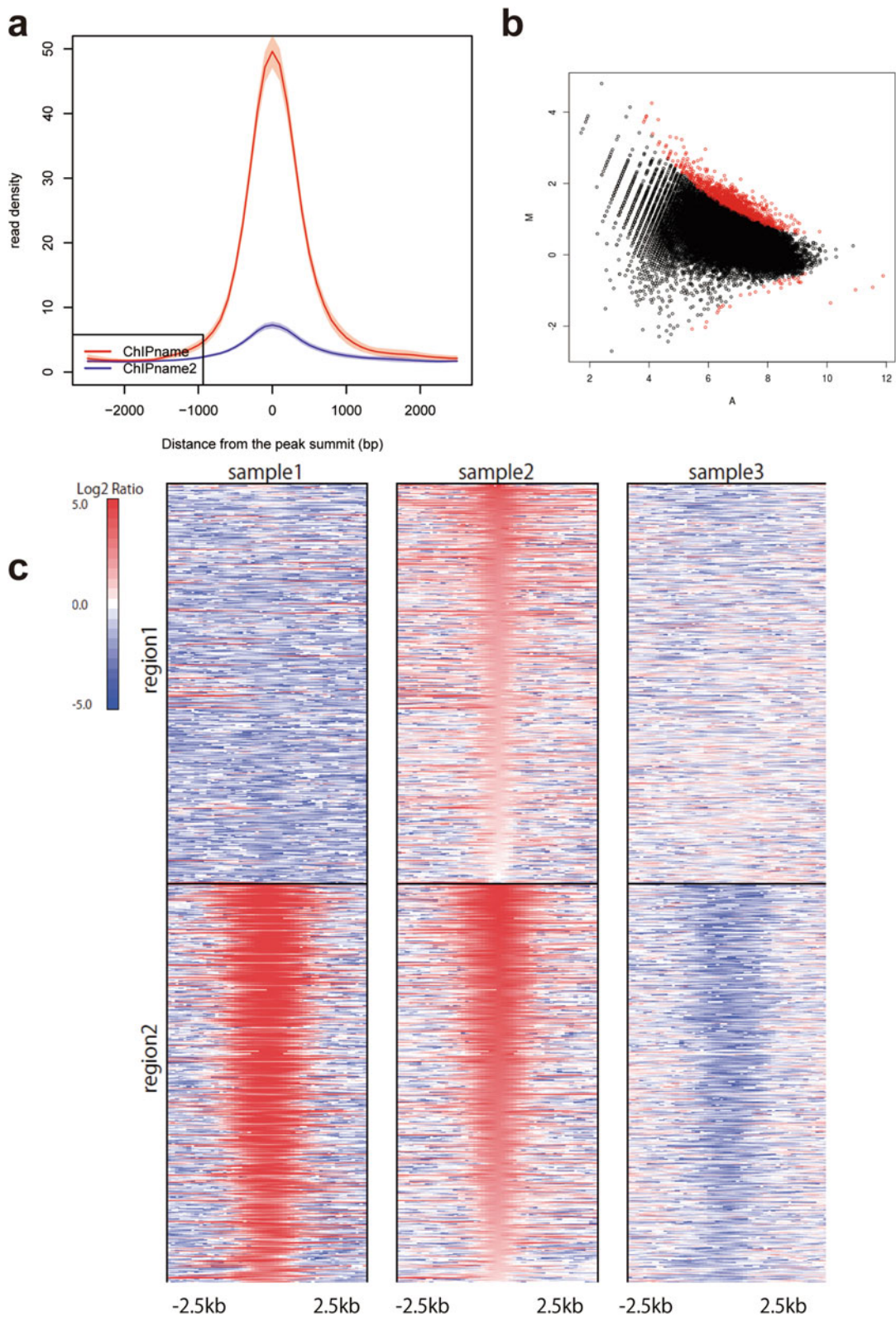
```
$ drompa_draw CI -p CIoutput -gt genometable.txt -bed sample.
bed -i $IP,,ChIPname -i $IP2,,ChIPname2
```

This mode takes two ChIP samples to be compared, and the output file contains the accumulated read number, the average log<sub>2</sub> read density (A), the log<sub>2</sub> ratio of read density between two samples (M) and the significance ( $-\log_{10}(p)$ ) of the difference based on a binomial test, for each peak specified by “-bed”. This output file can be used for making a MA plot of overlapped peak regions (Fig. 4b).

3. The HEATMAP mode outputs a heatmap of ChIP reads or ChIP/input enrichment profiles around target sites. The command:

```
$ drompa_draw HEATMAP -p heatmap -gt genometable.txt $s1 $s2
$s3 -stype 1 -ptype 4 -hmsort 2 -scale_ratio 5 -bed site1.bed,
region1 -bed site2.bed,region2 -png
```

generates the heatmap of ChIP/input enrichment (-stype 1) around the sites (-ptype 4) in two peak lists specified “-bed” option (Fig. 4c). The option “-hmsort 2” sorts the order of sites (rows) using the second ChIP sample. When specifying “-hmsort 0” (default), the sites are not sorted. With “-png” option, *drompa\_draw* outputs the figure in PNG format.



**Fig. 4** (a) PROFILE mode: Aggregation plot around peak regions. *Shaded regions* indicate a 95% confidence interval. (b) MA plot between M and A columns in the output file of CI mode. This MA plot is generated by *plot* function of R. The peaks in which  $-\log_{10}(p)$  is over 10 are colored in *red*. (c) HEATMAP mode: Heatmap for ChIP/input enrichment. Sample2 was used for sorting input sites. Each BED file is sorted individually

---

## 4 Notes

1. DROMPA accepts the following annotation data: Gene annotation (refFlat, gtf, and SGD\_features.tab format); Replication origin (for yeast); Mappability data; Interaction data (Hi-C, ChIA-PET); Repeat data (RepBase); GC contents; Bed12 annotations (e.g., peak regions, and chromatin state model generated by ChromHMM [9]). Some of these data are available on the DROMPA website (<http://www.iam.u-tokyo.ac.jp/chromosomeinformatics/rnakato/drompa/>).
2. It is noted that scaling a small number of reads up to a larger number (e.g., 1 million → 10 million) is not recommended because that will result in plenty of background noise.
3. For *drompa\_draw*, the option “-i” can take the following comma-separated multiple fields: (1) ChIP sample (required); (2) Input control sample; (3) Sample name to be shown in figure; (4) peak list to be highlighted; (5) binsize; (6) Scale of read; (7) Scale of ratio; (8) Scale of *p*-value. Except for the “ChIP sample”, all the other fields can be omitted. For example,

```
-i $ChIP,$Input,ChIPname,ChIPpeak.bed,1000,60
```

specifies 1000-bp bin size, 60 for *y*-scale of the read line, and the bed file “ChIPpeak.bed” as highlighted peak regions for this ChIP-input sample pair.

4. In addition to a peak file, *drompa\_peakcall* can also output bin data of the ChIP-read profile normalized the input sample with “-outputwig” option ((1) ChIP/Input ratio; (2) ChIP-internal *p*-value; (3) ChIP/Input enrichment *p*-value) for visualization with other tools. The bin data is outputted into the directory ‘drompadir’.

---

## Appendix

1. Script for the visualization of human (data obtained from the Sequence Read Archive (SRA) under accession SRP006944)

```
# convert map files to bin data
$ for sample in Control H3K27me3 H3K36me3 H3K4me3 H3K9me3
Pol2_b; do
$ for bin in 100 1000 100000;do parse2wig -f BAM -i HeLa-S3_
$$sample-n2-m1-hg19.sort.bam -o $$sample -gt genome_table -bin-
```

```

size $bin; done
$ done
# make pdf files
$ IP1="parse2wigdir/Pol2_b"
$ IP2="parse2wigdir/H3K4me3"
$ IP3="parse2wigdir/H3K27me3"
$ IP4="parse2wigdir/H3K36me3"
$ IP5="parse2wigdir/H3K9me3"
$ Input="parse2wigdir/Control"
# Figure 2a
$ s1="-i $IP1,$Input,Pol2_b,,100"
$ s2="-i $IP2,$Input,H3K4me3,,80"
$ s3="-i $IP3,$Input,H3K27me3,,1000,60"
$ s4="-i $IP4,$Input,H3K36me3,,1000,60"
$ drompa_draw PC_SHARP -gt genome_table -gene refFlat.txt $s1
$s2 $s3 $s4 -p Fig 2a -lpp 1 -chr 1 -ls 1000 -rmchr -show_itag 2
# Figure 2b
$ s1="-i $IP2,$Input,H3K4me3"
$ s2="-i $IP3,$Input,H3K27me3"
$ s3="-i $IP4,$Input,H3K36me3"
$ s4="-i $IP5,$Input,H3K9me3"
$ drompa_draw GV -gt genome_table $s1 $s2 $s3 $s4 -p Fig
2b_liner -GC GCcontents -gcsiz 500000 -GD genedensity -gdsiz
500000
$ drompa_draw GV -gt genome_table $s1 $s2 $s3 $s4 -p Fig 2b_log
-GC GCcontents -gcsiz 500000 -GD genedensity -gdsiz 500000 -
showratio 2

