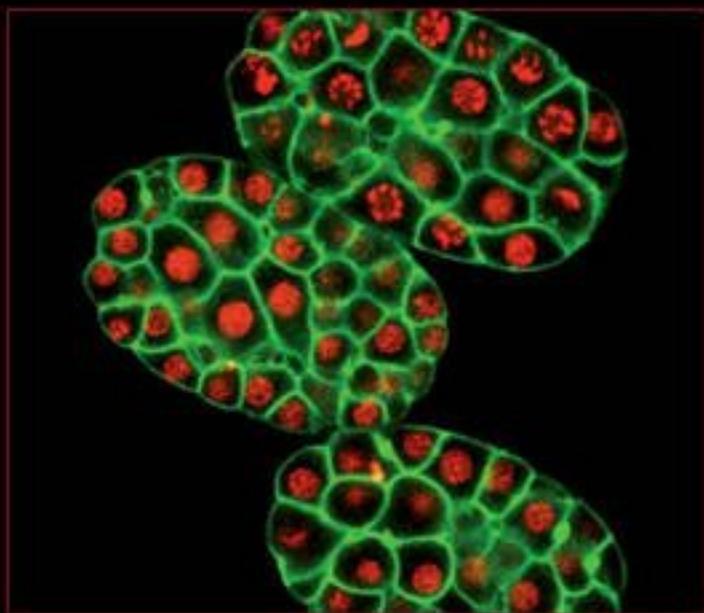


---

Methods in Cell Biology • Volume 106

---

# CAENORHABDITIS ELEGANS: MOLECULAR GENETICS AND DEVELOPMENT



Edited by

H. William Detrich III, Monte Westerfield  
and Leonard I. Zon



# Methods in Cell Biology

**VOLUME 106**

*Caenorhabditis elegans*: Molecular Genetics and Development  
2nd Edition

## **Series Editors**

**Leslie Wilson**

Department of Molecular, Cellular and Developmental Biology  
University of California  
Santa Barbara, California

**Paul Matsudaira**

Department of Biological Sciences  
National University of Singapore  
Singapore

---

---

# Methods in Cell Biology

---

## VOLUME 106

*Caenorhabditis elegans*: Molecular Genetics and Development  
2nd Edition

Edited by

**Joel H. Rothman**

Professor and Chair, Department of MCD Biology,  
University of California, Santa Barbara, CA, USA

**Andrew Singson**

Associate Professor, Rutgers University,  
Waksman Institute, Piscataway, NJ, USA



AMSTERDAM • BOSTON • HEIDELBERG • LONDON  
NEW YORK • OXFORD • PARIS • SAN DIEGO  
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier  
225 Wyman Street, Waltham, MA 02451, USA  
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA  
32, Jamestown Road, London NW1 7BY, UK  
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

Second edition 2011

Copyright © 2011 Elsevier Inc. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: [permissions@elsevier.com](mailto:permissions@elsevier.com). Alternatively you can submit your request online by visiting the Elsevier web site at <http://elsevier.com/locate/permissions>, and selecting *Obtaining permission to use Elsevier material*

#### **Notice**

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

ISBN: 978-0-12-544172-8  
ISSN: 0091-679X

For information on all Academic Press publications  
visit our website at [elsevierdirect.com](http://elsevierdirect.com)

Printed and bound in USA

11 12 13 14 10 9 8 7 6 5 4 3 2 1

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation

---

---

## CONTRIBUTORS

*Numbers in parentheses indicate the pages on which the author's contributions begin.*

**Shirin Bahmanyar** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA

**Jean-Louis Bessereau** (65), Ecole Normale Supérieure, Institut de Biologie de l'ENS, IBENS, Paris, France; INSERM U1024, Paris, France; CNRS, UMR 8197, Paris, France

**Thomas Blumenthal** (187), Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado, USA

**Gina Broitman-Maduro** (253), Department of Biology, University of California at Riverside, Riverside, California, USA

**Ana Carvalho** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA

**Shih-Peng Chan** (219), Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA; Current address: Department of Microbiology, College of Medicine, National Taiwan University, Taipei, Taiwan

**Iain Cheeseman** (289), Whitehead Institute for Biomedical Research, MIT, Cambridge, MA, USA

**Andrew D. Chisholm** (325), Section of Cell and Developmental Biology, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA

**Patricia G. Cipriani** (89), Center for Genomics and Systems Biology; Department of Biology, New York University, New York, NY 10003

**Arshad Desai** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA

**Julien Dumont** (289), Institut Curie-UMR144, Paris, France

**Brian D. Geldziler** (343), Waksman Institute, Rutgers University, Department of Microbiology and Molecular Genetics, Piscataway, New Jersey, USA

**Claudiu A. Giurumescu** (325), Section of Cell and Developmental Biology, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA

- Reto Gassmann** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA
- Jeff Hardin** (377), Department of Zoology, University of Wisconsin-Madison, Madison, Wisconsin, USA
- Pan-Young Jeong** (445), Present address: Department of Molecular, Cellular and Developmental Biology, University of California Santa Barbara, Santa Barbara, California, USA
- Yishi Jin** (413), Howard Hughes Medical Institute, Neurobiology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA
- Martin R. Jones** (23), Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
- Hyoe-Jin Joo** (445), Yonsei Proteome Research Center, Department of Biochemistry and Biomedical Science, College of Life Science and Biotechnology, World Class University Program, Graduate School, Yonsei University, Seoul, Korea
- Heekyeong Kim** (445), Yonsei Proteome Research Center, Department of Biochemistry and Biomedical Science, College of Life Science and Biotechnology, World Class University Program, Graduate School, Yonsei University, Seoul, Korea
- Kwang-Youl Kim** (445), Yonsei Proteome Research Center, Department of Biochemistry and Biomedical Science, College of Life Science and Biotechnology, World Class University Program, Graduate School, Yonsei University, Seoul, Korea
- Eric J. Lambie** (3), Biology Department II, Ludwig-Maximilians-University, Munich, Germany
- Jeeyong Lee** (445), Yonsei Proteome Research Center, Department of Biochemistry and Biomedical Science, College of Life Science and Biotechnology, World Class University Program, Graduate School, Yonsei University, Seoul, Korea
- Zoe Lohn** (23), Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
- Paul Maddox** (289), Institute for Research in Immunology and Cancer, Department of Pathology and Cell Biology, University of Montreal, Montreal, Quebec, Canada
- Morris F. Maduro** (161, 253), Department of Biology, University of California Riverside, Riverside, California, USA
- Matthew R. Marcello** (343), Waksman Institute, Rutgers University, Department of Microbiology and Molecular Genetics, Piscataway, New Jersey, USA; UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey, USA
- J. Jason Morton** (187), Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado, USA
- Sherry Niessen** (289), The Skaggs Institute for Chemical Biology and Department of Chemical Physiology, The Center for Physiological Proteomics, The Scripps Research Institute, La Jolla, California, USA

- Karen Oegema** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA
- Young-Ki Paik** (445), Yonsei Proteome Research Center, Department of Biochemistry and Biomedical Science, College of Life Science and Biotechnology, World Class University Program, Graduate School, Yonsei University, Seoul, Korea; Rm#423, Industry-University Research Bldg., Yonsei University, 134, Shinchon-dong, Sudaemoon-ku, Seoul, Korea
- Amy E. Pasquinelli** (219), Department of Biology, University of California San Diego, La Jolla, California, USA
- Fabio Piano** (89), Center for Genomics and Systems Biology; Department of Biology, New York University, New York, NY 10003; New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
- Vida Praitis** (161), Biology Department, Grinnell College, Grinnell, Iowa, USA
- Valérie J.P. Robert** (65), Ecole Normale Supérieure, Institut de Biologie de l'ENS, IBENS, Paris, France; INSERM U1024, Paris, France; CNRS, UMR 8197, Paris, France; Present address: Laboratory of Molecular and Cellular Biology, UMR5239 CNRS, Ecole Normale Supérieure de Lyon, Lyon, France
- Ann M. Rose** (23), Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
- Diane C. Shakes** (343), College of William and Mary, Department of Biology, Williamsburg, Virginia, USA
- Andrew Singson** (343), Waksman Institute, Rutgers University, Department of Microbiology and Molecular Genetics, Piscataway, New Jersey, USA
- David R. Sherwood** (113), Department of Biology, Duke University, Durham, North Carolina, USA
- Frank J. Slack** (219), Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA
- Albertha J.M. Walhout** (271), Program in Gene Function and Expression and Program in Molecular Medicine, University of Massachusetts Medical School, USA
- Zheng Wang** (113), Department of Biology, Duke University, Durham, North Carolina, USA
- Priscilla M. Van Wynsberghe** (219), Department of Biology, University of California San Diego, La Jolla, California, USA; Current address: Department of Biology, Colgate University, Hamilton, NY, USA
- Dong Yan** (413), Howard Hughes Medical Institute, Neurobiology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA
- John R. Yates III** (289), The Skaggs Institute for Chemical Biology and Department of Chemical Physiology, The Center for Physiological Proteomics, The Scripps Research Institute, La Jolla, California, USA
- Esther Zanin** (289), Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA

---

---

## PREFACE

### ***Caenorhabditis elegans*: Molecular Genetics and Development**

The allure of a model organism comes not from any special fascination for the creature itself; it is doubtful that most researchers studying a simple and tiny animal, the nematode *Caenorhabditis elegans*, are particularly attracted to these modest creatures *per se*. Rather than any fondness for the animal, it is the exceptional experimental methodology available with these high-performance vehicles of biological discovery that entice those driven by intellectual curiosity about the living world to investigate their inner workings. As this millimeter-long creature has amply proven itself to be of enduring utility for biological discovery, it is of value to continue to assemble and update experimental methodology on its use.

A century, in fact a millennium, has turned since the last *C. elegans* volume was published in the Methods in Cell Biology series (as volume 48). That volume, “*Caenorhabditis elegans*: Modern Biological Analysis of an Organism,” was the first major compendium of *C. elegans* methods and only the second complete published volume on this creature. Since that volume appeared, several other collections of methods have been published, notably a brief practical volume edited by I. Hope, a methods section edited by V. Ambros as a component of the online resource, Wormbook, and a published volume edited by K. Strange. Nonetheless, for over a decade and half the original *C. elegans* methods volume has served as an invaluable resource to both seasoned and new researchers who focus their scientific curiosity on *C. elegans*.

One might reasonably ask, in an era in which printed material is rapidly dissipating into cyberspace and vast information resources are available online, why a printed volume is of value. In our view, the accessibility of a printed form is still well-suited to the laboratory environment. Several copies of the 1995 *C. elegans* methods volume are stationed at ready access on our laboratory shelves. Members of our laboratories continually reach for the book, even many years after its publication. Just as laboratory notebooks have yet to be satisfactorily replaced by a digital medium, the ability to flip through the pages of a methods volume is not yet an anachronism in the setting of an experimental laboratory.

We are closing in on the 50th anniversary of Sydney Brenner’s 1963 letter to the director of the Medical Research Council Laboratory of Molecular Biology, immortalized in the first Cold Spring Harbor Press *C. elegans* monograph, in which he proposed to adopt a *Caenorhabditis* worm as a model organism. The period since the predecessor of this volume was published in 1995 is a fraction of that interval and yet has seen the majority of the prominent discoveries made with the animal. Six Nobel laureates in the *C. elegans* community have been celebrated since the last volume in

this series was published and key discoveries that led to some of those prizes were made during this period. This era also saw the first complete animal genome sequence, the discovery of RNAi, the generality of miRNA-mediated control across biology, and many other fundamental advances that have emerged from the laboratories of *C. elegans* investigators and have proven to be broadly transformative to biology. While advances made with the worm are quantifiable, their full impact on science is inestimable.

The original volume made an effort to be fully comprehensive. In earlier days, it was possible to craft a single volume that fully addressed the state of the art. But the precipitous growth in the field would make such an undertaking unwieldy, if not impractical. 85% of all publications in the Pubmed database containing the text word “*elegans*” (currently approximately 20,000) were published since the last volume appeared, many by labs that have not traditionally focused on the animal. The number of *C. elegans* laboratories with strain designations (now over 850) is more than five times that in 1995. A compendium of methods that scaled similarly might contain over 170 chapters spread over as many as 10 volumes. Thus, the two volumes assembled here no longer attempt to serve as a single source book for *C. elegans* methodology. Rather, we have chosen to include chapters on many of the methods that have evolved or dramatically altered since the 1995 volume, with the recognition that more has been left out than included. It is inevitable that additional volumes will come in the future that fill and update voids left by the current collection.

Owing to the length of this updated collection, it is now distributed over two volumes. This volume (vol. 106 in the series) comprises genetics, molecular biology, and development, while the subsequent volume (vol. 107) will focus on imaging, cell biology, and physiology. Many methods from the original 1995 volume (e.g., basic culturing, mutagenesis, mosaic analysis, and so on) are still relevant and useful and the experimentalist is encouraged to consult that volume for such methods. It is inevitable that some of the methods in the earlier volume (e.g., the physical map, genomic and cDNA sequencing, and use of the extinct database structures that preceded Wormbase) have become obsolete. On the other hand, many methods have been improved or refined for specific applications, for example, genetic mapping techniques (Chapter 1), reverse genetic approaches (Chapters 3 and 4), transgenesis (Chapter 6), and *in situ* hybridization using RNA probes (Chapter 9), all of which are covered in this volume as revised or entirely new chapters. We note that, unlike the previous edition, we have not included comprehensive appendices, as this information is now readily available in a continually updated manner online through the internet resources listed in the single Appendix of this volume.

Mastery of the varied tools of *C. elegans* biology is enhanced by the experience gained in a lab connected to those that grew up during formative stages of the field. The lore, philosophy, and strategies one uses to dissect biological processes are not coherently incorporated in the literature, but can be effectively transmitted through a sort of apprenticeship in such labs. The worm field is famed for the large fraction of practitioners who trace lineal roots to the early pioneers in the field. However, over recent years, the prominence of the worm system has lured many researchers not

formally linked to the “worm pedigree” to adopt the animal as a useful tool for their favorite subjects of inquiry. Thus, rather than covering discrete methods *per se*, some of the chapters are designed to transmit strategies that are not easily gleaned from the literature (most prominently featured in Chapter 5, which describes genetic strategies used to deconstruct the pathways that drive cellular and developmental processes, and also in the chapters on mapping and on specialized chromosomes in Chapters 1 and 2.) We believe that these strategies will be of particular value to newcomers who learn worm biology without the benefits of apprenticeship in a seasoned worm lab.

Among the most notable of the advances in *C. elegans* technology since the first volume was published was the discovery of RNAi and subsequent methods for adapting RNAi to broad functional genomics screens, which have revolutionized discovery of gene function. Such approaches, and the integration of the “phenome” with informatics studies of functional relationships between gene activities, are covered in chapter 4, of value to aficionados and newcomers alike. Similarly, the recognition that miRNAs function at many levels across animal biology make chapter 8, on analysis of miRNAs, an essential component of this volume. In addition, since publication of the earlier volume, it has become clear that a large fraction of worms genes are organized in operons and are trans-spliced. Any worm molecular geneticist must be mindful of this complexity of gene organization in the animal and the methods for analyzing RNA processing (Chapter 7) are therefore important to any researcher considering the structure of genes and effects of mutations and RNAi on gene expression.

An overarching goal articulated by Sydney Brenner when he inaugurated *C. elegans* research was to obtain a complete description of the animal: this began with the comprehensive analysis of the cell lineage and anatomy and later the whole genome. More recently, this goal has been extended to the level of gene function and interaction by techniques covered in chapters that describe functional and transcriptional network analysis (Chapters 4 and 10). Genetic approaches have dominated *C. elegans* research; however, biochemical methods have become increasingly more significant, particularly as the pathway from *in vitro* discovery to *in vivo* validation has shortened, and methods for analyzing protein complexes and other proteomics approaches are covered in Chapters 11 and 16. The pre-eminent focus of Brenner’s original vision to exploit *C. elegans* as a new model system was directed at unveiling the processes that drive development, the biological challenge that drew many researchers to the worm. In keeping with the predominance of this discipline, a major subdivision of this volume comprises five chapters (Chapters 12–16) that address varied approaches to problems in developmental biology, ranging from cell lineage analysis (including new advances in automated lineage analysis), fertilization, morphogenesis, nervous system development and regulation of the alternative developmental stage, the dauer larva.

Just as *C. elegans* develops rapidly, so do technological approaches to analyzing its biology. It is clear that we are able here to capture only an instant in this rapidly moving field, and methods have advanced even during the period in which these

chapters were being assembled and edited. For example, the tremendous advances in DNA sequencing technology is making whole-genome sequence identification of mutations inexpensive and routine, thereby superseding much of the traditional genetic mapping approaches. Moreover, effective new methods for generating genomic modifications based on synthetic nucleases have recently appeared, but came too late to include in the initial release of this volume. These and untold other technologies will no doubt occupy the pages of future editions in this series, devoted to this magnificent living tool for biological discovery.

Appreciation for the richness of technology available to *C. elegans* researchers, only partially captured in the current volumes, has been expressed in many ways, even beyond scientific activity. Two of the traditions at the biennial International *C. elegans* meeting are the Worm Show, an evening comedy variety show, and the Worm Art show, in which artistic members of the worm community pay homage to the animal through visual arts, films, and crafts, including clothing and even cuddly stuffed toys. Upon further reflection, perhaps those of us who have dedicated so many years to pursuing the wonderful mysteries of *C. elegans*, and appreciative of the many gifts that it has generously yielded, have indeed developed a deep and abiding fondness for the modest little creature after all.

Joel H. Rothman and Andrew Singson  
August, 2011

---

---

## ACKNOWLEDGMENTS

We would like to thank all of the chapter authors for their contributions and patience during the editing of this volume. The scientific community will greatly benefit from their efforts. We would like to thank the following people who critically reviewed the various chapters: Allison Abbott, Laura Bianchi, David Fey, Tina Gumienny, Tim Kroft, Matthew Marcello, Kim McKim, Jeremy Nance, Kevin O'Connell, Sara Olson, Richard Padgett, Gunasekaran Singaravelu, and Gillian Stanfield. We are also grateful to Stacey Nocciolo and Cari Donnelly for their assistance in managing the many files associated with this volume.

---

---

**PART I**

Genetics and Functional Genomics



---

---

## CHAPTER 1

# Mapping Mutations in *C. elegans*

**Eric J. Lambie**

Biology Department II, Ludwig-Maximilians-University, Munich, Germany

---

Abstract

- I. Introduction
  - II. Mutant Origination
  - III. Mutations That Can Be Propagated as Homozygotes
    - A. Preliminary Mutant Characterization
    - B. Mapping Crosses
  - IV. Mutations That Cannot Be Propagated as Homozygotes
    - A. Mapping with an Unbalanced Mutation
    - B. Mapping with a Balanced Mutation
  - V. Modifier Mutations
    - A. Types of Modifiers
    - B. Initial Characterization
    - C. An Example: Mapping Suppressors of *gon-2(q388)*
    - D. General Strategies for Modifier Mutations
  - VI. SNP Mapping
    - A. Bulked Lysate Analysis
    - B. After Bulked Lysates
  - VII. WGS and Beyond
- Acknowledgments  
References

---

---

## Abstract

At present, the principal goal of mapping is to establish correspondence between a mutation identified via a change in phenotype and an alteration in the DNA sequence of the genome. Recent advances in molecular biology and bioinformatics have greatly facilitated this procedure, but certain standard methods, such as the three-factor cross, continue to be extremely useful for high-resolution mapping and separation of tightly linked mutations. This chapter provides both general guidelines and specific procedures for the characterization and mapping of newly isolated mutations in *C. elegans*. Procedures are included for dealing with mutations that cannot be propagated as

homozygotes, as well as mutations that can only be scored in specialized genetic backgrounds, for example, suppressor, enhancer, and modifier mutations.

---

---

## I. Introduction

In 1974, Sydney Brenner outlined the essential methods for genetic mapping in *C. elegans* and described the relative positions of nearly 100 different loci distributed across the six linkage groups (Brenner, 1974). This provided the fundamental framework upon which the current genetic map has been built. Brenner used a combination of two- and three-factor mapping methods to determine the relative positions of loci along each chromosome. Subsequent workers have described methods for using deficiencies (i.e., deletions) (Rogalski *et al.*, 1982; Sigurdson *et al.*, 1984), free duplications (Bullerjahn and Riddle, 1988; McKim and Rose, 1990), and translocations (Albertson, 1984; McKim *et al.*, 1988; Rosenbluth and Baillie, 1981; Sigurdson *et al.*, 1986) to determine gene locations. See Fay (2006) for an extensive discussion of gene mapping methods in *C. elegans*. Since the three-point cross is arguably the single most useful standard mapping cross, its use is described in Sections VIB and VII.

With the advent of gene cloning technology, a progressively sophisticated series of methods have been enlisted to establish direct correlations between genetically defined loci and the DNA sequence of the genome. These began with Southern blot-based methods such as RFLP (restriction fragment length polymorphism) mapping (Ruvkun *et al.*, 1989) and transposon tagging (Collins *et al.*, 1987; Moerman *et al.*, 1986), and evolved into higher-resolution, high-throughput techniques for detecting polymorphisms (Davis *et al.*, 2005; Flibotte *et al.*, 2009; Fuhrman *et al.*, 2008; Jakubowski and Kornfeld, 1999; Shelton, 2006; Swan *et al.*, 2002; Wicks *et al.*, 2001; Williams, 1995). Drastic reductions in the cost of whole genome sequencing (WGS; also termed resequencing), have now made it feasible to “map” mutations simply by sequencing the entire genome of a strain that carries the newly isolated allele (Hobert, 2010). Since the standard conditions that are used for mutagenesis generate hundreds of mutations within the genome (Flibotte *et al.*, 2010; Sarin *et al.*, 2010), combining WGS and mapping can greatly reduce the number of candidate loci that need to be considered (Doitsidou *et al.*, 2010). The purpose of this chapter is to describe the fundamental principles and practices involved in establishing the basic genetic properties of a newly isolated mutation and also to discuss the use of methods for determining the correspondence between a newly isolated genetic mutation and the causative sequence alteration(s) within the genome.

---

---

## II. Mutant Origination

Mutations typically are identified via forward genetic screens, since these provide a powerful method for identifying both loss-of-function and alteration-of-function alleles of any given gene of interest. Hundreds of different single-mutant

phenotypes have been identified, but for the practical purpose of mapping, these can be divided into two categories: homozygous viable and homozygous inviable. Since different strategies are required for each of these, they are considered separately below.

Many genetic screens involve searching for mutations that have no detectable phenotype in an otherwise wild-type background. These include suppressor, enhancer, and synthetic lethal screens, as well as screens that depend on specialized markers (e.g., GFP reporters) for detection of the mutant phenotype. The underlying strategy for mapping such mutations is similar to that used for mutations that can be scored independently, but also includes consideration of events that affect the primary locus. Therefore, these are also considered in a separate section.

Although the procedures described herein will be applicable to most newly identified mutations, it is not possible to anticipate all potential idiosyncrasies that might be encountered for a given mutant stock. Therefore, it may be necessary to make appropriate modifications for temperature effects, growth rate, pleiotropy, sex-specificity, and other factors.

---

---

---

### III. Mutations That Can Be Propagated as Homozygotes

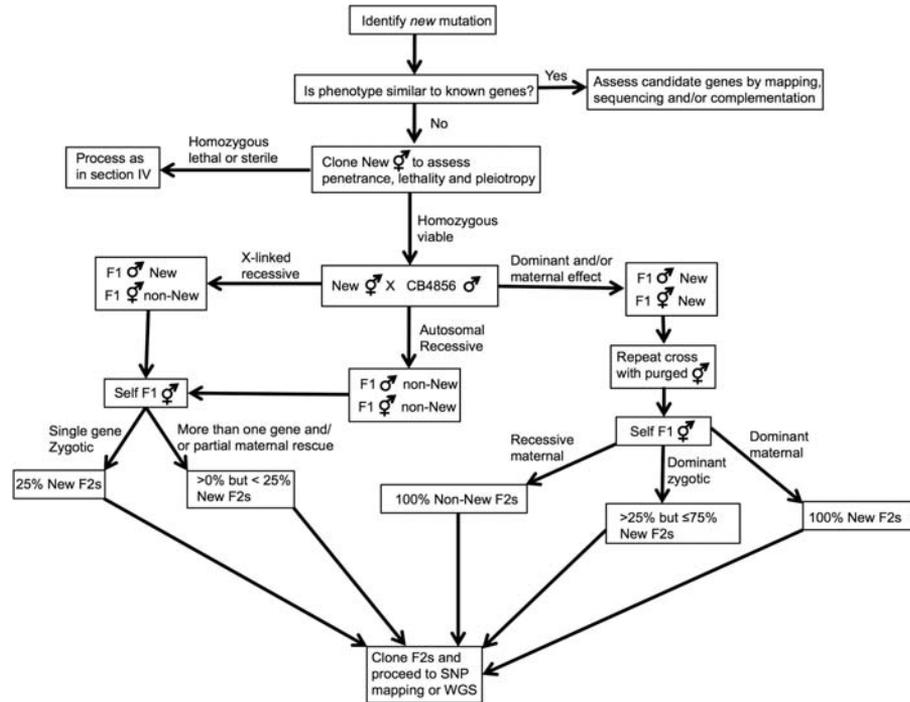
These include visible mutations such as the standard *dpy* and *unc* loci that were used to define the initial genetic map of *C. elegans*. In this chapter, I will use the term “New” to denote the phenotype that has been sought when isolating new mutations, and the terms *new-1(+)* and *new-1(\*)* to indicate, respectively, the wild-type and mutant alleles of the affected locus. Although unlikely, it is possible that New is a synthetic phenotype, that is, it results from mutations in more than one gene. In the case of two unlinked genes, the genotype would be written as *new-1(\*)*; *new-2(\*)*. If multiple mutations are responsible for the mutant phenotype, this is usually obvious from the results of the initial mapping cross (Fig. 1). However, if the mutations are linked to each other (written *new-1(\*) new-2(\*)*) this might not be clear until after further tests of gene identity have been performed.

#### A. Preliminary Mutant Characterization

Before proceeding with mapping crosses, address the following two issues.

##### 1. Mutant Similarity

Use available information resources (e.g., WormBase, PubMed) to determine whether the mutant phenotype fits the profile of any previously identified genes, or if it is similar to that of other mutations that were identified in the same mutant screen. If either of these conditions holds, then the most efficient course of action is to perform a complementation test to assess allelism and/or proceed directly with



**Fig. 1** Flowchart for characterization of newly isolated mutations. Overall progression and major conclusions are diagrammed. See text for details and caveats.

WGS. Despite the large number of mutations induced by standard mutageneses (Anderson, 1995; Flibotte *et al.*, 2010; Sarin *et al.*, 2010), if a particular gene is found to be mutated in each of two independently isolated, non-complementing mutants, then it is overwhelmingly likely that this is the gene of interest. In this case, it is important to sequence non-mutant siblings from the same mutant screen to rule out mutations that have occurred spontaneously during growth of the stock.

## 2. Assessment of Penetrance

Examine the full brood of at least three individual New hermaphrodites. To do this, pick a single L4-stage New hermaphrodite onto each of three seeded plates. After 24 h (at 20°), transfer each hermaphrodite to a new seeded plate, and repeat the transfer at 12 h intervals until the hermaphrodites are purged (have run out of sperm). Examine each plate daily for the presence of unhatched eggs (indicating embryonic lethality) or slow-growing/arrested larvae. If possible, score every animal on the plate for the New phenotype.

- a) If 100% of the progeny are New and the brood size is fairly large (>250), without evidence of any slow growers or lethals, then the stock is true breeding with complete penetrance.
- b) If a significant fraction of the progeny are sickly or inviable, but all of the survivors are New, then this is probably due to either the presence of a separate mutation (“*sick-1*(\*)”) in the strain, or multiple phenotypic effects (pleiotropy) of the *new-1*(\*) mutation. Although less likely, it is also possible that the strain has a balanced lethal genotype, for example, *new-1*(\*) *sick-2*(+)/*new-1*(+) *sick-2*(\*), where *new-1*(\*) has a dominant New phenotype, and *new-1*(\*) and *sick-2*(\*) are each lethal when homozygous. *sick-2*(\*) could be a single gene mutation in a locus that is fairly tightly linked to the *new-1* locus, or it could be a chromosomal alteration, such as a deletion, inversion, or translocation. It is not essential to distinguish between these possibilities prior to the initiation of mapping crosses.
- c) If there is no evidence of lethality, but <100% of the progeny exhibit the New phenotype, then either the mutation has incomplete penetrance or it is not homozygous (which could occur if *new-1*(\*) is dominant). Clone 12 New hermaphrodites and three non-New hermaphrodites and examine their progeny. In the case of a true-breeding stock with incomplete penetrance, each clone should exhibit the same percentage of New progeny. If the mutation was previously not homozygous, then a subset of the New clones should exhibit 100% penetrance. These should then be used for further analysis.

## B. Mapping Crosses

Traditional methods for mapping using visible markers and balancer chromosomes have been well described. These methods are still necessary in some cases; however, the overwhelming majority of mutations that have been induced in N2 Bristol-derived strains can be successfully mapped by outcrossing to the polymorphic wild-type Hawaiian strain, CB4856 (available from the Caenorhabditis Genetic Stock Center).

### 1. Mutations with 100% Penetrances

Set up a cross between CB4856 and your New stock. To do this, use a pick to transfer a small amount of *E. coli* (~3 mm diameter glob) onto the center of an unseeded plate. Add 8–10 young adult/L4-stage CB4856 males and 3–4 young adult New hermaphrodites (presumed genotype *new-1*(\*)/*new-1*(\*)). Incubate the plate at 20° for 3 days, and then inspect the progeny. If all of the progeny are New, and males are not present, then the cross was probably not successful and should be repeated, adding ~20 L4-stage males and 5–6 L4/young adult-stage hermaphrodites.

#### a) **Recessive Autosomal Mutations.**

The most likely result is that the progeny will consist of a mixture of New self-progeny hermaphrodites, plus non-New cross-progeny hermaphrodites and

males. If this result is observed, then *new-1*(\*) is recessive, autosomal, and (assuming a single locus is involved) a zygotically provided copy of *new-1*(+) is sufficient to confer a wild-type phenotype. However, this result does not exclude the possibility that a maternally expressed copy of *new-1*(+) can also confer a wild-type phenotype, and this will be tested by examining the F2s, as described below. Note that this result is also consistent with the unlikely possibility of a paternal effect mutation (phenotype usually determined by genotype of sperm; see Browning and Strome (1996), Hill *et al.* (1989), Seidel *et al.* (2008), but also Darby *et al.* (2007)).