```

## 2. Script for the visualization of *S. cerevisiae* (data obtained from SRA under accession SRP009385)

```

$ gt=genome_table_sacCer3
$ index=UCSC-sacCer3-cs # bowtie index for colorspace data
# read mapping and parse2wig
$ for num in $(seq 398609 398624); do
$ prefix=SRR$num
$ bowtie -C $index $prefix.fastq -p8 -S > $prefix.sam
$ parse2wig -f SAM -i $prefix.sam -o $prefix -gt $gt
$done
# generate pdf files
$ dir=parse2wigdir
$ IP1_60="$dir/SRR398612$postfix" # YST1019 Gal 60min
$ IP1_0="$dir/SRR398611$postfix" # YST1019 Gal 0min
$ IP2_60="$dir/SRR398610$postfix" # YST1019 Raf 60min
$ IP2_0="$dir/SRR398609$postfix" # YST1019 Raf 0min
$ IP3_60="$dir/SRR398616$postfix" # YST1053 Gal 60min
$ IP3_0="$dir/SRR398615$postfix" # YST1053 Gal 0min
$ IP4_60="$dir/SRR398614$postfix" # YST1053 Raf 60min

```

```

$ IP4_0="$dir/SRR398613$postfix" # YST1053 Raf 0min
$ IP5_60="$dir/SRR398620$postfix" # YST1076 Gal 60min
$ IP5_0="$dir/SRR398618$postfix" # YST1076 Gal 0min
$ IP6_60="$dir/SRR398619$postfix" # YST1076 Raf 60min
$ IP6_0="$dir/SRR398617$postfix" # YST1076 Raf 0min
$ IP7_60="$dir/SRR398624$postfix" # YST1287 Gal 60min
$ IP7_0="$dir/SRR398623$postfix" # YST1287 Gal 0min
$ IP8_60="$dir/SRR398622$postfix" # YST1287 Raf 60min
$ IP8_0="$dir/SRR398621$postfix" # YST1287 Raf 0min
$ s1="-i $IP1_60,$IP1_0,YST1019_Gal"
$ s2="-i $IP2_60,$IP2_0,YST1019_Raf"
$ s3="-i $IP3_60,$IP3_0,YST1053_Gal"
$ s4="-i $IP4_60,$IP4_0,YST1053_Raf"
$ s5="-i $IP5_60,$IP5_0,YST1076_Gal"
$ s6="-i $IP6_60,$IP6_0,YST1076_Raf"
$ s7="-i $IP7_60,$IP7_0,YST1287_Gal"
$ s8="-i $IP8_60,$IP8_0,YST1287_Raf"
$ drompa_draw PC_ENRICH -p SRP009385 -gt $gt -ars ARS-oriDB_s-
cer.txt -gene SGD_features.tab -gftype 3 $s1 $s2 $s3 $s4 $s5
$s6 $s7 $s8 -ls 250 -lpp 2 -scale_ratio 1 -bn 3 -ystep 14 -
ethre 1.5

```

## References

1. Nakato R, Shirahige K (2017) Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. *Brief Bioinform* 18:279. doi:[10.1093/bib/bbw023](https://doi.org/10.1093/bib/bbw023)
2. Gerstein MB, Rozowsky J, Yan KK, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Li JJ, Pei B, Harmanci AO, Duff MO, Djebali S, Alexander RP, Alver BH, Auerbach R, Bell K, Bickel PJ, Boeck ME, Boley NP, Booth BW, Cherbas L, Cherbas P, Di C, Dobin A, Drenkow J, Ewing B, Fang G, Fastuca M, Feingold EA, Frankish A, Gao G, Good PJ, Guigo R, Hammonds A, Harrow J, Hoskins RA, Howald C, Hu L, Huang H, Hubbard TJ, Huynh C, Jha S, Kasper D, Kato M, Kaufman TC, Kitchen RR, Ladewig E, Lagarde J, Lai E, Leng J, Lu Z, MacCoss M, May G, McWhirter R, Merrihew G, Miller DM, Mortazavi A, Murad R, Oliver B, Olson S, Park PJ, Pazin MJ, Perrimon N, Perouchine D, Reinke V, Reymond A, Robinson G, Samsonova A, Saunders GI, Schlesinger F, Sethi A, Slack FJ, Spencer WC, Stoiber MH, Strasbourger P, Tanzer A, Thompson OA, Wan KH, Wang G, Wang H, Watkins KL, Wen J, Wen K, Xue C, Yang L, Yip K, Zaleski C, Zhang Y, Zheng H, Brenner SE, Graveley BR, Celniker SE, Gingeras TR, Waterston R (2014) Comparative analysis of the transcriptome across distant species. *Nature* 512(7515):445–448. doi:[10.1038/nature13424](https://doi.org/10.1038/nature13424)
3. Sutani T, Sakata T, Nakato R, Masuda K, Ishibashi M, Yamashita D, Suzuki Y, Hirano T, Bando M, Shirahige K (2015) Condensin targets and reduces unbound DNA structures associated with transcription in mitotic chromosome condensation. *Nat Commun* 6:7815. doi:[10.1038/ncomms8815](https://doi.org/10.1038/ncomms8815)
4. Consortium RE, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfening AR, Wang X, Clausnitzer M, Liu Y, Coarfa C, Harris RA, Shores N, Epstein CB, Gjonjeska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJ, Li W, Marra MA, McManus MT,

- Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518(7539):317–330. doi:[10.1038/nature14248](https://doi.org/10.1038/nature14248)
5. Nakato R, Itoh T, Shirahige K (2013) DROMPA: easy-to-handle peak calling and visualization software for the computational analysis and validation of ChIP-seq data. *Genes Cells* 18(7):589–601. doi:[10.1111/gtc.12058](https://doi.org/10.1111/gtc.12058)
6. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25. doi:[10.1186/gb-2009-10-3-r25](https://doi.org/10.1186/gb-2009-10-3-r25)
7. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359. doi:[10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923)
8. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760. doi:[10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324)
9. Ernst J, Kellis M (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9(3):215–216. doi:[10.1038/nmeth.1906](https://doi.org/10.1038/nmeth.1906)



## Quantitative Analysis of DNA Damage Signaling Responses to Chemical and Genetic Perturbations

Francisco M. Bastos de Oliveira, Dongsung Kim, Michael Lanz, and Marcus B. Smolka

### Abstract

Phosphorylation-mediated signaling is essential for maintenance of the eukaryotic genome. The evolutionarily conserved kinases ATR and ATM sense specific DNA structures generated upon DNA damage or replication stress and mediate an extensive signaling network that impinges upon most nuclear processes. ATR/ATM signaling is highly regulated and can function in a context-dependent manner. Thus, the ability to quantitatively monitor most, if not all, signaling events in this network is essential to investigate the mechanisms by which kinases maintain genome integrity. Here we describe a method for the *Quantitative Mass-Spectrometry Analysis of Phospho-Substrates* (QMAPS) to monitor in vivo DNA damage signaling in a systematic, unbiased, and quantitative manner. Using the model organism *Saccharomyces cerevisiae*, we provide an example for how QMAPS can be applied to define the effect of genotoxins, illustrating the importance of quantitatively monitoring multiple kinase substrates to comprehensively understanding kinase action. QMAPS can be easily extended to other organisms or signaling pathways where kinases can be deleted or inhibited.

**Key words** DNA damage checkpoint, DNA damage signaling, Phosphorylation, Quantitative mass spectrometry, *Saccharomyces cerevisiae*

---

## 1 Introduction

The integrity of eukaryotic genomes relies on the action of the evolutionary conserved PI3K-like kinases ATR and ATM, which detect aberrant DNA structures and mediate complex signaling responses [1]. In *Saccharomyces cerevisiae*, DNA damage signaling is initiated by Mec1 and Tel1 (orthologs of human ATR and ATM) and further transduced to the downstream effector kinase Rad53 (ortholog of human CHK2) [2, 3]. These kinases phosphorylate a diverse array of proteins and recent proteomic studies have identified over 100 phosphorylation events that depend on Mec1 and/or Tel1 [4–7]. Despite the complexity of the DNA damage signaling network, most studies typically rely on a few selected substrates as



readouts of the overall activity of the kinases involved. Since Mec1 and Tel1 can target different substrates under different conditions [5], probing one or a few substrates will not capture all aspects of Mec1/Tel1 regulation and action.

To comprehensively assess DNA damage signaling, our laboratory has recently developed a method for the Quantitative Mass-Spectrometry Analysis of Phospho-Substrates (QMAPS). This proteomic approach allows us to monitor the dynamics of DNA damage signaling by looking at multiple *in vivo* kinase substrates in a systematic, unbiased, and quantitative manner [5]. Our method employs Stable Isotope Labeling of Amino acids in Cell culture (SILAC), followed by Immobilized Metal Affinity Chromatography (IMAC) and LC-MS/MS to compare the relative abundance of specific kinase substrates in different conditions [8, 9].

Importantly, QMAPS analysis implicitly relies on a predefined list of kinase-dependent phosphorylation events. In order to define a set of kinase-dependent phosphorylation events, the phosphoproteomes of wild type cells (WT) are compared against the phosphoproteomes of mutants lacking kinases involved in DNA damage signaling (*see* Fig. 1). Phosphorylation events dependent on these kinases can be quantitatively monitored upon a range of different conditions, such as treatment with different genotoxins or mutations/deletions of specific genes. As an example for this protocol, here, we use the Mec1/Tel1 and Rad53 phospho-substrates to monitor differences in signaling events induced by two different genotoxic drugs (*see* Figs. 2 and 3).

---

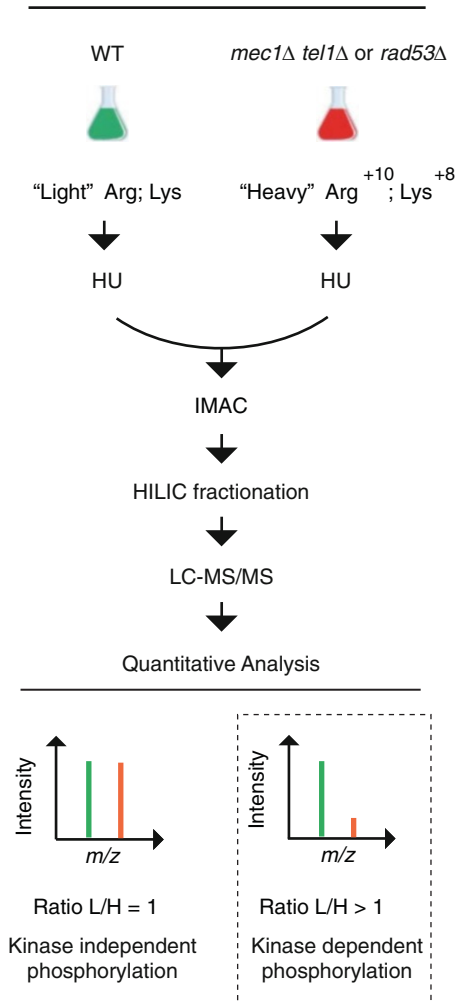
## 2 Materials

### 2.1 Stable Isotope Labeling of Amino Acids in Cell Culture (SILAC)

1. *Saccharomyces cerevisiae*: Strains MBS164 (*MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1::ΔTRP1, bar1::ΔHIS3*), MBS2042 (*MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1::ΔTRP1, bar1::ΔHIS3, mec1::ΔURA3, tel1::ΔkanMX6*) and MB188 (*MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1::ΔTRP1, bar1::ΔHIS3, rad53::ΔURA3*) (*see* Note 1).
2. Culture media: Synthetic Defined (SD) –arginine –lysine media for SILAC: 0.67 g of Complete Supplement Mixture (CSM) –arg –his –lys, 6.7 g of YNB, 80 mg of L-proline, 40 mg of L-histidine, and 960 ml of distilled water. After autoclaving, add 40 ml of a 50% dextrose solution to a final concentration of 2% (*see* Note 2).
3. “Light” arginine/lysine stock solution (1000×): 30 mg/ml of L-arginine (0.2 M), 20 mg/ml of L-lysine (0.115 M). Dissolve amino acids in distilled water and filter using a 0.22 μm disposable syringe filter.

## Step 1

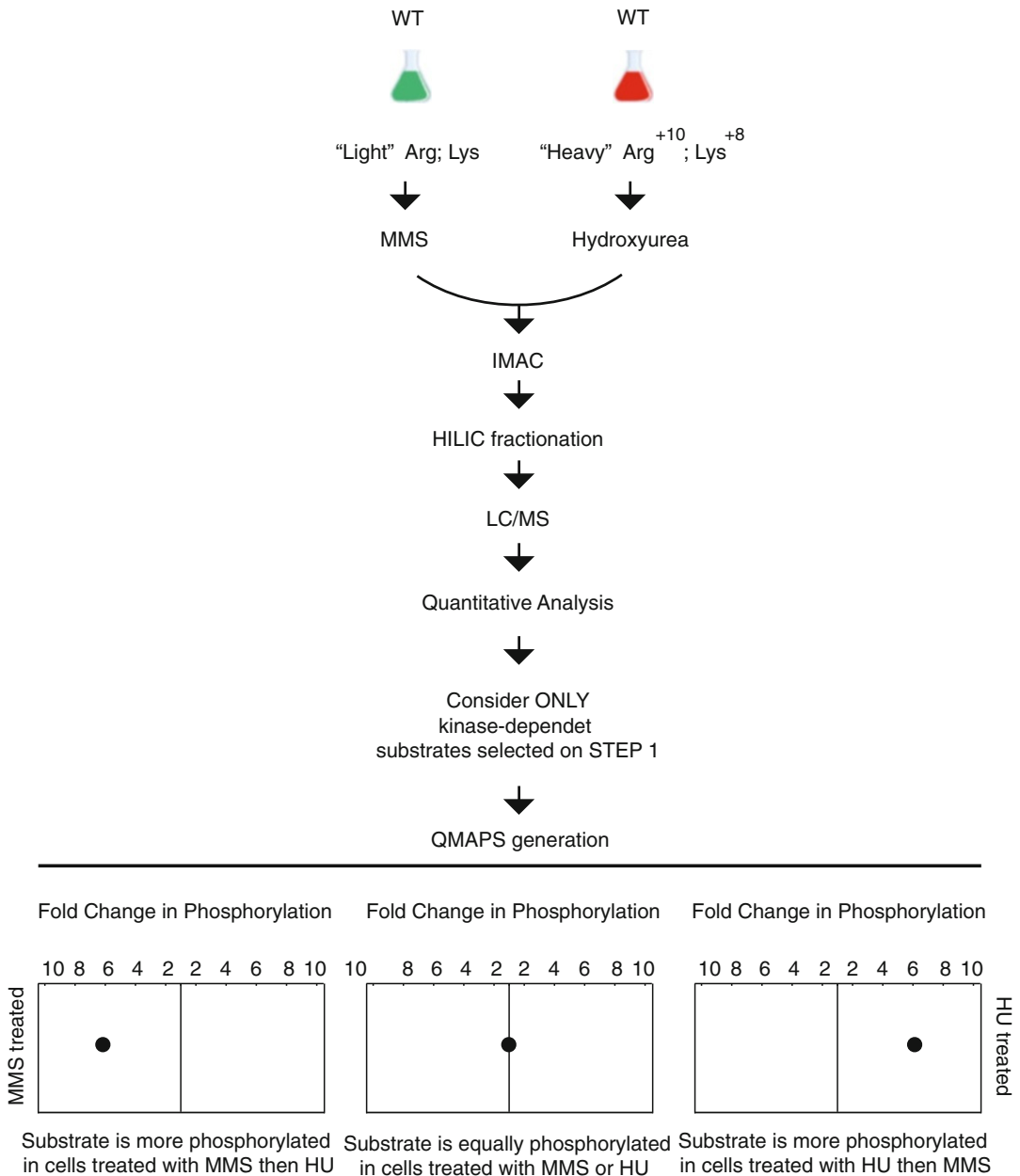
Defining kinase-dependent substrates during DNA damage signaling



**Fig. 1** Overview of the methodology for the identification of Mec1/Tel1 or Rad53-dependent substrates during DNA damage response. Wild type and kinase null mutant strains were grown in “Light” and “Heavy” media, respectively. Cells were arrested in G1 with alpha-factor and released in media containing 0.04% of methyl methanesulfonate (MMS) or 100 mM of hydroxyurea (HU). After treatment, cultures were combined, proteins were extracted, digested with trypsin, desalted, and subjected to phosphopeptide enrichment using IMAC. Phosphopeptide elutions were fractionated by HILIC and subjected to LC-quantitative mass spectrometry analysis. Phosphopeptide ratio equal or lower than 1 suggests a checkpoint kinase-independent or dependent phosphorylation, respectively. *Dashed line* highlights the quantitation profile expected for the checkpoint kinase-dependent event