Next, pick a single non-New F1 hermaphrodite onto each of 10 plates, incubate at 20° for 4 days (assuming that New is optimally scorable in adults), and then inspect the progeny.

i) **Zygotic-Effect Mutations.**

The most likely result is that ~1/4 of the F2 progeny are New. In this case, it is highly likely that a single gene recessive, zygotic-effect mutation is responsible for the mutant phenotype (i.e., phenotype is determined by the genotype of zygote). The New progeny will have genotype *new-1*(\*)/*new-1*(\*), and thus be homozygous for N2 sequences in the vicinity of the *new-1* locus. Pick New and non-New F2s for SNP mapping, as described in Section VI.

ii) **Partial Maternal Effect Mutations.**

If New F2s are present, but the fraction is less than 1/4, but clearly greater than 1/16, then it is likely that the *new-1*(+) allele present in the F1 hermaphrodite confers partial maternal rescue of the homozygous *new-1*(\*) progeny. If the fraction of New F2s is  $\leq 1/16$ , then it could be that more than one locus is responsible for the New phenotype and/or that partial maternal rescue occurs. These possibilities will be resolved by the mapping data. Regardless of the basis for incomplete penetrance in the F2s, animals that are New are likely to be homozygous for *new-1*(\*) (or both *new-1*(\*) and *new-2*(\*) if two loci are involved). Pick New and non-New F2s for SNP mapping, as described in Section VI.

iii) **Recessive X-linked Mutations**

The next most likely result is that all of the hermaphrodite progeny are non-New, but all of the F1 males are New. In this case, *new-1* is almost certainly X-linked. In this case, use the F1 hermaphrodites to proceed with SNP mapping as for autosomal recessive mutations. Note, this result is also consistent with the unlikely possibility that *new-1*(\*) is specifically dominant in males. This is rare, but has been observed (Barton and Kimble, 1990; E.J.L., unpublished), and should be considered if candidate genes on the X do not satisfy confirmatory tests.

iv) **Dominant and Maternal Effect Mutations**

If both males and hermaphrodites are present among the F1 progeny, but all are New, then it is likely that either *new-1*(\*) has a dominant zygotic effect or *new-1*(\*) has a strict maternal effect (i.e., phenotype of progeny is solely

determined by the genotype of the mother; *new-1*(\*) could be either dominant or recessive, since the mother is homozygous). An alternative possibility is that the cross was unsuccessful and the *new-1*(\*) stock has a Him (high-incidence males) phenotype; this can be resolved simply by inspecting the original *new-1*(\*) stock for males and then repeating the cross, using purged hermaphrodites if necessary (see below).

In cases where *new-1*(\*) appears to be dominant and/or has a maternal effect, the mapping cross will need to be repeated using purged hermaphrodites.

To obtain purged hermaphrodites, pick  $\geq 20$  young adult New hermaphrodites onto a seeded plate and incubate at 20°. After 48 h, transfer the adult worms to a fresh plate and inspect them to see if any have expended their sperm. Early-stage purged hermaphrodites will have unfertilized oocytes in the uterus. Unlike fertilized eggs, these lack an eggshell and give the uterus a fairly homogeneous, light brown appearance. As the hermaphrodites run out of sperm, they will begin to lay unfertilized oocytes on the plate. These superficially resemble fertilized eggs; however, they are soft and easily ruptured, typically degenerating into irregular brown blobs within the *E. coli* lawn. Late-stage purged hermaphrodites will have an empty uterus and the oocytes within the proximal oviduct will be compressed to give a “piano key” appearance. If you are uncertain, it is better to pick slightly later-stage purged hermaphrodites; however, if they are too old they might not produce progeny in response to mating.

Set up the mapping cross using 10–20 young adult CB4856 males and 3–10 purged New hermaphrodites. Incubate at 20° and inspect after 24 h. If no eggs are present on the plate, then the hermaphrodites were too old and the cross needs to be repeated. If eggs are present, then incubate at 20° for an additional 48 h, then inspect the F1 progeny. If the hermaphrodites were fully purged, then ~50% of the progeny should be males.

Clone 10 young adult F1 hermaphrodites. Incubate at 20° for 24 h, then transfer each to a new plate, and continue incubation at 20°. Inspect the F2 progeny 4 days after the hermaphrodites were initially cloned (assuming that New is optimally scorable in adults).

**b) Dominant Zygotic Effect**

If substantially more than 25% (and up to 75%) of the F2s are New, this is consistent with a dominant/semi-dominant, zygotic-effect mutation. Pick New and non-New F2s for SNP mapping, as described in Section VI.

A dominant maternal effect mutation that has incomplete penetrance could also produce this ratio, and this should be considered if the SNP mapping procedure does not provide evidence of linkage. Note that such results are also consistent with the unlikely possibility that the New phenotype is due to a mutation in the mitochondrial genome. If desired, this possibility can be tested by determining whether *new-1*(\*) can be transmitted through males (see Tsang and Lemire, 2002).

c) **Recessive Maternal Effect**

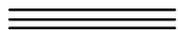
If 100% of the F2s are non-New, this suggests that *new-1*(\*) has a recessive maternal absence effect (a single maternal copy of *new-1*(+) is sufficient to rescue the mutant phenotype). In this case, homozygotes can be identified by examining the F3s. Pick 100 adult F2 hermaphrodites onto two or more seeded plates, allow them to lay eggs for 90 min, then either pick them off or rinse them off with M9 buffer (the eggs will stick to the plate). Once the F3s have reached a stage where the mutant phenotype is scorable, pick New and non-New animals for SNP analysis as described in Section VI. New animals will be *new-1*(\*)/*new-1*(\*) (assuming that no males are present on the plate), since they must have been derived from homozygous mothers. Non-New animals have only a 1/6 chance of being *new-1*(\*)/*new-1*(\*), since only 2/3 of them will be derived from *new-1*(\*)/*new-1*(+) mothers.

d) **Dominant Maternal Effect**

If 100% of the F2s are New, this suggests a dominant maternal effect mutation (or a mitochondrial mutation; see above). Both of these are expected to be rare classes. For dominant maternal effect mutations, the same procedure is used as for recessive maternal effect mutations, bearing in mind that the non-New progeny are the more informative class, since they will be homozygous for the CB-derived SNP.

## 2. Mutations with Incomplete Penetrance

If the New phenotype has incomplete penetrance in the original stock, then it is advisable to use purged hermaphrodites when crossing with CB4856 males since this will eliminate the background of self-progeny. The interpretations of cross results and subsequent mapping procedures are very similar to those for fully penetrant mutations. However, the fraction of progeny exhibiting the mutant phenotype will be diminished, which may make it difficult to determine whether maternal rescue occurs and/or multiple gene mutations are involved. Nevertheless, the surest strategy is to pick multiple independent New and non-New animals of the appropriate generation after outcrossing to CB4856, and then perform SNP mapping as described in Section VI.



## IV. Mutations That Cannot Be Propagated as Homozygotes

If the *new-1*(\*) mutation arises in a screen that was not specifically designed to recover sterile/lethal mutations, then the stock will have to be maintained by cloning the *new-1*(\*)/*new-1*(+) and *new-1*(+)/*new-1*(+) siblings of New animals and scoring their progeny to identify plates that carry *new-1*(\*). This is not a practical long-term solution, so *new-1*(\*) should be placed over a balancer chromosome as soon as possible (see Ann Rose's chapter on balancer chromosomes in this volume). As

for mutations that are homozygous viable/fertile, apply the criteria in Section III.A.1 prior to proceeding further.

### A. Mapping with an Unbalanced Mutation

If no map position has yet been established, set up a cross with 10–15 young adult/L4 males of strain CB5600 (genotype: *ccls4251[myo-3::GFP]* I; *him-8(e1489)* IV) and 6–10 hermaphrodite progeny obtained from a worm of genotype *new-1(\*)/new-1(+)*. *ccls4251* is an insertion near the center of chromosome I (map position + 4.0) that dominantly expresses GFP within body muscle nuclei. CB5600 is available from the Caenorhabditis Genetic Stock Center. Incubate at 20° for 3 days, then use a dissecting microscope equipped with GFP optics to inspect the progeny. If GFP(+) New males are present (and GFP(+) New hermaphrodites are absent), then *new-1* is very likely on the X chromosome. If none of the GFP(+) progeny are New, then clone 12 GFP(+) F1 hermaphrodites and incubate at 20°. After 4 days, inspect the progeny to identify plates that carry the *new* mutation. If *new* is unlinked to *ccls4251*, then 75% of the New progeny will be GFP(+).

In parallel to cloning the GFP(+) F1 hermaphrodites, set up a cross with 10–20 F1 males and 5–8 CB4856 hermaphrodites. If *new-1* is not found to be linked to *ccls4251*, then clone 24 GFP(+) hermaphrodite progeny. Incubate these at 20°, then score each plate for the presence of New progeny. Approximately four of these plates should segregate New offspring. Pick New and non-New progeny for SNP mapping as in Section VI.

### B. Mapping with a Balanced Mutation

The exact procedure to be followed will depend on the properties of the balancer chromosome. For example, if *new-1(\*)* is being kept across from a balancer chromosome that carries a GFP marker, then one can cross CB4856 males with the balanced strain and pick non-New non-GFP F1 hermaphrodites. Self these, then pick New and non-New F2s for SNP analysis as in Section VI.

---

---

---

## V. Modifier Mutations

### A. Types of Modifiers

A typical modifier screen begins with animals that are homozygous for a primary mutation, *prim-1(\*)*. The *prim-1(\*)* allele could be dominant, as in the case of an integrated GFP expression reporter, or recessive, as in the case of a typical loss-of-function mutation. To perform the screen, *prim-1(\*)* animals are mutagenized and progeny are then screened in the F1 or subsequent generations for animals that are Mod, that is, exhibit some modification of the Prim

phenotype. Most likely this will be either an exaggeration of the Prim phenotype (enhancement), or a reduction/elimination of the Prim phenotype (suppression). It is also possible to identify modifier mutations that cause a qualitative change in the Prim phenotype (neomorphism), particularly in the case of a GFP expression reporter, where loss-of-function mutations that affect the specification or differentiation of cells that express the reporter can lead to an altered expression pattern (e.g., if the labeled cell is eliminated, duplicated, or mispositioned). In some cases, modifier screens are performed that involve primary mutations in more than one locus, for example, a loss-of-function allele, *prim-1*(\*), in combination with an integrated GFP reporter, *prim-2*(\*) (Grote and Conradt, 2006; Hammarlund *et al.*, 2009; Hatzold and Conradt, 2008; Nehme *et al.*, 2010). When performing mapping and characterization, the same fundamental principles apply, but attention must be given to maintaining both *prim-1*(\*) and *prim-2*(\*) through the mapping procedures.

## B. Initial Characterization

The first experimental step in characterization of a Mod stock is to determine whether the *prim-1* locus itself has been affected. For example, if *prim-1*(\*) is a weak loss-of-function allele, then second-site mutations within *prim-1* could potentially result in either enhancement or suppression of the Prim phenotype, depending on the nature of the modifier allele. Or, if *prim-1*(\*) is an integrated GFP reporter, mutation of the reporter (or recombination, in the case of integrated extrachromosomal arrays) could result in attenuation of GFP expression.

The simplest method for determining whether the *prim-1* locus has been altered in a given Mod stock is to perform an outcross to a non-Mod *prim-1*(\*) stock. I will illustrate this procedure by using an example of a modifier screen that was performed in my laboratory:

## C. An Example: Mapping Suppressors of *gon-2(q388)*

We began with a strain of genotype *gon-2(q388) unc-29(e1072)*. *gon-2(q388)* is a recessive, temperature-sensitive, loss-of-function mutation that causes a highly penetrant gonadless/sterile phenotype when animals are raised at restrictive temperature (Sun and Lambie, 1997). *unc-29* is tightly linked to *gon-2*, and the *e1072* is a recessive allele that causes a mild, but easily scorable Unc phenotype. The *unc-29(e1072)* mutation serves two purposes: it enables cross progeny to be distinguished from self-progeny during outcrossing, and it allows identification of revertant mutations within *gon-2* based on linkage. We mutagenized *gon-2(q388) unc-29(e1072)* hermaphrodites, and then raised progeny at the restrictive temperature to select for fertile Mod derivatives (Church and Lambie, 2003).

For initial characterization, hermaphrodites from each Mod stock were crossed with males of genotype *gon-2(q388); him-8(e1489)* that had been raised at

permissive temperature. These crosses were done at permissive temperature and F1 cross progeny were identified based on their non-Unc phenotype. F1 hermaphrodites were then selfed at restrictive temperature to obtain information regarding linkage and dominance of the modifier mutations.

Approximately half of the modifier mutations behaved as expected for single gene dominant alleles, with >50% of the F2 progeny exhibiting fertility. Some of these appeared to be intragenic revertants in *gon-2*, since all of the Unc progeny were fertile, and this was confirmed by DNA sequencing (E.J.L., unpublished). Most of the dominant mutations were clearly unlinked to *gon-2*, since ~25% of the Unc progeny were sterile. Due to their dominant nature, these mutations could not be tested for allelism via complementation testing. However, we deduced that each was X-linked, because when we crossed F1 males (genotype *gon-2(q388); mod/0*) with hermaphrodites of genotype *gon-2(q388) unc-29(e1072)* and raised the progeny at the restrictive temperature, nearly all of the non-Unc hermaphrodite progeny were fertile, but nearly all of the male progeny were gonadless; through positional cloning and sequencing we subsequently determined that all were alleles of the same gene, *gem-1* (*gon-2* extragenic modifier) (Kemp *et al.*, 2009).

Among the recessive mutations, a few were clearly linked to *unc-29*, since all of the Unc F2 progeny were fertile. These were confirmed to be intragenic revertants of *gon-2* by DNA sequencing. The other Mod strains segregated  $\leq 25\%$  fertile F2 progeny, consistent with a single gene recessive mutation. We tested these for linkage to *him-8* by cloning fertile F2s and scoring the Him phenotype (i.e., examining their broods for the presence of high frequency males). In two cases, approximately 25% of the fertile F2s were homozygous for *him-8(e1489)*, indicating a lack of linkage. We subsequently used conventional mapping methods to assign these mutations to chromosomes II (*gem-2*) and III (*gem-3*) (E.J.L., unpublished). In the other cases, few if any of the fertile F2s were Him, suggesting that the mutations were balanced by *him-8(e1489)* on chromosome IV. These mutations failed to complement each other, and we used positional cloning and sequencing to assign them to the same gene, *gem-4* (Church and Lambie, 2003).

#### D. General Strategies for Modifier Mutations

Essentially the same methods that we used for characterizing suppressors of *gon-2* can be applied for the initial characterization of any Mod mutations. In the case of extragenic Mod mutations, one would then apply the following strategy (e.g., with *prim-1* tightly linked to *unc-29*). Cross hermaphrodites of genotype *prim-1(\*) unc-29(e1072); mod-1(\*)* with CB4856 males. Pick one non-Unc F1 hermaphrodite onto each of 10 plates. For a zygotic-effect mutation, proceed with SNP mapping as in Section VI by picking Mod and non-Mod F2s. For maternal effect mutations, pick 100 F2s onto each of two plates, allow them to lay eggs for 1–2 h, and then rinse off with M9 buffer to remove adults. Pick Mod and non-Mod F3s for SNP mapping.

---

---

## VI. SNP Mapping

Assuming that only a single gene mutation is involved, the genotype of the F1s derived from crossing a New mutant stock with CB4856 is *new-1*(\*)/*new-1*(+), where the *new-1*(+) allele is present on a CB4856-derived chromosome. Since each F2 contains two independently derived recombinant genomes, for a section of the genome that is unlinked to *new-1* a given individual has equal likelihood (25%) of being homozygous for either N2 sequence or CB4856 sequence, and 50% likelihood of being heterozygous. However, at the *new-1* locus, individuals that are homozygous for *new-1*(\*) will have a 100% likelihood of being homozygous for N2 sequence (and 0% chance of being heterozygous). As the distance between a particular sequence and the *new-1* locus increases, the likelihood of SNP heterozygosity increases (up to a maximum of 50% at 50 cM). Since *C. elegans* chromosomes are only 50 cM in length, and the majority of the genes on each chromosome are located within the central gene clusters, assignment to linkage group can usually be accomplished by testing only five centrally situated SNPs, as described below (for recessive mutations, X-linkage can usually be determined by inspecting the F1 progeny; dominant X-linked mutations are distinguishable since F1 mutant males will transmit to 100% of their hermaphrodite progeny, but none of their male progeny. If the mutation prevents males from mating, methods comparable to those used for autosomal mapping must be used).

### A. Bulk Lysate Analysis

Currently, the easiest method for SNP mapping is bulked lysate analysis (Wicks *et al.*, 2001). Here, I describe the use of ARMS-PCR (annealing restricted marker system; Ye *et al.*, 2001) for bulked lysate SNP analysis. This method uses sequence-specific primer annealing to permit the assessment of SNPs that do not alter restriction enzyme cleavage sites. Note, that by the time this chapter is published technical advances and cost reductions might permit WGS to be economically substituted for SNP mapping at this stage of the analysis (Doitsidou *et al.*, 2010).

1. Set up two experimental lysates, one with New worms and the other with non-New worms. Also set up three control lysates, one with only N2 worms, one with only CB4856 worms, and one with equal numbers of CB4856 and N2 worms. To make a lysate, pick 100 animals into a PCR tube containing 200  $\mu$ L of lysis buffer (50 mM KCL, 10 mM Tris pH 8.3, 2.5 mM MgCl<sub>2</sub>, 0.45% NP-40, 0.45% Tween-20, 0.01% gelatin; keep stock frozen at  $-20^{\circ}$ , thaw and add proteinase K to 100  $\mu$ g/mL immediately prior to use (Williams *et al.*, 1992)). Transfer the tube to  $-70^{\circ}$ . After incubating at  $-70^{\circ}$  for at least 10 min, transfer the tube to a PCR machine set to  $60^{\circ}$  for 4 h, followed by  $95^{\circ}$  for 15 min.

2. Set up five reaction tubes for each lysate, one for each primer set (Table I):

Reagent	Amount per reaction:
H <sub>2</sub> O	3.5 $\mu$ L
50 mM MgCl <sub>2</sub>	0.1 $\mu$ L
Primer mix (6.25 pmol/ $\mu$ L each) <sup>a</sup>	0.4 $\mu$ L
Taq Master Mix (Promega 9PIM750)	5.0 $\mu$ L

<sup>a</sup> Except for primers o1295 and o1312, which should be at 3.12 pmol/ $\mu$ L)

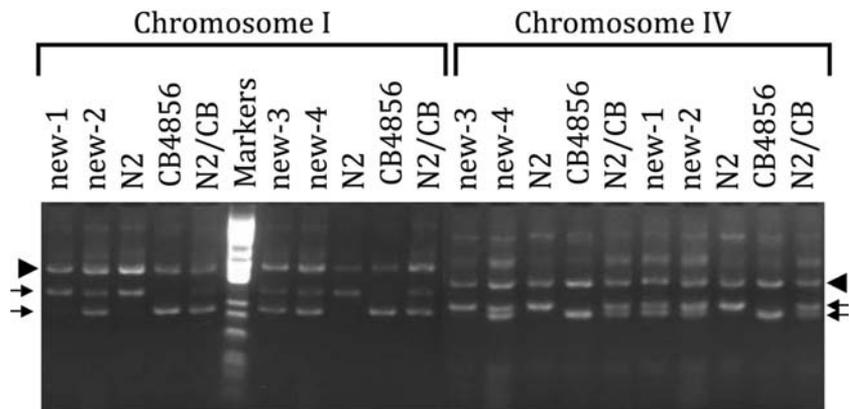
**Table I**  
ARMS-PCR primers

Chromosome I = pkP1114	
o1287	TTTCGGAATCAGTTTTATCTTGACGGAT
o1288	TCCAACGATGCTTCAACCGC
o1289	TGTTCAATTCTACGAGGCAAGATATCG
o1290	TCATCCTTGTCCCAGACTACAATCTCAA
N2 product: ~280 bp	
CB4856 product: ~200 bp	
Chromosome II = pkP2149	
o1273	AAAAAACTGAAGGAGACCGAGACACAT
o1274	ACCATCGCATTTGCTTCAGATCAG
o1275	TTCTTCTCGTGAAACAGCTTCATCAACT
o1276	ATTTCAATTTGCAGCTCGGTAACATTTTC
N2 product: ~225 bp	
CB4856 product: ~350 bp	
Chromosome III = ce3-171	
o1311	GGTGTGGACCAGCTTTTGCG
o1312	GGATTTGGCGGAATCTATCACCTG dilute 1:2 (125 uM);
o1313	CGACCTTCTGCGTCAGTATTGTTGA
o1314	AAAATTCGGCGGTCTTCATGAACTT
N2 product: ~350 bp	
CB4856 product: ~450 bp	
Chromosome IV = ce4-167	
o1204	GCATTATTGTTTCCGATTTTAAATGGTC
o1205	GCTCAGGACGATACATGTTTAAATGGTA
o1206	TCACACGATACATCAAAGGATATCGTC
o1207	TTCAACAACATAAACGTTCTGTGGATTG
N2 product: ~190 bp	
CB4856 product: ~175 bp	
Chromosome V = ce5-15	
o1295	TTTTTACTGTCAAATCGTCATGAGAGCT diluted 1:2 (125 uM)
o1296	AAGTACTGGCCAAATTTCAACGATCAAT
o1297	CCTCGTTGTTTCGTTTTCAGGAAATTATGT
o1298	GGGCATATTAATGATAAGGGTTGCAAAA
N2 product: ~200 bp	
CB4856 product: ~350 bp	

3. Preheat the PCR machine block to 95°, add the tubes, and then run the following program:

1.	95°	:30
2.	67°	:30–1.0 per cycle
3.	72°	:30
4.	Go to 1 12 times	
5.	95°	:30
6.	55°	:30
7.	72°	:30
8.	Go to 5 24 times	
9.	72°	2:00
10.	22°	For ever

4. Run half of the sample for each reaction on a 4% agarose gel. If your reaction conditions are working properly, for each primer set you should see an upper band that is present in all lanes (arrowhead in Fig. 2), one lower band that is specific to N2, and another lower band that is specific to CB4856 (arrows in Fig. 2). The two lower bands may not be of equal intensity in the lysate that contains both N2 and CB4856 worms; however, their intensities relative to each other should remain constant in any lysate that contains equal amounts of N2 and CB4856 DNA at a



**Fig. 2** An example of ARMS-PCR mapping. Results for SNPs on chromosomes I and IV are shown for four different zygotic-effect recessive mutants, *new1-new4*. *new-1* is clearly linked to the SNP on chromosome I, whereas *new-3* is linked to the SNP on chromosome IV. Arrowhead indicates upper band that is amplified from both N2 and CB4856. Arrows indicate lower products specific to N2 and CB4856. The size markers used are pGEM<sup>®</sup> markers from Promega (the bands in the triplet near the N2- and CB4856-specific bands are 222, 179, and 126 bp).

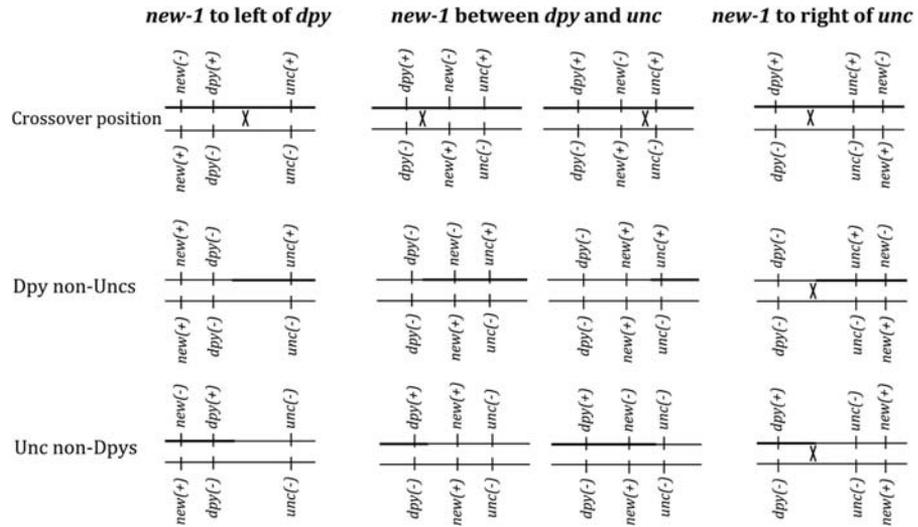
given site. If you are working with a recessive mutation, then linkage to a SNP is indicated by an increase in the ratio of N2:CB in the lysate made from New animals. If you are working with a dominant mutation, then linkage is indicated by an increase in the ratio of CB:N2 in the lysate made from non-New worms. If two unlinked mutations are responsible for the New phenotype, then you are likely to see evidence of linkage to two different SNPs. Note that [Seidel \*et al.\* \(2008\)](#) have shown that the distal left arm of chromosome I is refractory to becoming homozygous for CB4856 sequences after intercrossing between N2 and CB4856.

## B. After Bulk Lysates

Thousands of sites that are polymorphic between N2 and CB4856 have been identified and these can be accessed through WormBase. For extensive lists of SNPs and methods for their detection, see [Davis \*et al.\* \(2005\)](#), [Fuhrman \*et al.\* \(2008\)](#), [Jakubowski and Kornfeld \(1999\)](#), [Shelton \(2006\)](#), [Swan \*et al.\* \(2002\)](#), and [Wicks \*et al.\* \(2001\)](#). These SNPs can be used for bulked lysate analysis, or they can be used to analyze lysates made from individual animals if one wishes to obtain multipoint mapping data. However, fairly large sample sizes are needed in order to achieve high resolution when multipoint mapping is done this way.

In order to achieve high-resolution SNP mapping data, it is necessary to place the *new-1*(\*) mutation on the same chromosome as a marker mutation. This is typically done using a standard three-point cross ([Fig. 3](#)). For example, if SNP mapping data indicated that *new-1* is situated within/near the chromosome II gene cluster, then one would construct a strain of genotype *new-1*(\*)/*dpy-10*(*e124*) *unc-4*(*e120*), then pick Dpy non-Unc and Unc non-Dpy recombinants (*e124* and *e120* are both recessive). If *new-1* is to the left of *dpy-10*, then approximately all of the Unc non-Dpy recombinants will carry *new-1*(\*) on a recombinant chromosome of genotype *new-1*(\*) *unc-4*(*e120*). If *new-1* is to the right of *unc-4*, then approximately all of the Dpy non-Unc recombinants will carry *new-1*(\*) on the recombinant chromosome. If *new-1* is between *dpy-10* and *unc-4*, some recombinants of each class will carry *new-1*(\*) and some will not. Thus, this procedure not only generates a chromosome where *new-1*(\*) is adjacent to a recessive visible marker, but also provides information about map position.

In order to perform SNP mapping, CB4856 males are crossed with a strain that carries a marked *new-1*(\*) chromosome, for example, *dpy-10*(*e124*) *new-1*(\*), to generate F1s of genotype *dpy-10*(*e124*) *new-1*(\*)/*dpy-10*(+) *new-1*(+). This strain is not balanced, but it can be propagated by cloning wild-type animals and verifying that they segregate Dpy New progeny. Mapping is then done by cloning Dpy non-New and New non-Dpy recombinants, making lysates and scoring SNPs by PCR. Given the high density of SNPs, this is a very powerful method for gene mapping. However, depending on the nature of the New and marker phenotypes, picking recombinants can be tedious and/or difficult, so this method is being replaced by WGS technology whenever possible.



**Fig. 3** An example of a three-factor cross. A mutation that has been only roughly mapped could be either within or outside the interval defined by the marker mutations. Note that if *new-1* were situated between the *dpy* and *unc* loci, but very near the *dpy* locus, it would produce the same results as if *new-1* were to the left of *dpy* (and likewise for the *unc* locus). Since the *dpy* and *unc* loci are fairly close to each other, nearly all of the Unc non-Dpy and Dpy non-Unc recombinants will carry one recombinant chromosome and one parental chromosome of genotype *dpy*(-) *new*(+) *unc*(-).

## VII. WGS and Beyond

Once the approximate map position of *new-1* has been determined by SNP mapping, the most efficient strategy is to proceed with WGS using the methods described by Sarin and Flibotte (Flibotte *et al.*, 2010; Sarin *et al.*, 2010). Since these methods are designed for strains that are homozygous viable and fertile, they need to be adapted for sterile/lethal mutations. In the case of sterile mutants, it is technically feasible (albeit laborious) to pick 5000–10,000 homozygotes derived from a balanced stock, then purify DNA and proceed with WGS (O. Hobert, personal communication). For lethal mutations, WGS could be done using DNA purified from animals of genotype *new-1*(\*)/Bal. However, much higher-fold coverage than usual would be necessary in order to distinguish heterozygous mutations from background (O. Hobert, personal communication), so this could be an expensive option.

Regardless of whether the mutation being mapped is homozygous viable or not, it is highly likely that multiple mutations will be present within the vicinity of the *new-1* locus, so additional work will be necessary to determine which corresponds to *new-1*(\*) (Flibotte *et al.*, 2010; Sarin *et al.*, 2010). Multiple criteria need to be applied in order to make this determination. The correct locus is likely to: a) confer transformation rescue of *new-1*(\*), b) produce a New phenotype when inactivated by

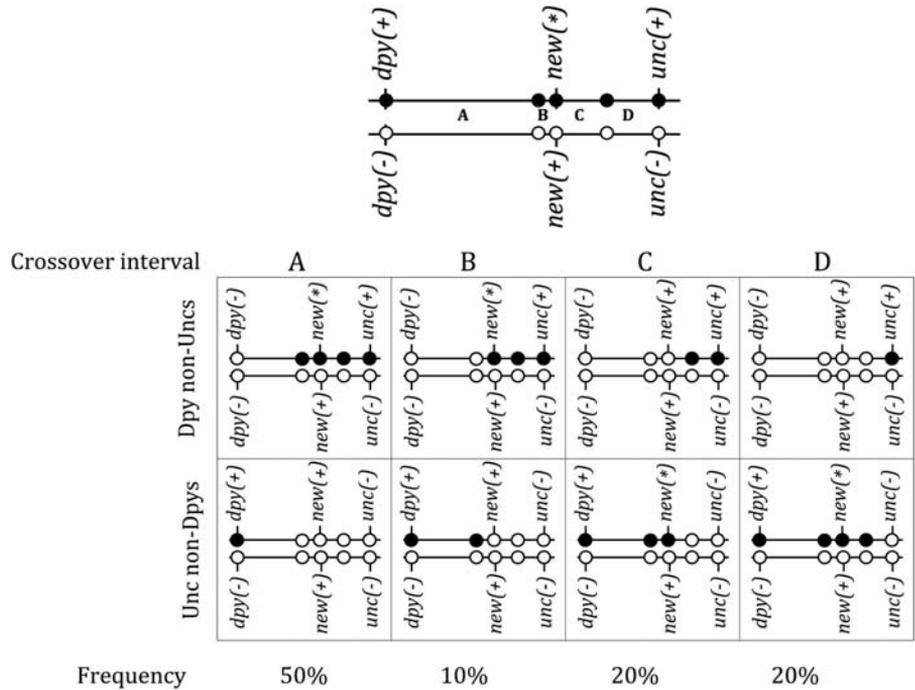
RNAi, c) produce a New phenotype (or fail to complement *new-1*(\*)) when inactivated by deletion, d) contain sequence alterations when *new-1*(\* is reverted, e) contain sequence alterations when additional alleles are isolated by screening for alleles that fail to complement *new-1*(\*), f) invariably correlate with the New phenotype when recombination events are selected for in the vicinity of *new-1*. Since this last criterion is arguably the most definitive and also the most directly pertinent to the subject of this chapter, I will describe the procedure below.

First, based on the combined data from SNP mapping and WGS, choose a pair of recessive visible mutations that are likely to flank *new-1*. Ideally, neither mutation will interfere with scoring the New phenotype. For example, if *new-1* is situated within the chromosome I gene cluster, *dpy-5*(*e61*) and *unc-29*(*e1072*) would be a good choice since these encompass the entire cluster.