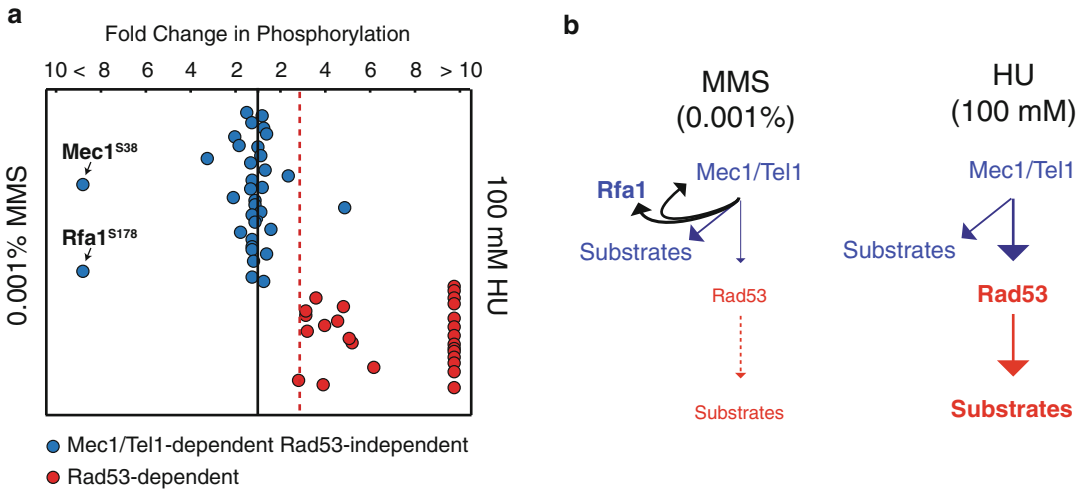
## Step 2

Defining chemical perturbations to kinase-dependent DNA damage signaling



**Fig. 2** Overview of the methodology for the QMAPS in cells treated with methyl methanesulfonate or hydroxyurea. Wild type cells were arrested in G1 with alpha-factor and released in “Heavy” or “Light” media containing 0.001% of methyl methanesulfonate (MMS) or 100 mM of hydroxyurea (HU), respectively. After incubation, cultures were combined, proteins were extracted, digested with trypsin, desalted, and subjected to phosphopeptide enrichment using IMAC. Phosphopeptide elutions were fractionated by HILIC and subjected to LC–quantitative mass spectrometry analysis. To generate the QMAPS, results of the

4. “Heavy” arginine/lysine stock solution (1000×): 38 mg/ml of L-arginine  $^{13}\text{C}_6$ ,  $^{15}\text{N}_4$  (0.2 M), 25 mg/ml of L-lysine  $^{13}\text{C}_6$ ,  $^{15}\text{N}_2$  (0.115 M). Dissolve amino acids in distilled water and filter using a 0.22  $\mu\text{m}$  disposable syringe filter.
5. To prepare 1 L of SD “Light” or SD “Heavy” add 1 ml of 1000 × “Light” arg/lys or 1000 × “Heavy” arg/lys stock solution to 1 L of SD –arg –lys.



**Fig. 3** QMAPS reveals uncoupling of Mec1/Tel1 and Rad53 signaling. **(a)** QMAPS depicting the fold change in phosphopeptide abundance between cells treated with MMS or HU. *Blue dots* correspond to Mec1/Tel1-dependent and Rad53-independent substrates and *red dots* correspond to Rad53-dependent substrates. *Red dashed line* highlights the uncoupling between Mec1/Tel1 and Rad53-dependent signaling in cells treated with MMS or HU. **(b)** Schematic representation of Mec1/Tel1 and Rad53 signaling in cells treated with MMS or HU. Global analysis of DNA damage signaling allowed us to identify an increase in Mec1 autophosphorylation and Mec1-dependent Rfa1 phosphorylation in cells treated with 0.001% MMS compared to HU-treated cells. We also observed an uncoupling between Mec1/Tel1 and Rad53 signaling in cells treated with 0.001% MMS. *Blue arrows* represent Mec1/Tel1-dependent and Rad53-independent phosphosignaling and *red arrows* represent Rad53-dependent phosphosignaling. *Large, medium, and small arrows* represent the variations on phosphosignal intensity

**Fig. 2** (continued) phosphoproteome analysis were filtered using the list of kinase-dependent phosphopeptides identified in step 1 (see Fig. 1). In the QMAPS, we depict the fold change in phosphopeptide abundance converted from “Light” versus “Heavy” relative abundances. Each dot corresponds to a different checkpoint kinase-dependent phosphopeptide identified at least three times in two independent biological replicates. Phosphopeptides deviated to the left or to the right side of the plot indicate an increase in their abundance in cells treated with MMS or HU, respectively. *Dot* placed in the middle of the plot indicates same level of phosphorylation in both conditions

## **2.2 Cell Cycle Arrest, Genotoxin Treatment, and Cell Harvesting**

1. Alpha-factor mating pheromone.
2. Protease from *Streptomyces griseus*.
3. Hydroxyurea.
4. Methyl methanesulfonate.
5. TE buffer: 10 mM Tris-HCl pH 8.0 and 5 mM EDTA.

## **2.3 Cell Lysis**

1. Lysis buffer: 50 mM Tris-HCl pH 8.0, 0.2% Tergitol, 150 mM NaCl, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 × complete EDTA-free protease inhibitor cocktail (Roche), and 1 × phosphatase inhibitor (*see Note 3*).
2. 100 × phosphatase inhibitor: 100 mM Na<sub>3</sub>VO<sub>4</sub>, 500 mM NaF, and 1 M β-glycerophosphate.
3. Glass beads, 0.5 mm dia.
4. Protein assay dye reagent.
5. Bead Beater type cell disruptor.

## **2.4 Protein Precipitation and Tryptic Digestion**

1. Alkylation solution: 1 M Tris-HCl pH 8.0 and 0.5 M iodoacetamide (*see Note 4*).
2. Protein precipitation (PPT) solution: 50% acetone, 49.9% ethanol and 0.1% biochemistry grade acetic acid.
3. Water bath sonicator.
4. Urea-Tris solution: 100 mM Tris pH 8.0 and 8 M urea (*see Note 5*).
5. NaCl-Tris solution: 50 mM Tris-HCl pH 8.0 and 150 mM NaCl.
6. TPCK treated trypsin 10 mg/ml in 0.1% biochemistry grade acetic acid (*see Note 6*).

## **2.5 Protein Sample Clean-Up**

1. 10% LC-MS grade formic acid (*see Note 7*).
2. 10% trifluoroacetic acid (*see Note 7*).
3. C<sub>18</sub> cartridges Vac 1 cm<sup>3</sup>, 1 G (Sep-Pak).
4. C<sub>18</sub> buffer A: 0.1% trifluoroacetic acid (*see Note 7*).
5. C<sub>18</sub> buffer B: 0.1% biochemistry grade acetic acid (*see Note 7*).
6. C<sub>18</sub> buffer C: 80% HPLC grade acetonitrile and 0.1% biochemistry grade acetic acid (*see Note 7*).
7. Silanized glass vial, 300 μl.
8. Silanized glass vial, 2 ml.

**2.6 Preparation of Immobilized Metal Affinity Chromatography (IMAC) Resin**

1. Ni-NTA (silica) spin columns (Qiagen).
2. Stripping solution: 50 mM EDTA and 1 M NaCl.
3. Iron (III) chloride solution: 100 mM FeCl<sub>3</sub> in 0.3% biochemistry grade acetic acid (*see Note 8*).
4. IMAC washing solution 1: 25% acetonitrile, 0.1 M NaCl, and 0.1% biochemistry grade acetic acid.

**2.7 Phosphopeptide Enrichment Using IMAC**

1. Glass wool.
2. Gel loading tips, 200  $\mu$ l.
3. 1 ml slip-tip disposable syringe w/o needle.
4. IMAC washing solution 1: 25% acetonitrile, 0.1 M NaCl, and 0.1% biochemistry grade acetic acid.
5. IMAC washing solution 2: 1% biochemistry grade acetic acid.
6. IMAC eluting solution: 12% ammonia and 10% HPLC grade acetonitrile.

**2.8 Hydrophilic Interaction Liquid Chromatography (HILIC)**

1. This protocol assumes access to high performance liquid chromatography (HPLC) equipment including an elution gradient programmer with data processing software, UV absorbance detector, and automated fraction collector.
2. 99.9% HPLC grade acetonitrile (*see Note 7*).
3. Column: 2.0  $\times$  150 mm TSK gel Amide-80 5  $\mu$ m particle.
4. Buffer D: 90% HPLC grade acetonitrile (*see Note 7*).
5. Buffer E: 80% HPLC grade acetonitrile with 0.005% trifluoroacetic acid (*see Note 7*).
6. Buffer F: 0.025% trifluoroacetic acid (*see Note 7*).

**2.9 Reverse Phase Liquid Chromatography–Tandem Mass Spectrometry**

1. The protocol assumes access to an online nano LC system interfaced with a mass spectrometer capable of performing tandem MS/MS. The present study was performed using an on-line Nano LC-Ultra<sup>®</sup> system (Eksigent) coupled with a Q Exactive mass spectrometer (Thermo Fisher Scientific).
2. Reverse phase analytical column: 125  $\mu$ m ID  $\times$  20 cm in-house packed with 3  $\mu$ m C<sub>18</sub> resin (*see Note 9*).
3. Buffer I: 0.1% LC-MS grade formic acid (*see Note 7*).
4. Buffer II: 0.1% LC-MS grade formic acid with 80% HPLC grade acetonitrile (*see Note 7*).

**2.10 Database Searching, Phosphopeptide Identification, Quantification, and QMAPS Generation**

1. For database search: SEQUEST software on a SORCERER system (Sage-N Research).
2. For quantification of relative protein abundances: XPRESS software, part of the Trans-Proteomic Pipeline (Seattle Proteome Center-SPC).

3. To increase the confidence of phosphopeptide identification, we performed a parallel search on Proteome Discoverer 1.4 software (Thermo Fisher Scientific) running SEQUEST and Percolator.
4. For phospho-site localization probabilities we used PhosphoRS within Proteome Discoverer (version 1.4.1.14, Thermo Fisher Scientific).

---

### 3 Methods

#### 3.1 *Stable Isotope Labeling of Amino Acids in Cell Culture (SILAC)*

The basic principle of SILAC for quantitative phosphoproteomic analysis consists of growing two cell cultures: one in a medium complemented with normal (“Light”) amino acids and the other in a medium complemented with stable-isotope labeled (“Heavy”) amino acids. When “Heavy” and “Light” samples are combined and analyzed by mass spectrometry, each phosphopeptide is detected as a pair in the mass spectra, with a predictable mass shift between the “Heavy” and “Light” forms of the phosphopeptide. In this case, because the “Light” and “Heavy” amino acids are chemically identical, the ratio of peak intensities in the mass spectrometer directly represents the ratio of protein phosphorylation abundance between the two cultures.

1. Inoculate fresh colonies from either strain MBS164 and MBS2042 (or MBS188) into 100 ml of SD “Light” or SD “Heavy” media, respectively. Then grow cultures for at least 12 h, at 30 °C with constant shaking. For **step 2** of the analysis, inoculate 100 ml of SD “Light” and SD “Heavy” media with fresh colonies from MBS164 (*see* Figs. 1 and 2 and **Note 10**).
2. On the next day the cultures should have an optical density at 600 nm ( $OD_{600}$ ) around 0.5. Dilute cultures in their respective SD media to an  $OD_{600} = 0.1$  in a final volume of 200 ml (*see* **Note 11**).

#### 3.2 *Cell Cycle Arrest, Drug Induced-DNA Damage, and Cell Harvesting*

Hydroxyurea (HU) is an inhibitor of ribonucleotide reductase. By depleting pools of dNTPs, HU stalls the progression of replication forks, leading to single-stranded DNA exposure and activation of DNA damage signaling [10]. Also, the DNA alkylating agent methyl methanesulfonate (MMS) can efficiently activate the DNA damage response during S-phase but not during G1 [11]. For the purpose of this protocol, we use 0.04% of MMS and 100 mM of HU to assess Mec1/Tel1 and Rad53-dependent signaling during S-phase.

1. When cells reach an  $OD_{600} = 0.2$  add alpha-factor to final concentration of 50 ng/ $\mu$ l for MBS164 and 50  $\mu$ g/ $\mu$ l for MBS2042/MBS188 and keep cultures growing for 3 h.

2. After 3 h, cell cultures should be around  $OD_{600} = 0.4$ . Centrifuge the cultures for  $2500 \times g$  for 3 min at RT. Discard supernatant and resuspend cells in prewarmed SD media containing pronase at 100  $\mu\text{g}/\text{ml}$  and HU or MMS to final concentration of 0.1 M and 0.04%, respectively. Keep cultures growing (*see Note 12*).
3. After drug treatment, normalize the amount of cells in each culture based on their respective  $OD_{600}$ .
4. Mix equal amounts of cells from both, light and heavy cultures and divide the final volume into two separated 500 ml conical centrifugation bottles.
5. Centrifuge the cultures for  $1000 \times g$  for 5 min at 4 °C. Discard supernatant.
6. Add 10 ml of ice-cold TE buffer to each centrifugation bottle, resuspend cell pellets and transfer each to a separate 15 ml Falcon tubes
7. Repeat **step 5**.
8. Resuspend each cell pellet in 4 ml of ice-cold TE buffer and divide the solution in 2 ml conical screw cap tube (total of  $4 \times 2$  ml conical screw cap tubes, each one containing 1 ml of cells).
9. Spin down cells for 1 min in a bench top microcentrifuge for  $10,000 \times g$  at room temperature (RT) and remove supernatant. Cell pellets can be stored at  $-80$  °C for up to 2 weeks.

### 3.3 Cell Lysis

1. Add 600  $\mu\text{l}$  of ice-cold glass beads to each tube.
2. Add 1 ml of lysis buffer to each tube and break cells at 4 °C for 30 min, with a 1 min pause at each 10 min interval, using the bead beater cell disruptor.
3. Transfer the lysate together with glass beads to a falcon tube and leave it on ice until glass beads decant.
4. Transfer the lysate to a 50 ml polycarbonate centrifuge tube and clear lysate by centrifugation at  $45,000 \times g$  at 4 °C for 30 min.
5. Collect supernatant in a new conical tube and determine the protein concentration by using protein assay dye reagent. Final protein concentration should be around 2 mg/ml with a total of 20 mg of protein.

### 3.4 Protein Precipitation and Tryptic Digestion

1. To denature and reduce disulfide bonds prior to alkylation, add SDS and DTT to a final concentration of 1% and 5 mM, respectively. Incubate samples at 60 °C for 10 min.
2. For alkylation of cysteines, add iodoacetamide to a final concentration of 25 mM and incubate samples for 15 min at RT.



Transfer the solution (approximately 12 ml) to a 200 ml conical bottle.

3. Add 36 ml of PPT solution. Mix it and keep on ice for 10 min (*see Note 13*).
4. Centrifuge sample for at  $4700 \times g$  for 10 min at RT.
5. Pour supernatant out and keep the bottles upside down for 5 min to drain the remaining PPT solution.
6. Add 4.5 ml of distilled water and 45  $\mu$ l of urea–Tris solution and resuspend protein pellet using a pipet-aid (*see Note 14*).
7. Transfer the solution to a 50 ml polycarbonate centrifuge tube and centrifuge at  $45,000 \times g$  for 4 min at RT.
8. Carefully remove supernatant with a tip connected to a vacuum line.
9. Wash the walls twice with 1 ml of distilled water without disturbing the pellet and repeat **step 8**.
10. Add 6 ml of urea–Tris solution and solubilize pellet by pipetting up and down.
11. Add 18 ml of Tris–NaCl solution (*see Note 15*).
12. Add 400  $\mu$ g of TPCCK treated trypsin (40  $\mu$ l of a 10 mg/ml stock) and incubate sample overnight at 37 °C under constant agitation.

### **3.5 Protein Sample Clean-Up**

1. On the next day, using a glass micro-syringes, add 500  $\mu$ l of 10% formic acid and 500  $\mu$ l of 10% TFA.
2. To remove particulates prior to sample cleanup, centrifuge tubes at  $2500 \times g$  for 5 min at RT.
3. By applying air pressure with a pipette bulb, condition 1 g of C<sub>18</sub> column by adding 2 ml of C<sub>18</sub> buffer D (*see Note 16*).
4. Add 4 ml of C<sub>18</sub> buffer A to equilibrate column.
5. Apply the sample through the column and let it flow by gravity.
6. Add 4 ml of C<sub>18</sub> buffer B to wash the column. Let it flow by gravity.
7. Repeat **step 6**.
8. Wipe residual volume of C<sub>18</sub> buffer C off column tip using a paper wipe.
9. By applying air pressure with a bulb, elute sample in 400  $\mu$ l of C<sub>18</sub> buffer C in a 2 ml silanized glass vial.
10. Using a gel loader tip, mix sample by pipetting up and down.
11. Dry sample completely at 45 °C using a speed-vac concentrator.
12. Add 400  $\mu$ l of 1% acetic acid and resuspend sample by pipetting up and down.