In the case of homozygous viable mutations, cross N2 males with *new-1*(\* hermaphrodites, then mate the F1 males with *dpy-5*(*e61*) *unc-29*(*e1072*) hermaphrodites. Clone 12 non-Dpy non-Unc L4-stage hermaphrodites and allow them to produce self-progeny. Animals of genotype *dpy-5*(+) *new-1*(\* *unc-29*(+)/*dpy-5*(*e61*) *new-1*(+) *unc-29*(*e1072*) will segregate New:wild type:Dpy Unc progeny in a 1:2:1 ratio. This is a balanced stock that can be maintained by picking wild-type animals and ensuring that they continue to segregate the proper phenotypes. Dpy non-Unc and Unc non-Dpy recombinant progeny will be produced by such animals and these can be cloned to obtain recombinant chromosomes (Fig. 4). Dpy non-Unc and Unc non-Dpy recombinants should be cloned and their progeny scored for the New phenotype. If desired, animals can be subcloned to obtain stocks that are homozygous for the recombinant chromosome, but this is not essential prior to scoring candidate mutations by PCR/sequencing. The *new-1*(\* mutation should invariably correlate with the segregation of New progeny. If the New phenotype is synthetic, then more than one mutation will always be present in New animals.

In the case of lethal/sterile mutations (e.g., if *new-1* is in the chromosome I gene cluster), cross *new-1*(\**Bal* hermaphrodites with wild-type males, then pick non-Bal male progeny and cross them with *dpy-5*(*e61*) *unc-29*(*e1072*) hermaphrodites. Clone non-Dpy non-Unc F1 progeny and self them to identify clones of genotype *dpy-5*(*e61*) *new-1*(+) *unc-29*(*e1072*)/*dpy-5*(+) *new-1*(\* *unc-29*(+). This stock can be maintained by picking wild-type animals and examining their progeny to ensure that they segregate both Dpy Unc and New progeny. Next, clone Dpy non-Unc and Unc non-Dpy recombinants from this stock and examine their progeny to determine whether or not they carry *new-1*(\*). Use PCR/sequencing to score candidate mutations among the recombinants.

The number of recombinants that will need to be picked is dependent on the spacing between candidate mutations. For example, the interval flanked by *dpy-5* and *unc-29* spans 3.5 megabases (Mb). If two candidate mutations present within this interval are separated by 0.5 Mb, then the likelihood that they are not separated by a given recombination event is  $6/7 = 0.857$  (under the simplifying assumption that recombination frequency is uniform across the interval). Therefore, if



**Fig. 4** Selecting for recombination in a given interval in order to separate candidate *new-1*(\*) mutations. The actual mutation corresponding to *new-1*(\*) is likely to be flanked by other mutations induced by the mutagenesis (or occurring spontaneously during strain propagation). Sites of sequence discrepancy between the *dpy*(-) *unc*(-) marker chromosome and the *new-1*(\*) chromosome are indicated by open and filled circles, respectively. Recombination events in intervals B and C permit the bona fide *new-1*(\*) mutation to be distinguished from background mutations.

21 recombinants are picked, the likelihood that at least one has a breakpoint between these two candidate mutations is  $1 - (0.857)^{21} = 0.96$ .

## Acknowledgments

The author is grateful to Barbara Conradt, David Fay and Andy Singson for constructive comments on the manuscript, Jess Whitaker and Corinne Pierce for testing ARMS-PCR conditions, and Oliver Hobert for helpful discussions and providing unpublished information.

## References

- Albertson, D. G. (1984). Localization of the ribosomal genes in *Caenorhabditis elegans* chromosomes by in situ hybridization using biotin-labeled probes. *EMBO J.* **3**, 1227–1234.
- Anderson, P. (1995). Mutagenesis. *Methods Cell Biol.* **48**, 31–58.

- Barton, M. K., and Kimble, J. (1990). *fog-1*, a regulatory gene required for specification of spermatogenesis in the germ line of *Caenorhabditis elegans*. *Genetics* **125**, 29–39.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- Browning, H., and Strome, S. (1996). A sperm-supplied factor required for embryogenesis in *C. elegans*. *Development* **122**, 391–404.
- Bullerjahn, A. M., and Riddle, D. L. (1988). Fine-structure genetics of *ama-1*, an essential gene encoding the amanitin-binding subunit of RNA polymerase II in *Caenorhabditis elegans*. *Genetics* **120**, 423–434.
- Church, D. L., and Lambie, E. J. (2003). The promotion of gonadal cell divisions by the *Caenorhabditis elegans* TRPM cation channel GON-2 is antagonized by GEM-4 copine. *Genetics* **165**, 563–574.
- Collins, J., Saari, B., and Anderson, P. (1987). Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature* **328**, 726–728.
- Darby, C., Chakraborti, A., Politz, S. M., Daniels, C. C., Tan, L., and Drace, K. (2007). *Caenorhabditis elegans* mutants resistant to attachment of *Yersinia* biofilms. *Genetics* **176**, 221–230.
- Davis, M. W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S., and Jorgensen, E. M. (2005). Rapid single nucleotide polymorphism mapping in *C. elegans*. *BMC Genom.* **6**, 118.
- Doitsidou, M., Poole, R. J., Sarin, S., Bigelow, H., and Hobert, O. (2010). *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. *PLoS One* **5**, e15435.
- Fay, D. Genetic mapping and manipulation: Chapter 1-Introduction and basics (February 17, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.90.1, <http://www.wormbook.org>.
- Flibotte, S., Edgley, M. L., Chaudhry, I., Taylor, J., Neil, S. E., Rogula, A., Zapf, R., Hirst, M., Butterfield, Y., Jones, S. J., Marra, M. M., Barstead, R. J., and Moerman, D. G. (2010). Whole-genome profiling of mutagenesis in *Caenorhabditis elegans*. *Genetics* **185**, 431–441.
- Flibotte, S., Edgley, M. L., Maydan, J., Taylor, J., Zapf, R., Waterston, R., and Moerman, D. G. (2009). Rapid high resolution single nucleotide polymorphism-comparative genome hybridization mapping in *Caenorhabditis elegans*. *Genetics* **181**, 33–37.
- Fuhrman, L. E., Shianna, K. V., and Aballay, A. (2008). High-throughput isolation and mapping of *C. elegans* mutants susceptible to pathogen infection. *PLoS One* **3**, e2882.
- Grote, P., and Conradt, B. (2006). The PLZF-like protein TRA-4 cooperates with the Gli-like transcription factor TRA-1 to promote female development in *C. elegans*. *Dev. Cell* **11**, 561–573.
- Hammarlund, M., Nix, P., Hauth, L., Jorgensen, E. M., and Bastiani, M. (2009). Axon regeneration requires a conserved MAP kinase pathway. *Science* **323**, 802–806.
- Hatzold, J., and Conradt, B. (2008). Control of apoptosis by asymmetric cell division. *PLoS Biol.* **6**, e84.
- Hill, D. P., Shakes, D. C., Ward, S., and Strome, S. (1989). A sperm-supplied product essential for initiation of normal embryogenesis in *Caenorhabditis elegans* is encoded by the paternal-effect embryonic-lethal gene, *spe-11*. *Dev. Biol.* **136**, 154–166.
- Hobert, O. (2010). The impact of whole genome sequencing on model system genetics: get ready for the ride. *Genetics* **184**, 317–319.
- Jakubowski, J., and Kornfeld, K. (1999). A local, high-density, single-nucleotide polymorphism map used to clone *Caenorhabditis elegans cdf-1*. *Genetics* **153**, 743–752.
- Kemp, B. J., Church, D. L., Hatzold, J., Conradt, B., and Lambie, E. J. (2009). *gem-1* encodes an SLC16 monocarboxylate transporter-related protein that functions in parallel to the *gon-2* TRPM channel during gonad development in *Caenorhabditis elegans*. *Genetics* **181**, 581–591.
- McKim, K. S., Howell, A. M., and Rose, A. M. (1988). The effects of translocations on recombination frequency in *Caenorhabditis elegans*. *Genetics* **120**, 987–1001.
- McKim, K. S., and Rose, A. M. (1990). Chromosome I duplications in *Caenorhabditis elegans*. *Genetics* **124**, 115–132.
- Moerman, D. G., Benian, G. M., and Waterston, R. H. (1986). Molecular cloning of the muscle gene *unc-22* in *Caenorhabditis elegans* by Tc1 transposon tagging. *Proc. Natl Acad. Sci. U.S.A.* **83**, 2579–2583.
- Nehme, R., Grote, P., Tomasi, T., Loser, S., Holzkamp, H., Schnabel, R., and Conradt, B. (2010). Transcriptional upregulation of both *egl-1* BH3-only and *ced-3* caspase is required for the death of the male-specific CEM neurons. *Cell Death Differ.* **17**, 1266–1276.

- Rogalski, T. M., Moerman, D. G., and Baillie, D. L. (1982). Essential genes and deficiencies in the *unc-22* IV region of *Caenorhabditis elegans*. *Genetics* **102**, 725–736.
- Rosenbluth, R. E., and Baillie, D. L. (1981). The genetic analysis of a reciprocal translocation, *eTl*(III; V), in *Caenorhabditis elegans*. *Genetics* **99**, 415–428.
- Ruvkun, G., Ambros, V., Coulson, A., Waterston, R., Sulston, J., and Horvitz, H. R. (1989). Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14*. *Genetics* **121**, 501–516.
- Sarin, S., Bertrand, V., Bigelow, H., Boyanov, A., Doitsidou, M., Poole, R., Narula, S., and Hobert, O. (2010). Analysis of multiple EMS-mutagenized *Caenorhabditis elegans* strains by whole genome sequencing. *Genetics* **185**, 417–430.
- Seidel, H. S., Rockman, M. V., and Kruglyak, L. (2008). Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science* **319**, 589–594.
- Shelton, C. A. (2006). Quantitative PCR approach to SNP detection and linkage mapping in *Caenorhabditis elegans*. *Biotechniques* **41**, 583–588.
- Sigurdson, D. C., Herman, R. K., Horton, C. A., Kari, C. K., and Pratt, S. E. (1986). An X-autosome fusion chromosome of *Caenorhabditis elegans*. *Mol. Gen. Genet.* **202**, 212–218.
- Sigurdson, D. C., Spanier, G. J., and Herman, R. K. (1984). *Caenorhabditis elegans* deficiency mapping. *Genetics* **108**, 331–345.
- Sun, A. Y., and Lambie, E. J. (1997). *gon-2*, a gene required for gonadogenesis in *Caenorhabditis elegans*. *Genetics* **147**, 1077–1089.
- Swan, K. A., Curtis, D. E., McKusick, K. B., Voinov, A. V., Mapa, F. A., and Cancilla, M. R. (2002). High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res.* **12**, 1100–1105.
- Tsang, W. Y., and Lemire, B. D. (2002). Stable heteroplasmy but differential inheritance of a large mitochondrial DNA deletion in nematodes. *Biochem. Cell. Biol.* **80**, 645–654.
- Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H., and Plasterk, R. H. (2001). Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* **28**, 160–164.
- Williams, B. D. (1995). Genetic mapping with polymorphic sequence-tagged sites. *Methods Cell Biol.* **48**, 81–96.
- Williams, B. D., Schrank, B., Huynh, C., Shownkeen, R., and Waterston, R. H. (1992). A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics* **131**, 609–624.
- Ye, S., Dhillon, S., Ke, X., Collins, A. R., and Day, I. N. (2001). An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* **29**, E88–E98.

---

---

## CHAPTER 2

# Specialized Chromosomes and Their Uses in *Caenorhabditis elegans*

**Martin R. Jones, Zoe Lohn and Ann M. Rose**

Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada

---

Abstract

I. Introduction

A. Nomenclature for Chromosomal Rearrangements

II. Chromosomal Rearrangements

A. Translocations

B. Inversions

C. Deletions (Deficiencies)

D. Duplications

III. Characterization of Chromosomal Rearrangements

A. Restriction Enzyme Digestion of Single Nucleotide Polymorphisms (snip-SNP)

B. Oligonucleotide Array Comparative Genome Hybridization (aCGH)

C. Inverse PCR

D. Whole Genome Sequencing

IV. Engineered Constructs: Transgenic Arrays

A. Experimental Applications of Transgenic Arrays

B. Methods for Construction of Transgenic Arrays

C. Properties for Construction of Transgenic Arrays

D. Methods for Creating Transgenic Animals

E. Use of Transgenic Arrays as Balancers

Acknowledgments

References

Glossary

---

---

### Abstract

Research on *Caenorhabditis elegans* involves the use of a wide range of genetic and molecular tools consisting of chromosomal material captured and modified for specific purposes. These “specialized chromosomes” come in many forms ranging from relatively simple gene deletions to complex rearrangements involving endogenous

chromosomes as well as transgenic constructs. In this chapter, we describe the specialized chromosomes that are available in *C. elegans*, their origins, practical considerations, and methods for generation and evaluation. We will summarize their uses for biological studies, and their contribution to our knowledge about chromosome biology.

---

---

---

## I. Introduction

*Caenorhabditis elegans* researchers have at their disposal a large collection of mutant strains and specialized chromosomes, constituting an extensive genetic toolkit for genetic analyses. Specialized chromosomes are chromosome rearrangements that have either been recovered in genetic screens or specifically constructed and adapted for research purposes. They range in size from large reciprocal translocations to single-gene deletions and in structure from major chromosomal rearrangements to transgenic arrays. They may be inherited either as chromosomal insertions or extra-chromosomal fragments. The collection includes a variety of chromosomal rearrangements, several types of transgenic arrays containing plasmid, cosmid, fosmid, or reporter fusion constructs, inserted marker sequences, and naturally occurring variants. Modifying chromosomes in order to address biological questions or obtain tools for better research methodology is standard practice for *C. elegans* researchers, who have adopted specialized chromosomes for a variety of purposes, including maintenance of mutations, mapping, investigating gene expression, and the study of specialized processes such as meiosis.

Traditionally, the best known and most commonly used type of rearrangement is generically called a balancer, a term inherited from *Drosophila* genetics where chromosomal rearrangements have been used to maintain lethal mutations (Muller, 1918). Lethal mutations are so-called because they cannot be propagated as homozygotes. The category includes those that arrest development as embryos, larvae, sterile adults, or that produce progeny that cannot be maintained over successive generations. To maintain these mutants, they must be kept as heterozygous strains. The wild-type copy of the gene may be present on the homolog or provided by gene duplication.

Balancers can be applied to a variety of tasks including strain construction, maintaining existing mutations, and screening for new mutations. The best balancers have reduced crossover frequencies in the balanced region, phenotypes that are distinguishable as heterozygotes and are genotypically stable with low spontaneous mutation frequencies.

It is possible to balance (maintain a genotype) using tightly linked markers (Rose and Baillie, 1979) or by selecting for a dominant phenotype of the heterozygote (Moerman and Baillie, 1979, 1981; Rogalski *et al.*, 1982; Rogalski and Riddle, 1988) but in general chromosomal rearrangements are easier to use and more effective (Herman, 1978). Chromosomal rearrangements that can be used as balancers are generally of two types. A major class comprises those that reduce or eliminate recombination between a mutation-bearing chromosome and its wild-type homolog. These include translocations, inversions, and some deletions. A second

type of balancer provides a wild-type allele by means of an extrachromosomal or integrated segment of DNA. This group includes duplications and transgenic arrays.

It is useful to have marked balancers expressing a recognizable phenotype (Tables I–III). In some cases, the rearranged chromosome is lethal and the heterozygote the only viable phenotype, which can be especially useful for growing large numbers of animals for experimental purposes (Jones *et al.*, 2007). Equally valuable for strain maintenance are dominant markers. An excellent example of a dominant marker is an insertion of a transgene that express GFP, such as *myo-2::GFP*, which expresses GFP in the pharynx. Popular examples of balancers that have insertions of *myo-2::GFP* are the inversion, *mIn1* and the translocations, *nT1* and *hT2* (Table I).