13. Transfer the vial to 2 ml microcentrifuge tube and spin it at  $5000 \times g$  for 1 min at room temperature to remove particulates prior to IMAC chromatography.

### 3.6 Preparation of Immobilized Metal Affinity Chromatography (IMAC) Resin

To increase coverage of the phosphoproteome we use Immobilized Metal Affinity Chromatography (IMAC) as a method to enrich for low abundance phosphopeptides prior to mass spectrometry analysis. The IMAC was performed using a homemade  $\text{Fe}^{3+}$  resin [8].

1. With the help of a paper clip, disassemble one Ni-NTA spin column and add resin (approximately 60 mg) to a 50 ml falcon tube containing 25 ml of stripping solution.
2. To strip Ni out from the resin, rotate for 40 min at RT on a nutator.
3. Spin resin for at  $1000 \times g$  for 1 min at RT.
4. Remove most of the solution using a glass pipette coupled to a vacuum line.
5. Add 25 ml of water.
6. Repeat **steps 3** and **4**.
7. Add 25 ml 0.6% acetic acid.
8. Repeat **steps 3** and **4**.
9. To bind iron to the resin, add 20 ml of iron solution and rotate for 40 min at RT.
10. Repeat **steps 3** and **4**.
11. Add 20 ml of 0.6% acetic acid.
12. Repeat **steps 3** and **4**.
13. Add 20 ml of 0.1% acetic acid.
14. Repeat **steps 3** and **4**.
15. Resuspend the resin in 1 ml of 0.1% acetic acid and transfer into a microcentrifuge tube. Wait until the resin is settled and then remove the supernatant leaving only 100  $\mu\text{l}$  (*see Note 18*).

### 3.7 Phosphopeptide Enrichment Using IMAC

1. Place glass wool fiber at the end of a loading tip and clip to stabilize.
2. Using a micropipette, add 100  $\mu\text{l}$  of IMAC slurry from Sub-heading 3.6, **step 15** (30  $\mu\text{l}$  of IMAC resin) to the loading tip.
3. Using a slip-tip disposable syringe, apply constant air pressure to the column and pack the resin (*see Note 16*).
4. Add trypsinized protein sample to the fresh packed IMAC resin. Run the sample through the resin by applying constant air pressure with a 1 ml slip-tip disposable syringe and collect the flow through as well (*see Note 17*).

5. Add 30  $\mu\text{l}$  of IMAC washing solution 1 and run through column.
6. Add 60  $\mu\text{l}$  of IMAC washing solution 2 and run through column.
7. Add 30  $\mu\text{l}$  of water and run through column.
8. Add 90  $\mu\text{l}$  of IMAC eluting solution and elute phosphopeptides into a 300  $\mu\text{l}$  silanized conical base insert vials.
9. Dry the sample completely at 45 °C using a speed-vac concentrator.
10. Add 15  $\mu\text{l}$  of water and mix by pipetting up and down.
11. Using a glass microsyringes, add 10  $\mu\text{l}$  of 10% formic acid.
12. Using a glass microsyringes, add 60  $\mu\text{l}$  of acetonitrile and mix well.
13. Spin it at  $5000 \times g$  for 1 min at room temperature to remove any precipitates prior to HILIC chromatography.

### **3.8 Hydrophilic Interaction Liquid Chromatography (HILIC)**

The unique separation ability of hydrophilic interaction liquid chromatography (HILIC) and its orthogonality toward other reverse phase separation techniques make it an ideal method for multidimensional fractionation of phosphopeptides prior to MS analysis [12]. Furthermore, because the HILIC mobile phase contains only acetonitrile, water and trifluoroacetic acid, fractions do not require an extra purification step for desalting. The present study was performed using a Dionex Ultimate 3000 HPLC controlled by Chromeleon 6.8 software. A TSK gel Amide-80 column (2 mm  $\times$  150 mm, 5  $\mu\text{m}$ ; Tosoh Bioscience) was used for the HILIC fractionation experiment. Three buffers were used for the gradient: buffer D (90% acetonitrile); buffer E (80% acetonitrile and 0.005% trifluoroacetic acid); and buffer F (0.025% trifluoroacetic acid).

1. Load samples into HILIC TSK gel Amide-80 column via a 100  $\mu\text{l}$  loop and set the fraction collector to time-based mode, collecting 1 min fractions between 10 and 22 min of the gradient in 300  $\mu\text{l}$  silanized conical base insert vials. The gradient used consists on a 100 % buffer D at time = 0 min, 88 % of buffer E and 12 % of buffer F at time = 5 min, 60 % of buffer E and 40 % of buffer F at time = 30 min and 5 % of buffer E and 95 % of buffer F from time = 35 to 40 min in a flow rate of 150  $\mu\text{l}/\text{min}$ .
2. Once the samples are fractionated, dry the sample fractions completely at 45 °C using a speed-vac concentrator.
3. Resuspend each fraction in 7  $\mu\text{l}$  of 0.1% trifluoroacetic acid.

### **3.9 Reverse Phase Liquid Chromatography- Tandem Mass Spectrometry**

As part of our multidimensional separation platform, we perform a  $C_{18}$  reverse phase fractionation prior to MS analysis.  $C_{18}$  reverse phase chromatography is a high resolution fractionation technique that presents a selectivity orthogonal to HILIC [12]. The present study was performed by using an on-line Nano LC-Ultra system coupled with a Q Exactive mass spectrometer.

1. Load sample fractions into a 20-cm column with 125  $\mu\text{m}$  inner diameter, packed in-house with 3  $\mu\text{m}$   $C_{18}$  particles. Reverse phase chromatography is performed with a binary buffer system consisting of buffer I (0.1% formic acid) and buffer II (0.1% formic acid with 80% acetonitrile). The peptides are separated by a linear gradient of buffer II up to 95% for 80 min with a flow rate of 200 nl/min.
2. Xcalibur 2.2 software (Thermo Fischer Scientific) was used for the data acquisition and the Q Exactive [13] was operated in the data-dependent mode. Survey scans were acquired in the Orbitrap mass analyzer over the range of 380–2000  $m/z$  with a resolution of 70,000. The maximum ion injection time for the survey scan was 80 ms with a 3e8 automatic gain-control target. MS/MS spectra was performed selecting up to the ten most abundant precursor ions with a charge state  $\geq 2$  and  $< 5$  within an isolation window of 2.0  $m/z$ . Selected precursor ions were fragmented by higher energy collisional dissociation with normalized collision energy of 27 and analyzed in the Orbitrap mass analyzer with a mass resolution of 17,500.

### **3.10 Database Searching, Phosphopeptide Identification, Quantification, and QMAPS Generation**

Raw MS/MS spectra were searched in a SORCERER (Sage N Research, Inc.) system using SEQUEST software and a composite yeast protein database, consisting of both the normal yeast protein sequences and their reversed protein sequences as a decoy to estimate the false discovery rate (FDR) in the search results [14]. To increase the confidence of phosphopeptide identification, we performed a parallel search on Proteome Discoverer 1.4 software (Thermo Fisher Scientific) running SEQUEST and Percolator. Only phosphopeptides identified with high confidence in both SORCERER and Proteome discoverer/percolator searches were considered, which resulted in extremely low false discovery rates (FDR < 0.02%) [15].

1. The following parameters were used in the database search: Semitryptic requirement, a mass accuracy of 15 ppm for the precursor ions, differential modification of 8.0142 Da for lysine, 10.00827 Da for arginine, 79.966331 Da for phosphorylation of serine, threonine and tyrosine and a static mass modification of 57.021465 Da for alkylated cysteine residues.