### A. Nomenclature for Chromosomal Rearrangements

*C. elegans* is a self-fertilizing hermaphrodite that produces both sperm and oocytes. Males, which are XO (no Y chromosome), occur spontaneously as a result of X-chromosome nondisjunction or loss (Brenner, 1974). The diploid genome consists of six chromosomes, five autosomes (I–V), and a sex chromosome (X). The genome is approximately 100 Mb in size and contains approximately 20,000 genes. To facilitate communication, *C. elegans* researchers have adopted a consistent nomenclature (Horvitz *et al.*, 1979). Genotypes are written as three or four italicized (or underlined) lower case letters, for example, *dpy*, refers to the wild-type DNA sequence of the gene. Protein products encoded by the genes are written in upper case, for example, DPY. Phenotypes are written nonitalicized with an upper case letter, for example, Dpy or written out in full, for example, dumpy, refers to the phenotypic consequence of a mutation. The category of Dpy may be used for any animal that has a short fat body and includes mutants in at least 30 *dpy* genes (different DNA coding sequences). Each gene is distinguished by an Arabic number separated from the general name by a hyphen, for example, *dpy-5* and *dpy-11* are different genes. Gene names may also include a Roman numeral to indicate the chromosome in which that gene maps (I, II, III, IV, V, X, as well as f for “free” or unattached to a chromosome), for example, *dpy-5* (I) indicates that *dpy-5* is on chromosome I. Specific mutations in the gene (alleles) are referred to by a lowercase letter and Arabic number, for example, *dpy-5(e61)*. The letter identifies the laboratory in which the mutation originated, for example, “e” is the laboratory in which Sydney Brenner did his initial research on *C. elegans*. Each individual strain carrying one or more mutations is identified by uppercase letters and Arabic numbers, for example, CB61 is the original strain carrying *dpy-5(e61)*. The laboratory in which a strain was generated is identified by an uppercase letter combination, for example, CB for Cambridge. However, whether or not a particular allele is dominant or recessive, temperature sensitive, etc. is not indicated in the allele name and requires further notation. When a strain is moved into a different genetic background (either by outcrossing or mutagenesis), the new strain gets a new strain name. A full list of laboratory designations is available in WormBase ([www.wormbase.org](http://www.wormbase.org)).

**Table I**Marked balancers commonly used in *C. elegans*

Balancer	Region covered	Marker	Phenotype
<b>Balancers with a dominant GFP marker</b>			
<i>hT2 [bli-4 (e937)]</i>	I <sup>L</sup> III <sup>R</sup>	<i>[myo-2::GFP]</i>	Pharyngeal GFP
<i>mIn1 [mIs14]</i> (Edgley and Riddle, 2001)	II <sup>C</sup>	<i>[myo-2::GFP]</i>	Pharyngeal GFP
<i>nT1[qIs51]</i> (Belfiore <i>et al.</i> , 2002; Cui <i>et al.</i> , 2006)	IV <sup>R</sup> V <sup>L</sup>	<i>[myo-2::GFP]</i> <i>[pes-10::GFP]</i> <i>[F22B7.9::GFP]</i>	Pharyngeal GFP Embryonic GFP Intestinal GFP
<b>Balancers with a recessive lethal marker</b>			
<i>eT1</i>	III <sup>L</sup> V <sup>R</sup>	<i>[let-x (s2165)]</i>	Larval arrest
<i>hIn1</i> (Zetka and Rose, 1992)	I <sup>R</sup>	<i>[unc-75 (h1041)]</i> <i>[unc-75 (h1042)]</i>	Larval arrest
<i>hT2 [bli-4 (e937)] [q782]</i> (Jones <i>et al.</i> , 2007a)	I <sup>L</sup> III <sup>R</sup>	<i>[let-x (q782)]</i>	Larval arrest
<i>mIn1</i> (Edgley and Riddle, 2001)	II <sup>C</sup>	<i>[let-552 (e2542)]</i>	Larval arrest
<i>nT1</i>	IV <sup>R</sup> V <sup>L</sup>	<i>[let-x (qIs50)]</i>	Larval arrest
<b>Balancers with a recessive visible marker</b>			
<i>eDp6</i> (Hodgkin, 1980)	III <sup>L</sup>	<i>[unc-119]</i>	Uncoordinated
<i>eT1</i> (Rosenbluth and Baillie, 1981)	III <sup>R</sup> V <sup>L</sup>	<i>[unc-46 (e177)]</i> <i>[bli-5 (s277)]</i> <i>[dpy-11 (s287)]</i> <i>[dpy-18 (e364)]</i> <i>[sma-2 (s262)]</i> <i>[sma-3 (e491)]</i>	Uncoordinated Blistered Dumpy Small
<i>hIn1</i>	I <sup>R</sup>	<i>[unc-54 (h1040)]</i> <i>[unc-101 (sy241)]</i> <i>[unc-75 (h1041)]</i>	Uncoordinated
<i>hT1</i> (McKim <i>et al.</i> , 1988a)	I <sup>L</sup> V <sup>L</sup>	<i>[unc-29 (e403)]</i>	Uncoordinated
<i>hT2 [bli-4 (e937)]</i> (McKim <i>et al.</i> , 1988a, 1993)	I <sup>L</sup> III <sup>R</sup>	<i>[dpy-5 (h659)]</i> <i>[dpy-18 (h662)]</i> <i>[unc-54 (e190)]</i> <i>[unc-29 (h1011)]</i> <i>[unc-59 (e261)]</i>	Dumpy Uncoordinated
<i>hT3</i> (McKim <i>et al.</i> , 1993)	I <sup>L</sup> X <sup>R</sup>	<i>[bli-4 (e937)]</i> <i>[dpy-5 (e61)]</i>	Blistered Dumpy
<i>mIn1</i> (Edgley and Riddle, 2001)	II <sup>C</sup>	<i>[unc-4 (e120)]</i> <i>[dpy-10 (e128)]</i> <i>[rol-1 (e91)]</i>	Uncoordinated Dumpy Roller
<i>mT1</i> (Edgley and Riddle, 2001)	II <sup>R</sup> III <sup>R</sup>	<i>[dpy-10 (e128)]</i>	Dumpy
<i>nT1 [unc-x (n754)]</i> (Ferguson and Horvitz, 1985)	IV <sup>R</sup> V <sup>L</sup>	<i>[unc-x (n754)]</i>	Uncoordinated
<i>qC1</i> (Austin and Kimble, 1989)	III <sup>L</sup>	<i>[dpy-19 (e1259)]</i> <i>[glp-1 (q339)]</i>	Dumpy, Germ Line Proliferation Abnormal
<i>sC1</i> [D. Baillie, pers. comm.]	III <sup>L</sup>	<i>[dpy-1 (s2170)]</i>	Dumpy
<i>sC4</i> [D. Baillie, pers. comm.]	V <sup>R</sup>	<i>[dpy-28 (e428)]</i> <i>[dpy-21 (e428)]</i>	Dumpy
<i>szT1 [lon-2 (e678)]</i> (McKim <i>et al.</i> , 1992)	I <sup>L</sup> X <sup>R</sup>	<i>[lon-2 (e678)]</i> <i>[unc-29 (e403)]</i>	Long Uncoordinated

Note: <sup>L</sup> left; <sup>R</sup> right; <sup>C</sup> center. In *C. elegans*, details of the generation, types, and applications of chromosomal rearrangements, which allow lethal mutations to be maintained stably over generations, have been reviewed in the WormBook, the online review of *C. elegans* biology ([www.wormbook.org](http://www.wormbook.org)).

**Table II**Fully characterized balancers that are in common use in *C. elegans*

Balancer	Region covered	Reference strain	Use	Stability	Origin	Derivatives
<i>eDp6[unc-119] (III:f)</i> (Hodgkin, 1980)	<i>tra-1</i> and <i>vab-7</i>	CB1517	Screens (Hodgkin, 1986; Hunter and Wood, 1990)		Acetaldehyde	<i>ctDp2 (III:f)</i>
<i>eT1[unc-36] (III:V)</i> (Rosenbluth and Baillie, 1981; Rosenbluth <i>et al.</i> , 1985)	Left LG V through <i>unc-23</i> and right LG III to <i>unc-36</i>	BC2200	Strain construction screens (Adames <i>et al.</i> , 1998; Johnsen and Baillie, 1991; Stewart <i>et al.</i> , 1991)	Extremely stable	32P (120 J/m <sup>2</sup> )	
<i>hIn1 (I)</i> (Zetka and Rose, 1992)	<i>unc-75</i> to <i>unc-54</i>	KR2267	Strain construction screens (Lee <i>et al.</i> , 1994)	Very stable	1500 R gamma	<i>hDf11</i> <i>hDf12</i> <i>hDp131</i> <i>hDp132</i>
<i>hT1 [let-x] (I:V)</i> (McKim <i>et al.</i> , 1988a)	Left LG I through <i>let-80</i> and left LG V through <i>dpy-11</i>		Screens (Howell and Rose, 1990)	Very stable	Gamma irradiation	
<i>hT2 [bit-4] (I:III)</i> (McKim <i>et al.</i> , 1988a, 1993)	Left LG I through to between <i>unc-101</i> and <i>unc-59</i> , right LG III through to <i>dpy-17</i>		Strain maintenance		1500 R gamma irradiation	<i>hDp134</i> (I:III:f)
<i>hT3 [unc-29] (I:X)</i> (McKim <i>et al.</i> , 1993)	Left LG I through to <i>let-363</i> and right LG X through to between <i>dpy-7</i> and <i>unc-3</i>			Very stable	1500–3000 R gamma irradiation	<i>hDp135</i> (I:X:f)
<i>mIn1 (II)</i> (Edgley and Riddle, 2001)	<i>lin-31</i> through <i>rol-1</i>	DR1753	Strain construction	Very stable	1500 R (450 R/min) Gamma irradiation	<i>GFP</i>
<i>mT1 (II:III)</i> (Edgley and Riddle, 2001)	Right LG II through to between <i>bit-2</i> and <i>dpy-10</i> , and the right of LG III through to between <i>daf-2</i> and <i>unc-93</i>	DR1753		Very stable	1500 R (450 R/min) gamma irradiation	

(Continued)

Table II (Continued)

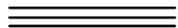
Balancer	Region covered	Reference strain	Use	Stability	Origin	Derivatives
<i>mnC1 (II)</i> (Edgley and Riddle, 2001)	Right LG II from approximately <i>dpy-10</i> through <i>unc-52</i>		Strain construction screens (Sigurdson <i>et al.</i> , 1984)	Very stable	7500 R (450 R/min) X-ray mutagenesis	<i>mnC1</i>
<i>mndp1 (X;V)</i> (Herman <i>et al.</i> , 1976, 1979)	Right LG X through <i>let-4</i>		Strain construction screens (Meneely and Herman, 1981)	Stable	~7000–8000 R X-ray mutagenesis	<i>ctDp6 (III;f)</i>
<i>mndp2 (X;f)</i> (Herman <i>et al.</i> , 1979)	Right LG X through <i>unc-3</i>		Strain construction	Stable	7000–8000 R X-ray mutagenesis	
<i>mndp3 (X;f)</i> (Herman <i>et al.</i> , 1976, 1979)	Right LG X through <i>unc-9</i>		Strain construction	Stable	7000–8000 R X-ray mutagenesis	
<i>mndp10 (X;f)</i> (Herman <i>et al.</i> , 1979; Meneely and Herman, 1981)	Right LG X through <i>lin-2</i>	SP117	Strain construction		EMS and 5400 R X-ray mutagenesis	
<i>mndp33 (X;IV)</i> (Herman <i>et al.</i> , 1979)	Left LG X from <i>lin-18</i> to <i>osm-5</i>	SP309	Strain construction		7000–8000 R X-ray mutagenesis	<i>mndp35 (II)</i>
<i>mndp34 (II)</i> (Herman <i>et al.</i> , 1979)	Right LG II from <i>unc-52</i> to <i>unc-53</i>	SP306	Strain maintenance		7000–8000 R X-ray mutagenesis	<i>mndp36 (II)</i>
<i>nT1 (IV;V)</i> (Ferguson and Horvitz, 1985)	Right LG IV through <i>lin-1</i> and left LG V through <i>unc-76</i>		Screens (Clark <i>et al.</i> , 1988; Rogalski and Riddle, 1988)	Very stable	Spontaneous	
<i>qC1 (III)</i> (Austin and Kimble, 1989)	Left LG III from <i>tra-1</i> to <i>dpy-1</i>			Very stable	7200 R ( $1.29 \times 10^3$ R/min) gamma irradiation	<i>hDp3</i> <i>hDp30</i> <i>hDp59</i> <i>hDp74-</i> <i>hDp77</i>
<i>sDp2 [unc-13] (I;f)</i> (McKim <i>et al.</i> , 1992, 1993; Rose <i>et al.</i> , 1984)	Left LG I through to <i>unc-15</i>		Screens (Clark <i>et al.</i> , 1988; Howell and Rose, 1990; McKim and Rose, 1990)		7500 R gamma irradiation mutagenesis	
<i>sDp3 [let-x] (III;f)</i> (Rosenbluth <i>et al.</i> , 1985)	Left LG III from <i>unc-86</i> through to <i>dpy-1</i>				1500 R gamma irradiation mutagenesis	
<i>stDp2 [let-x] (X;II)</i> (Meneely and Wood, 1984)	Center LG X from <i>unc-58</i> to <i>unc-6</i>					

<i>szDp1</i> [ <i>let-x</i> ] ( <i>I;X;f</i> ) (McKim <i>et al.</i> , 1988a, 1992, 1993)	Left LG I through to <i>unc-13</i> and one half translocation [ <i>szTI(X)</i> ]	Strain construction screens (Howell and Rose, 1990; McKim and Rose, 1990)	Very stable	Exceptional Unc-3 segregant from <i>dpy-5/szTI</i> [ <i>lon-2</i> ] <i>I;unc-3/szTI X</i> .	<i>hDp31- hDp58 hDp60- hDp73 ILXLszT</i>
<i>szTI</i> [ <i>let-x</i> ] (Fodor and Deák, 1985; McKim <i>et al.</i> , 1988a)	Left LG I through to <i>unc-13</i> and right LG X through to <i>dpy-3</i>	Screens (Plenefisch <i>et al.</i> , 1989)	Very stable	7000 R (1000 R/min) X-ray mutagenesis	
<i>yDp1</i> [ <i>let-x</i> ] ( <i>II;V;f</i> ) (DeLong <i>et al.</i> , 1987)	Right LG IV from <i>dpy-4</i> through to <i>unc-22</i> and left LG V from <i>unc-34</i> through to <i>unc-46</i>	mosaic analysis (Yuan and Horvitz, 1990)		EMS mutagenesis	<i>nDp3</i>

Note: Roentgens (R), a form of irradiation. EMS, ethyl methanesulfate.

Specialized chromosomes also have an accepted nomenclature. Df for deficiencies or deletions that are known to affect multiple genes, Dp for duplications, T for translocations, In for inversions, and C for crossover suppressors (of unknown structure). Extrachromosomal transgenic arrays are designated Ex, while transgenic constructs that are inserted into a chromosome are designated Is. Transgenes integrated by a recently developed transposon-based method, MosSCI (Frokjaer-Jensen *et al.*, 2008), are designated by the two letters *Si*. Full names include the laboratory allele prefix, the two letters, and a number. For example, a MosSCI insert from the Jorgensen lab (*ox*) will be named *oxSi31*. As for alleles, the name is preceded by the small letter italic designation of the originating laboratory. If the SC carries a mutation in a known gene, the gene is described using square brackets following the SC name, for example, the reciprocal translocation *eTI*, which has a breakpoint in *unc-36* (III), is *eTI[unc-36]* (III;V). The components of a reciprocal translocation can be identified by their pairing properties. Chromosome identity in *C. elegans* is defined by the end of the autosome that pairs with its homolog and that contains a pairing center, also known as the homolog recognition region (HRR) (McKim *et al.*, 1988b). Thus, *eTI* (III) pairs with the normal chromosome III, and *eTI* (V) with the normal chromosome V. In some cases, half-translocations can be maintained as extrachromosomal duplications and may in these cases be given a *Dp* designation. For example, a half-translocation from the reciprocal translocation *hTI*(I;V) that was isolated in the Rose laboratory at the University of British Columbia (*h*), recombines and segregates from chromosome I. It can be maintained as a duplication of regions of chromosomes I and V in addition to the normal diploid complement and has been designated *hDp133*, rather than the precise but more cumbersome *hTI*(I<sup>R</sup>V<sup>L</sup>).