2. XPRESS software, part of the Trans-Proteomic Pipeline (Seattle Proteome Center), was used to quantify all the identified peptides [16].
3. The phosphorylation localization probabilities were determined by using PhosphoRS within Proteome Discoverer (version 1.4.1.14, Thermo Fisher Scientific) [17].
4. In the first step of our analysis we established the following criteria to assign Mec1/Tel1-dependent phosphopeptides: (1) At least sixfold increase in phosphopeptide abundance in WT relative to *mec1Δtel1Δ* cells. (2) For Rad53-dependent phosphopeptides, we established a threshold of twofold increase in abundance in WT relative to *rad53Δ* cells. (3) Moreover, to be considered a Mec1/Tel1-dependent substrate, each phosphopeptide must be Mec1/Tel1-dependent but not Rad53-dependent and to be considered a Rad53-dependent substrate, each phosphopeptide has to be Mec1/Tel1-dependent and Rad53-dependent substrate. (4) Finally, we considered only phosphopeptides identified in at least two biological replicates.
5. In the QMAPS, we depict the fold change in normalized phosphopeptide abundance between cells treated with MMS or HU. Ratios of phosphopeptide abundance were converted to fold change by using MATLAB software (MathWorks®) as part of a custom-designed web tool that allowed upload of data files from SORCERER. For the QMAPS generation, results of the phosphoproteome analysis were filtered using the list of preassigned kinase-dependent phosphopeptides identified in step 1 (see Figs. 1 and 2). In the QMAPS plot, each dot corresponds to the average fold change of one Mec1/Tel1 and/or Rad53-dependent phosphopeptide identified at least three times in two independent biological replicates [5]. Phosphopeptides deviated to the left or right side of the plot indicate, respectively, a decrease or increase in their phosphorylation levels between cells treated with MMS or HU (see Figs. 2 and 3). To assess the probability of having each phosphopeptide with a ratio significantly distinct from the ratios of the total set of phosphopeptides identified in the analysis, we applied the Mann–Whitney–Wilcoxon Test (U-test). Finally, all plotted phosphopeptides were manually inspected for phosphor-site assignment and quantitation.

---

## 4 Notes

1. Strains MBS164, MBS2042 and MBS188 are SILAC strains auxotrophic for lysine and arginine. These strains carry a deletion of the ribonucleotide reductase inhibitor *SML1* that suppresses the lethality of *mec1Δ* or *rad53Δ* allele [18].

2. Proline is added to prevent the metabolic conversion of heavy arginine to heavy proline.
3. Add PMSF, EDTA-free protease inhibitor cocktail and phosphatase inhibitor to cell lysis buffer right before proceeding to cell lysis.
4. Prepare a fresh alkylation solution for every experiment.
5. Prepare a fresh urea-Tris solution for every experiment.
6. After the trypsin is diluted in 0.1% acetic acid, keep small aliquots frozen at  $-80^{\circ}\text{C}$ .
7. To avoid sample contamination with polymers always use a glass syringe/glass pipette/glass graduated cylinder to dilute strong acids and high concentrated organic solvents.
8. Prepare a 1 M stock  $\text{FeCl}_3$  solution in 0.3% acetic acid and let it decant for 2 weeks prior to use. Prepare the iron chloride solution by adding 2 ml of 1 M stock solution to 18 ml of 0.3% of acetic acid.
9. The analytical column was generated by pulling capillary to 5  $\mu\text{m}$ -ID tip. Reverse-phase particles were packed directly into the pulled column at 1000 psi until 20 cm long. The column was further packed, washed, and equilibrated at 1000 psi in buffer II followed by buffer I.
10. Use fresh yeast culture colonies.
11. Avoid culture saturation in SD media ( $\text{OD}_{600} > 0.5$ ). Once saturated, SILAC strains will not grow well. If culture is saturated, redilute and let it grow for at least four generations before proceeding with the experiment.
12. Incubation time should be tested empirically depending on the kinase mutant. For the WT and *rad53* $\Delta$  cells, 45 min incubation should be enough to allow cell cycle progression through S-phase. For *mec1* $\Delta$ *tel1* $\Delta$  double mutant 3–4 h incubation may be necessary for proper release from G1 and progression through S phase. To check cell cycle progression, FACS analysis should be performed at different time points.
13. Proteins will precipitate and solution will get cloudy.
14. This is a critical step to remove hydrophilic contaminants that interfere with HILIC.
15. This step is required to dilute urea prior to protein digestion with trypsin.
16. Avoid letting the column dry in any step. Keep resin moist by leaving a fine layer of liquid over it.
17. Apply constant and slow air pressure in all steps during phosphopeptide enrichment using IMAC.

## Acknowledgments

We thank Beatriz S. Almeida for technical support. M.B.S. is supported by grants from the National Institutes of Health (R01-GM097272), F.M.B.d.O. is supported by grants from FAPERJ No E-26/010.002831/2014 and No E-26/010.003001/2014 and from CNPq No 446143/2014 and D.K. is supported by Cornell Vertebrate Genomic Scholarship.

## References

- Awasthi P, Foiani M, Kumar A (2015) ATM and ATR signaling at a glance. *J Cell Sci* 128 (23):4255–4262
- Cimprich KA et al (1996) cDNA cloning and gene mapping of a candidate human cell cycle checkpoint protein. *Proc Natl Acad Sci U S A* 93(7):2850–2855
- Sanchez Y et al (1999) Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms. *Science* 286(5442):1166–1171
- Chen SH et al (2010) A proteome-wide analysis of kinase-substrate network in the DNA damage response. *J Biol Chem* 285 (17):12803–12812
- Bastos de Oliveira FM et al (2015) Phosphoproteomics reveals distinct modes of Mec1/ATR signaling during DNA replication. *Mol Cell* 57(6):1124–1132
- Hustedt N et al (2015) Yeast PP4 interacts with ATR homolog Ddc2-Mec1 and regulates checkpoint signaling. *Mol Cell* 57:273
- Zhou C et al (2016) Profiling DNA damage-induced phosphorylation in budding yeast reveals diverse signaling networks. *Proc Natl Acad Sci U S A* 113(26):E3667–E3675
- Albuquerque CP et al (2008) A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Mol Cell Proteomics* 7(7):1389–1396
- Ong SE et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1(5):376–386
- Alvino GM et al (2007) Replication in hydroxyurea: it's a matter of time. *Mol Cell Biol* 27(18):6396–6406
- Tercero JA, Longhese MP, Diffley JF (2003) A central role for DNA replication forks in checkpoint activation and response. *Mol Cell* 11 (5):1323–1336
- Horvatovich P et al (2010) Multidimensional chromatography coupled to mass spectrometry in analysing complex proteomics samples. *J Sep Sci* 33(10):1421–1437
- Michalski A et al (2011) Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* 10 (9):M111.011015
- Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5 (11):976–989
- Kall L et al (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* 4(11):923–925
- Han DK et al (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol* 19(10):946–951
- Taus T et al (2011) Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res* 10(12):5354–5362
- Zhao X, Muller EG, Rothstein R (1998) A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Mol Cell* 2(3):329–340

# INDEX

## A

- Adaptor ligation .....64, 217, 252
- A-like faker (ALF) ..... 1–8
- Anaphase bridges ..... 481–491, 495–507
- Atomic force microscopy (AFM) ..... 557–571

## B

- Base excision repair .....63, 64
- Bromodeoxyuridine (BrdU)..... 209–224
  - immunoprecipitation ..... 214, 219, 220, 229–231

## C

- Canavanine ..... 26, 28, 29, 36, 38–41, 44, 46, 576–579, 584, 585, 616, 617, 618, 620, 621, 624
- Cell labeling..... 156, 158
- Checkpoint ..... 35, 143, 156, 178, 180, 181, 183, 211, 485, 486, 537, 575, 599, 647, 649
- Chromatin
  - fibers..... 509–525
  - preparation ..... 83, 85–89, 98
- Chromatin immunoprecipitation (ChIP) ..... 80, 83–86, 90, 92, 93, 229–231, 631
- Chromosome instability (CIN)..... 16, 478
- Chromosome segregation ..... 11
- Chromosome transmission fidelity (CTF)..... 11–18
- Circular ligation ..... 72–73
- Comet assay..... 63, 312–314
- Confocal microscopy ..... 528
- Copy number variation (CNV)..... 196, 627
- Coulter counter..... 552
- Crosslinking
  - formaldehyde..... 86, 98
  - psoralen..... 262–266, 283, 286, 287

## D

- Data analysis
  - atomic force microscopy images ..... 557, 558
  - ChIP-seq..... 631, 633
  - colony fluorescence ..... 621–622
  - colony size ..... 605
  - DNA fibers ..... 196, 520–524
  - mass spectrometry..... 651, 657
  - microarray..... 91, 93
  - normalization ..... 94, 95