These standardized nomenclature guidelines have been modified and adopted to describe genetic constructs marked with a reporter gene; the most commonly used is the green fluorescent protein (GFP) (Chalfie *et al.*, 1994). In these cases, double colons are used to indicate the covalent linkage to the reporter, for example, a promoter engineered to connect 5' to a GFP is written promoter::GFP.



## II. Chromosomal Rearrangements

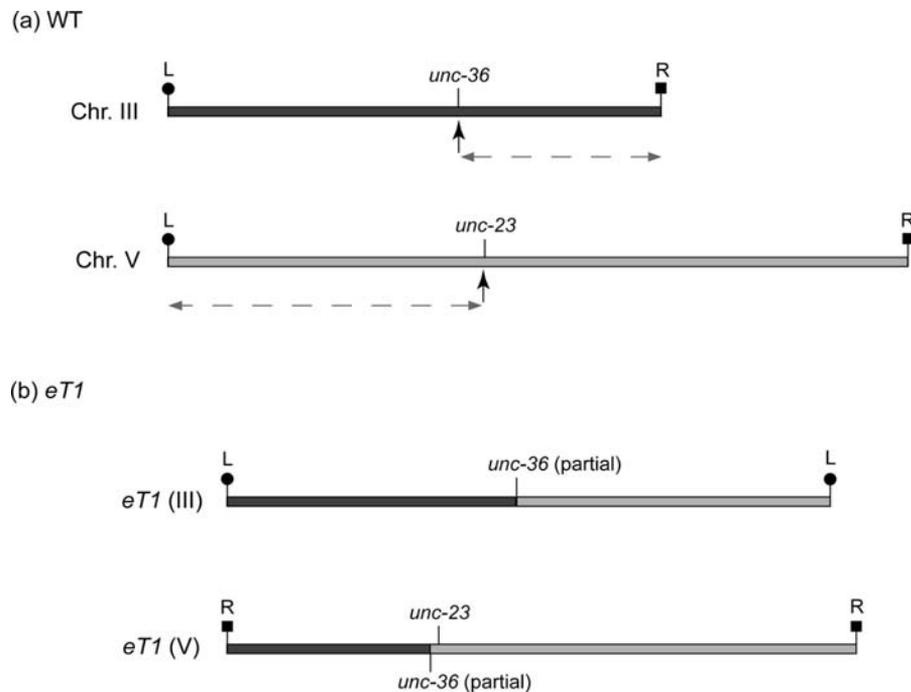
Gross rearrangements refer to chromosomes that involve an extensive alteration to the content or structure of the genome. Such rearrangements can arise naturally, although many have been engineered purposefully by the use of mutagenesis. Translocations, inversions, duplications, and deletions all fall into this category and each have specific properties that make them useful for a variety of housekeeping tasks and experimental approaches in *C. elegans* research.

### A. Translocations

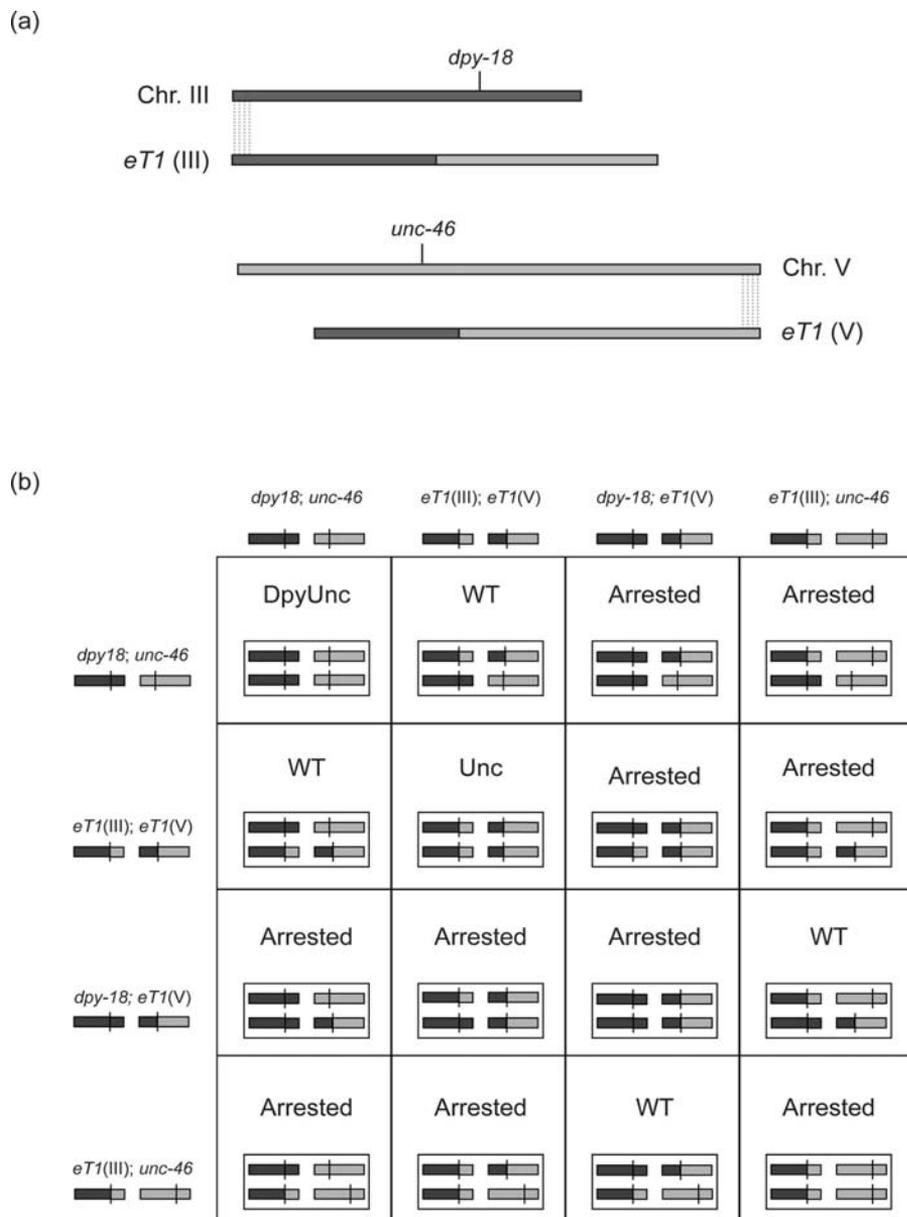
Translocations involve the displacement of a segment of DNA from one region of the genome to another. Consequently, nonhomologous DNA aligns during meiotic

pairing, resulting in a corresponding suppression of recombination in those regions. A heterozygous mutation in the recombination suppressed regions can subsequently be retained for an indefinite number of generations. Translocations are often marked with a mutation, which results in a phenotype that makes the homozygote readily observable (“Visible”), thus facilitating selection of the heterozygote. In this way, either lethal or visible mutants can be balanced and the strain can be maintained in a heritably stable manner.

The reciprocal translocation, *eT1*, illustrates the properties and uses of translocations in *C. elegans*. *eT1*, originally thought to be a mutation in an *unc* gene, was shown to have exchanged the right portion of chromosome III for the left portion of chromosome V without the loss of any essential material, and thus is homozygous viable (Rosenbluth and Baillie, 1981). The physical positions of the breakpoints in *eT1* have been characterized. On chromosome III, the break occurs within the *unc-36* gene, and on chromosome V is in cosmid H14N18 between *rol-3* and *unc-42* (Zhao *et al.*, 2006) (Fig. 1).



**Fig. 1** Structure of the reciprocal translocation *eT1*. *eT1* is a reciprocal translocation between the right end of chromosome III and the left half of chromosome V. The breakpoint on chromosome III falls within the gene *unc-36*, such that homozygous *eT1* animals display an Unc phenotype. *eT1* (III) pairs with the WT chromosome III while *eT1* (V) pairs with WT chromosome V. The recombination suppressed regions available for balancing mutations are shown by dashed lines in (a).



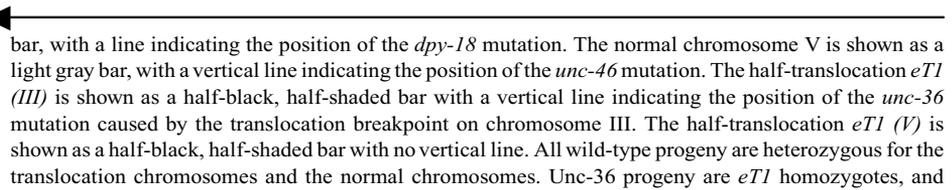
**Fig. 2** Structure of the *dpy-18/eT1* III; *unc-46/eT1* V chromosomes showing the genotypes and phenotypes of the self-progeny. (a) *dpy-18* (III) and *unc-46* (V) mutations are maintained heterozygously by *eT1*. *eT1* (III) segregates from chromosome III and *eT1* (V) segregates from chromosome V. The initial pairing regions are represented by gray dashed lines. (b) Punnet square showing resulting progeny. Phenotypes are indicated for each progeny class. The normal chromosome III is shown as a dark gray

## 1. Translocations for the Maintenance of Lethal Mutations

As heterozygotes, crossing over is eliminated within the regions no longer *cis*-lined to an HRR, making translocations valuable tools for the isolation and maintenance of lethal mutations (Clark *et al.*, 1990; Clark *et al.*, 1988; Ferguson and Horvitz, 1985; Johnsen and Baillie, 1991; Rogalski and Riddle, 1988; Rosenbluth *et al.*, 1990; Zhao *et al.*, 2006). In the case of *eTI*, the translocated regions cover a large fraction (one-sixth) of the genome and thus balance a very large number of genes. Fortunately, in the case of *eTI*, the breakpoint on chromosome III provided a viable visible marker, *unc-36*. In order to easily distinguish the heterozygotes, morphological mutations in the recombination-suppressed region of the normal homologs were used. A frequently used strain of *eTI[unc-36]* (III:V) has *dpy-18* (*e364*) (III) and *unc-46*(*e177*) (V) mutations on the normal homologs. The heterozygote has a wild-type phenotype and can be easily distinguished from either homozygote. *eTI* heterozygotes segregate: (1) *eTI* homozygotes (which have an Unc-36 phenotype because the translocation breakpoint lies in the *unc-36* gene on chromosome III); (2) Dpy-18 Unc-46 homozygotes; and (3) a large fraction of aneuploid progeny (10/16) that arrest development as embryos or larvae as a consequence of abnormal chromosome numbers (Adames *et al.*, 1998; Rosenbluth and Baillie, 1981) (Fig. 2).

## 2. Translocations for the Study of Meiotic Processes

It is not immediately obvious why translocations in *C. elegans* suppress crossing over. If they paired as described in textbooks, forming a cruxiform structure, only limited crossover disruption would be expected. Thus, based on the genetic evidence, it was proposed that in animals heterozygous for a reciprocal translocation, the “translocated” arms did not pair, recombine, or disjoin (Rosenbluth and Baillie, 1981). Subsequent molecular analysis confirmed the predictions (Adames *et al.*, 1998), which are shown in Fig. 2. An examination of all available translocations leads to the proposal that one end of each chromosome contained a region that facilitated pairing between homologs and initiated the meiotic cascade of pairing, recombination, and disjunction. This region was designated as the HRR. Subsequently, the regions were molecularly identified (Phillips *et al.*, 2009).



bar, with a line indicating the position of the *dpy-18* mutation. The normal chromosome V is shown as a light gray bar, with a vertical line indicating the position of the *unc-46* mutation. The half-translocation *eTI* (III) is shown as a half-black, half-shaded bar with a vertical line indicating the position of the *unc-36* mutation caused by the translocation breakpoint on chromosome III. The half-translocation *eTI* (V) is shown as a half-black, half-shaded bar with no vertical line. All wild-type progeny are heterozygous for the translocation chromosomes and the normal chromosomes. Unc-36 progeny are *eTI* homozygotes, and Dpy-18 Unc-46 progeny are homozygous for the normal chromosomes. Aneuploid progeny account for 10/16ths of the total progeny.

(Adapted from Edgley and Riddle (2001)).

A nonreciprocal type of translocation is a chromosome fusion. In a screen for genes required for germ line immortality, the *mrt-2* locus was discovered (Ahmed and Hodgkin, 2000). *mrt-2* encodes an ortholog of a checkpoint gene and in *C. elegans* is required to prevent telomere shortening leading to end-to-end chromosome fusions. The meiotic behavior of heterozygous fusion chromosomes has been studied in the context of crossover control and it has been found that fusions containing two or even three whole chromosomes behave as a single chromosome in regard to having mainly only one crossover per meiosis, similar to wild-type (Hillers and Villeneuve, 2003). X-autosome fusions have been used to investigate dosage compensation and the relationship between compensations and crossover control (Gladden *et al.*, 2007; Gladden and Meyer, 2007; Lieb *et al.*, 2000).

### 3. Translocations to Determine Forward Mutation Frequency

The number of lethal events recovered using a translocation heterozygote under defined conditions can be used to calculate forward mutation frequencies for different doses of specific mutagens, such as ethylmethane sulfonate (EMS) (Raja E. Rosenbluth *et al.*, 1983), ionizing radiation (Rosenbluth *et al.*, 1985), ultraviolet radiation (Stewart *et al.*, 1991), formaldehyde (Johnsen and Baillie, 1988), transposable element mobilization (Clark *et al.*, 1990), exposure to radiation on the International Space Station (Zhao *et al.*, 2006), and mutator strains such as the deleter of G's, *dog-1* (Zhao *et al.*, 2008). *eT1* has been used extensively for these purposes, reinforcing its advantages, such as ease of use, size of region balanced, and stability of genotype.

### 4. Advantages and Disadvantages

In general, translocations make excellent balancers (Table III). They are heritably stable and relatively easy to use. However, it is important to have some knowledge of their structure, complexity, and segregational properties in order to use them effectively. Reciprocal translocations balance regions on two different chromosomes, a feature that may not be as desirable as balancing a single chromosome. In addition, the large fraction of aneuploid self-progeny segregating from reciprocal translocation heterozygotes can complicate recognition of a lethal phenotype segregating with the normal homologs.