- sequencing ..... 64, 162–163
- Data visualization
  - ChIP-seq..... 638
  - replication profiles..... 196, 204
- Deep sequencing ..... 195–206
- Deletion ..... 1–4, 8, 13, 44, 45, 57, 78, 133, 143, 167, 247, 301, 302, 325, 329, 330, 404, 422–424, 430, 431, 436, 575, 576, 580, 581, 585, 592, 594, 618–620, 624, 6614–616
- DNA
  - adducts..... 107
  - barcoding..... 213–214, 217–218
  - combing..... 149
  - damage..... 35, 77–99, 101, 102, 149, 155, 156, 167, 195, 282, 312, 320, 477, 480, 527–539, 546–549, 615, 645–659
  - polymerase ..... 47, 52, 53, 66, 68–72, 74, 80, 92, 93, 157, 158, 162, 214, 215, 218, 221, 227, 228, 230, 235, 240, 248, 250, 300, 319, 320, 322, 324, 326, 330, 331, 334, 337, 473
  - repair ..... 78–80, 101–105, 375, 376, 381–383, 475, 478, 529, 537
  - replication ..... 11, 102, 195–206, 209, 211, 227–237, 245, 261–293, 295–308, 329, 347, 439, 472, 474, 477, 496, 510, 518, 520
- DNA purification
  - genomic ..... 52, 330–332
  - high molecular weight ..... 159, 160
  - Xenopus* egg extracts ..... 267, 272
- DNA purification (extraction, isolation) from human cells..... 52, 82
- DNA purification (extraction, isolation) from yeast..... 52, 82
- DNA-RNA hybrids ..... 347–358
- DNA-RNA Immunoprecipitation..... 349–350, 352–356
- Double-strand breaks (DSBs) ..... 65, 119–128, 131–144, 147, 148, 155–159, 162, 163, 165, 167–193, 364, 375–384, 404, 424, 529, 547
  - mapping and localization ..... 170
- Double-stranded DNA (dsDNA) ..... 33–36, 38, 48, 55, 120, 121, 138, 143, 169, 170, 174, 197, 200, 243, 251, 252, 258, 262, 282, 288, 291, 292, 337, 368



**E**

*E. coli* ..... 64, 67–69, 107,  
108, 110, 158, 160, 231, 234, 312, 315, 316,  
390, 395, 397, 615

Electron microscopy ..... 261–292

Electrophoresis  
alkaline ..... 135, 139, 144, 247, 249  
native ..... 367–368, 370  
single cell ..... 311–317  
two dimensional ..... 298

Electroporation ..... 109, 111, 116, 118

End resection ..... 121, 131, 132, 140, 148, 172

**F**

Fiber spreads ..... 509, 517

Fitness ..... 575–597

Fixation ..... 171, 172, 174, 176, 184,  
394–395, 400, 496–500, 502–507

Flow cytometry ..... 196, 198,  
200, 203–205, 224, 302, 307, 514, 614, 617,  
620, 624–626, 629

Fluctuation test ..... 51, 59, 421, 422,  
424, 428, 434

Fluorescence in situ hybridization (FISH) ..... 387,  
388, 389, 391, 396, 397, 399, 474, 476, 510

Fluorescence scanning ..... 617

Fluorescent in situ hybridization ..... 387–401

Fluorescent protein ..... 485, 486, 491, 528,  
537, 546, 547, 613–629

Formaldehyde crosslinking ..... 86, 98

Fragile site  
common ..... 405, 472, 473, 475, 476, 484, 496  
rare ..... 472, 473, 478, 480

Fragility ..... 403–437, 472–474, 477, 478, 480

**G**

Gene conversion ..... 2, 3, 133

Gross chromosomal rearrangement ..... 43

Growth curve ..... 449, 466, 576, 589, 592

**H**

Hemocytometer ..... 28, 37, 38, 85, 149

High throughput sequencing ..... 163, 197,  
200, 203, 209–224

HO endonuclease ..... 119, 122,  
133, 364–366

Homologous recombination (HR) ..... 2, 117, 119,  
131, 147, 172, 186, 193, 240, 376–378, 444,  
445, 615

Hybridization  
microarray ..... 92–93  
Southern ..... 249–250

Hydroxyurea (HU) ..... 211, 212, 215, 224,  
229, 232, 237, 289, 486, 487, 489, 511, 514,  
524, 547–549, 647–649, 650, 652, 658

**I**

Image analysis ..... 163, 293, 513,  
525, 580, 581, 592, 597, 617, 628

Immunofluorescence ..... 103, 104, 148,  
154, 209, 351, 356–358, 388, 399, 400, 495,  
496, 510, 512–513, 516–517

Immunoprecipitation (IP) ..... 77, 81, 82,  
89, 90–93, 99, 148, 196, 209–224, 227–231,  
349–350, 352–356, 631, 635

**L**

Libraries for sequencing ..... 55, 74

**M**

Mapping sequencing reads ..... 55

Marker loss ..... 8, 404

Mass spectrometry (MS) ..... 645–660

Microarrays ..... 45, 77–99, 196, 228, 628

Microfluidics ..... 537–554

Microsatellite instability ..... 422, 439

Mutagenesis ..... 33–41, 107

Mutation rate ..... 8, 21–30, 195, 418,  
422, 424–426, 432–436, 449

**N**

Next-generation sequencing ..... 155–165,  
234–236, 240, 241, 243, 244

Non-homologous end joining ..... 147, 167

Nucleotide excision repair (NER) ..... 78, 101,  
102, 104, 311

**P**

PCR  
colony ..... 381, 383, 404,  
405, 410–413, 416, 417, 426–427, 446,  
450–452, 454–459, 461, 467, 468  
quantitative real-time ..... 81, 91–92

Peak calling ..... 193, 223, 632, 634–635

Photo-toxicity ..... 529

Protein  
phosphorylation ..... 652  
recruitment ..... 102

Protein-protein interactions ..... 599

Pulsed-field gel electrophoresis ..... 45

**R**

Recombination  
foci ..... 529

homology-directed ..... 119  
 site-specific ..... 111  
 Repeat contractions ..... 404, 407  
 Repeat expansions ..... 404, 423  
 Replication  
   fork barrier ..... 295  
   intermediates ..... 261, 263, 267–268,  
     275–276, 279, 283, 285, 286, 305, 477  
   origins ..... 111, 195, 196, 209, 228, 230,  
     234–237, 240, 299, 341–343, 423, 475  
   stress ..... 168, 169, 192, 193, 262, 329,  
     480, 489, 496  
 Reporter gene ..... 33–36, 38, 298, 376, 422,  
   444, 445, 613, 614, 619  
 Ribonucleic acid (RNA) localization ..... 387, 388  
 Ribonucleotide  
   excision repair ..... 311  
   incorporation ..... 239–258, 329, 330  
 R loop ..... 347–359  
 Robotics ..... 580–584, 585, 587,  
   588, 602, 603, 618

**S**

*Saccharomyces cerevisiae* ..... 1, 11, 43–60, 64,  
   73, 78, 119–121, 126, 142, 196, 200, 211, 222,  
   239, 248, 264–266, 270, 273–274, 275, 289,  
   339, 341, 363, 365, 369, 370, 377–379, 381,  
   387, 403–437, 439–468, 485, 528, 540, 600,  
   613–629, 637, 645, 646  
*Schizosaccharomyces pombe* ..... 239, 240, 242,  
   248, 255, 257, 445, 509, 511, 540  
 Single molecule analysis ..... 147–154, 561  
 Single-strand breaks (SSBs) ..... 156, 157,  
   165, 227, 312  
 Single-strand DNA (ssDNA) ..... 34–36, 41,  
   119–121, 124, 126, 127, 131, 132, 138, 140,  
   142–144, 147–149, 156, 228, 229, 234–236,  
   243, 247, 251, 257, 258, 262–264, 281–283,  
   286, 288–290, 293, 390, 391, 397  
 Sonication of DNA ..... 188

Southern blot ..... 120, 121, 124, 125,  
   131–154, 239–242, 246–247, 249, 304, 305,  
   363–373, 444–446, 460–464  
 Spheroplasting ..... 134, 136, 137, 142,  
   266, 273, 349, 354, 366, 369, 371, 394, 395,  
   400, 515  
 Synchronization ..... 196, 198–200,  
   211–212, 215–216, 224, 232, 270, 271, 289,  
   298, 302, 490, 503  
 Synthetic genetic array (SGA) ..... 539–597, 576, 577,  
   580–583, 585–588, 593, 613–629

**T**

Telomerase ..... 363, 364, 372, 375, 376,  
   387–401, 406, 474, 480  
 Telomere addition ..... 44, 57, 363–373,  
   406, 407  
 Time-lapse imaging ..... 537–554  
 Transformation  
   *E. coli* ..... 615  
   yeast ..... 297, 298, 300–301  
 Translesion DNA synthesis (TLS) ..... 34, 35,  
   107, 111, 116  
 Trinucleotide repeats (TNR) ..... 404, 439–469

**U**

Ultrafine bridges (UFBs) ..... 477, 483–491,  
   495–497, 503–507  
 Uracil incorporation ..... 63  
 UV damage ..... 64, 71, 101, 103, 104  
 UV photodimers ..... 65, 74

**W**

Whole chromosomal loss ..... 3  
 Whole-genome sequencing ..... 45, 169, 170, 184

**Y**

Yeast artificial chromosome (YAC) ... 405–407, 409, 410,  
   413, 415, 416–418