## B. Inversions

Inversions are chromosomal rearrangements in which a segment of the chromosome reinserts in the same location but in the reverse orientation. Although crossing over can occur in the inverted region, recombination between the breakpoints is almost completely suppressed in heterozygotes due to topological restraints, absence of DNA alignment, or disruption of sites required for pairing. When recombination

**Table III**A listing of genomic balancers that are available for use in *C. elegans* but are not fully characterized

Balancer name	Region covered	Use	Stability	Origin
<i>hDp102 (I;X)</i> (McKim <i>et al.</i> , 1993)	Duplication of chromosome I ( <i>unc-40</i> to just past <i>unc-9</i> ) inserted between <i>unc-7</i> and <i>dpy-3</i> on the X chromosome		Very stable	1500–3000 R gamma irradiation
<i>hDp14 (I;X)</i> (McKim and Rose, 1990; McKim <i>et al.</i> , 1993)	Duplication of chromosome I inserted between <i>unc-2</i> and <i>dpy-8</i> on the X chromosome		Very stable	1500–3000 R (0.9–7.5 R/s) gamma irradiation
<i>sC1 (III)</i> [D. Baillie, pers. comm.]	LG III from <i>unc-45</i> to $\sim$ <i>daf-2</i>		Very stable	2000 R gamma irradiation mutagenesis
<i>sC4 (V)</i> [D. Baillie pers. comm.]	Right LG V from <i>rol-9</i> to <i>unc-76</i>			2000 R gamma irradiation mutagenesis
<i>sDp8</i> (Stewart <i>et al.</i> , 1991)	Covers <i>unc-36</i> but does not cover <i>sma-2</i>		Stable	UV mutagenesis (120 J/m <sup>2</sup> )
<i>sDp9</i> (Stewart <i>et al.</i> , 1991)	Covers <i>unc-36</i> and <i>sma-2</i>		Stable	UV mutagenesis (120 J/m <sup>2</sup> )
<i>sDp30 (V;X)</i> (McKim <i>et al.</i> , 1993)	Inserted between <i>dpy-7</i> and <i>unc-3</i> on chromosome V		Stable	
<i>sT1 (III;X)</i> [D. Baillie, pers. comm.]	Possible translocation	Unconfirmed Possibly Reduces crossing over on LG III between <i>sma-2</i> and <i>unc-64</i>		1500 R gamma irradiation mutagenesis
<i>sT2 (IV;V)</i> [D. Baillie, pers. comm.]	Possible translocation; covers <i>unc-46</i>	Unconfirmed		1500 R gamma irradiation mutagenesis
<i>sT3 (III;V)</i> [D. Baillie pers. comm.]	Possible translocation; putatively covers <i>unc-36</i> , <i>sma-2</i> and <i>unc-46</i>	Unconfirmed		1500 R gamma irradiation mutagenesis
<i>sT4 (III; likely V)</i> [D. Baillie pers. comm.]	Possible translocation; left LG III to between <i>unc-36</i> and <i>sma-2</i> and covers <i>unc-46</i> on LGV	Unconfirmed		1500 R gamma irradiation mutagenesis
<i>sT5 (II;III)</i> (Stewart <i>et al.</i> , 1991) [D. Baillie pers. comm.]	Possible translocation	Unconfirmed Reduces crossing over on LG III between <i>sma-2</i> and <i>unc-64</i>		120 J/m <sup>2</sup> UV irradiation mutagenesis

does occur, crossover products generate duplications and deletions, and often result in inviable animals (Zetka and Rose, 1992). As homozygotes, inversions are generally viable, pair properly, and exhibit normal levels of recombination. In *C. elegans*, two inversions have been well characterized. *hln1*(I) (Zetka and Rose, 1992) and *mIn1* (II) (Edgley and Riddle, 2001). In *hln1*, a large portion of the right half of

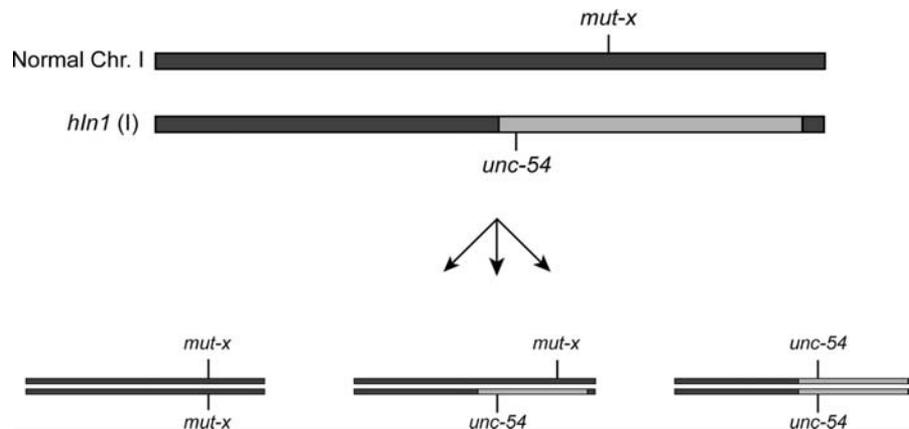
chromosome I is inverted, and in *mIn1*, a large section of the center of chromosome II from *lin-31* to *rol-1* is inverted.

## 1. Uses of Inversions

Inversions have been used effectively both as balancers (Fig. 3) and for capturing novel mutations. Lethal mutations, along with more complex rearrangements such as deletions, have been recovered with both *hIn1* and *mIn1* (Bosher *et al.*, 2003; Edgley and Riddle, 2001; Zetka and Rose, 1992). Similar to other balancers, inversions may be more effective when marked with either a lethal or visible mutation. In the case of *hIn1*, mutations in *unc-75* and *unc-54* were induced on the inverted chromosome (Zetka and Rose, 1992). Eight variant strains of *mIn1* have been isolated, an unmarked form, ones that carry recessive morphological or lethal markers, and one that carries an integrated transgene that confers a semidominant GFP phenotype, making *mIn1* useful for a wide variety of applications (M. Edgley, pers. comm.).

## 2. Advantages and Disadvantages

An advantage in the use of inversions is that the genomic disruption is limited to one chromosome and as a result inversions tend to be stable. Any recombination events that occur with the wild-type chromosome are detrimental to the organism, and are thus eliminated from the population. A disadvantage can be that inversions,



**Fig. 3** Schematic showing segregation from an inversion. In this example, the inversion *hIn1* is balancing a mutation (*mut-x*). From the self-progeny of this strain  $\frac{1}{4}$  will display the homozygous inversion phenotype (*unc-54*),  $\frac{1}{2}$  will be the balanced heterozygous animals, and  $\frac{1}{4}$  will be homozygous for the balanced mutation and will display the *mut-x* phenotype. The strain can be successfully maintained by selecting the WT animals in each generation.

although stable, present some technical difficulties in their characterization. Genetically, they generate rearranged linkage between markers. Molecularly, detection of the two breakpoints is needed to provide information about the rearrangement. Methods for specifically isolating and characterizing inversions have not been well developed and this in part explains the paucity of available ones for use.

### C. Deletions (Deficiencies)

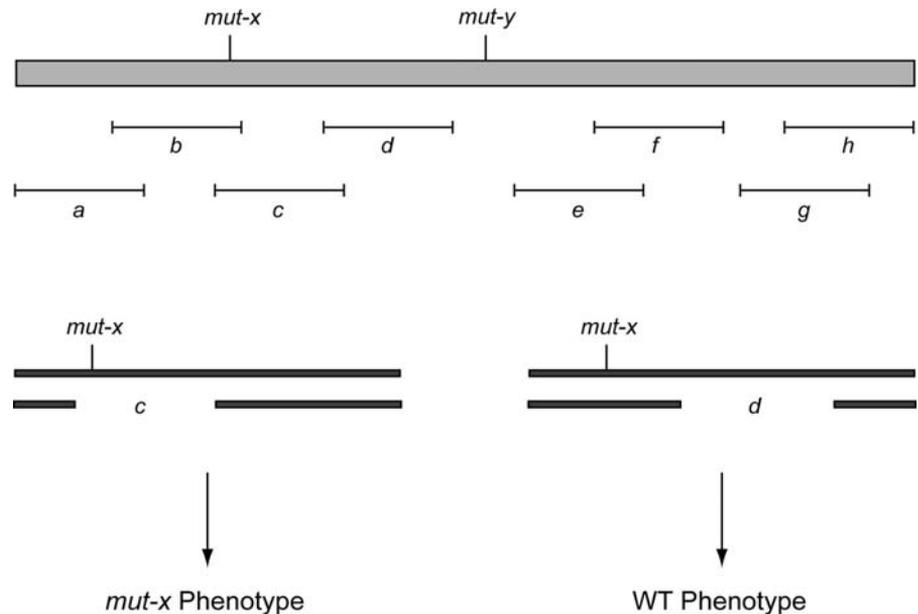
Chromosomes that have lost a region of DNA are often used in *C. elegans* research. These chromosomes are known as deletions, a term that normally refers to the DNA that is missing (deleted) or deficiencies, referring to the failure to complement two or more adjacent genetic loci (phenotypic deficiency). The nomenclature for both is *Df*. *Dfs* range in size from small deletions that disrupt a few genes to those that remove megabases (Mb) of sequence involving hundreds of genes. They can arise spontaneously but are more usually generated in mutagenesis screens. Deletions are one of the most useful mapping tools and as such a large resource of *Df* strains covering over 70% of the genome has been generated (Ahnn and Fire, 1994; Chanal and Labouesse, 1997). Specific information about deletions in a defined region of the genome can be obtained from WormBase ([www.wormbase.org](http://www.wormbase.org)).

#### 1. Deletions as Mapping Tools

Probably, the most common use of deletions is for mapping uncloned mutations (Clark *et al.*, 1990; Fay, 2006; Johnsen and Baillie, 1991; Meneely and Herman, 1979; Rogalski *et al.*, 1982; Sigurdson *et al.*, 1984; Stewart *et al.*, 1998). A typical mapping scheme would involve crossing each mutant strain to a *Df* strain and assaying (scoring) for rescue (complementation) of the mutant phenotype. If the mutation is rescued by the deficiency strain, the mutation does not fall within the region covered by the deficiency. If the mutant phenotype is observed in the F1 generation, the mutation is not rescued (failure to complement) and maps within the extent of the deficiency. If deficiencies overlap, it is possible to limit the position of the mutation to a relatively small region defined by the overlap of two or more deficiencies (Fig. 4).

#### 2. Deletions for the Study of Meiotic Recombination

Animals hemizygous for portions of the X-chromosome were used to identify the left end as critical for proper X-chromosome disjunction in XX hermaphrodites (Herman *et al.*, 1982). Subsequently, deletions that affected the pairing properties of the X-chromosome identified the X-chromosome pairing center (Villeneuve, 1994). Deletions of this region that can pair, have reduced levels of crossing over (Broverman and Meneely, 1994).



**Fig. 4** Deficiency mapping. In this example, the two mutations, x and y, fall within a genomic region that is covered by the deficiencies *a–g*. Complementation mapping will reveal that *mut-x* falls within the region defined by the overlap of deficiencies *b* and *c*. In the case of *mut-y* complementation tests will show that the mutation falls within the region not covered by a deficiency. If two or more such regions are present *mut-y* will be ambiguously assigned to all of them.

On an autosome, chromosome V deletions were described as that which could pair normally and recombine on the pairing (HRR) side of the chromosome and severely reduce crossing over in the other half of the chromosome. Molecular characterization of the rearranged chromosomes using oligoarray Comparative Genomic Hybridization (aCGH) (see Section III B.) revealed that the extent of the disruption between homologs correlated with the degree of crossover suppression (Jones *et al.*, 2009). The results demonstrated that homolog lengths may contribute to the fidelity of meiotic crossing over and is consistent with the proposed existence of multiple initiation sites occurring along the chromosome that mediate crossing over, possibly by facilitating tight alignment of homologs.

### 3. Allelic Characterization

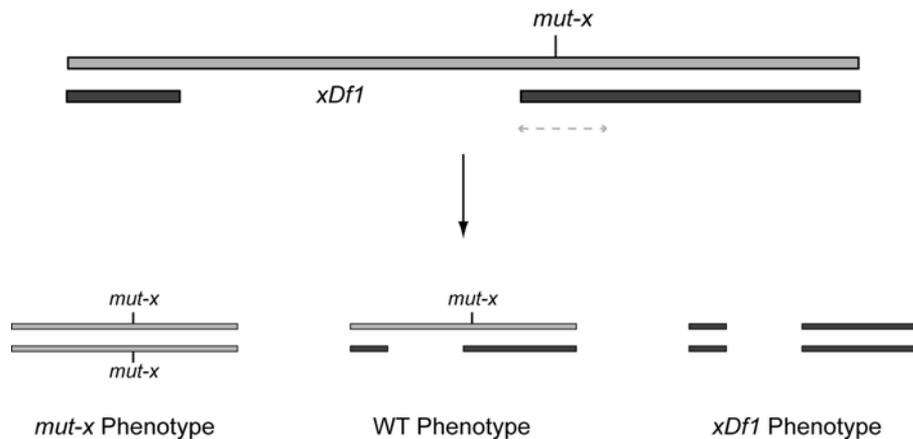
Deletions have value in the characterization of mutant alleles (Muller, 1918). Mutations that are phenotypically null are critical for interpretation of gene function, and valuable for determining the relative severity of other alleles that may alter the phenotype without eliminating it. Examination of the phenotype of a heterozygote

that contains a mutation *in trans* to the deletion allele will give an indication of the severity of the mutation. If the phenotype of the unknown mutation is unchanged it is considered to lack gene product (amorph), if more severe, it is considered to have reduced gene product (hypomorph); if less severe, to be an overproducer (hyper-morph) or to have taken on a new function (neomorph), which may not be detectable over a deletion (Muller, 1918).

A clever extension of this type of approach has been to use deletions for screening of genetic loci whose zygotic expression is required for formation of specific tissue types (Ahnn and Fire, 1994; Chanal and Labouesse, 1997; Labouesse, 1997; Terns *et al.*, 1997). In one study, 77 genetic deficiencies covering an estimated 72% of the genome were screened for staining with antibodies to body wall myosin (Ahnn and Fire, 1994) and in another 90 deficiencies covering up to 75% of the genome were screened for loci affecting hypodermal development (Chanal and Labouesse, 1997).

#### 4. Deletions as Balancers

Deletions that suppress crossing over (Broverman and Meneely, 1994; Jones *et al.*, 2009; Rosenbluth *et al.*, 1990) are potentially useful as balancers (Fig. 5). Mutations close to the deletion that fall within the suppressed region can be effectively balanced though rare recombination events can occur. As such deletions should only be used when a more stable alternative is not available and balanced mutations should be monitored regularly.



**Fig. 5** Balancing a mutation using a deficiency. In this example, a deficiency is placed heterozygously over a mutation that is outside the deletion but within the recombination suppressed region (dashed gray line). Progeny from these animals will display either the deficiency phenotype (one quarter), the *mut-x* phenotype (one quarter), or be WT in appearance (one half). Because *mut-x* is unlikely to recombine with the deficiency chromosome the heterozygous strain can be maintained by selecting the WT progeny. If recombination does occur, the breakdown of the strain can be observed by the loss of the *mut-x* phenotype in the progeny.

## 5. Single-Gene Deletions

Single-gene deletions, often referred to as gene knockouts, are valuable tools for determining gene function. Knockouts are generated using mutagens that create small deletions, most commonly trimethyl psoralen (TMP) followed by UV-radiation (UV-TMP). In *C. elegans*, there is an international effort to generate gene knockouts for the research community (Barstead and Moerman, 2006; Mitani, 2009; Moerman and Barstead, 2008). Many knockout strains are available from the Caenorhabditis Genetics Center (CGC) ([www.cbs.umn.edu/CGC/strains/](http://www.cbs.umn.edu/CGC/strains/)) and are also listed in WormBase ([www.wormbase.org](http://www.wormbase.org)). Requests to have a knockout allele generated for a specific gene should be sent to either the National Bioresource Project, Tokyo Women's Medical University School of Medicine, Japan (<http://www.shigen.nig.ac.jp/c.elegans/index.jsp>), or The Gene Knockout Project, University of British Columbia, Canada (<http://www.zoology.ubc.ca/~dgmweb/>).

## 6. Advantages and Disadvantages

Mapping using *Dfs* has the advantage of being able to position a mutation of interest to a region of the genome with higher resolution than can be done using more traditional genetic approaches such as three-factor mapping. However, deletions can be complex and not delete contiguous stretches of DNA. Without the aid of cytology to characterize deletions, most deletions have not been described molecularly reducing the mapping resolution with regard to the DNA sequence.

Many deletions disrupt at least one essential gene and are therefore lethal. In fact, they often affect more than one essential gene and are arrested during the embryo stage when homozygous. Thus, they, as for any lethal mutation, need to be maintained using a balancer. This is however normally straightforward. Single-gene deletions greatly facilitate correlating phenotype with genotype and provide a substrate for molecular PCR-based assays to follow phenotype. Mutations that are phenotypically null are critical for interpretation of gene function. Single-gene knockout are important because they are more likely to be null for gene function than point mutations that can often be hypomorphic. Experimentally, they are relatively easy to use and genes that do not display a viable visible phenotype can be followed using a simple PCR assay to determine if the strain contains the mutation either homozygously or heterozygously. A special subset of deletions reduce crossing over in adjacent regions and these can, in principal, be used to balance lethals on the homolog.

## D. Duplications

Genomic duplications have proven useful for a number of applications in *C. elegans* research. In general, two types of duplications are available, ones that are inserted into a chromosome and ones that exist as extrachromosomal “free” elements. Duplications that are inserted into the genome segregate in a Mendelian fashion and are stably maintained. Free duplications generally exist as single copies,

but strains carrying two copies have been documented (Herman *et al.*, 1976; Rogalski and Riddle, 1988). Those that are “free” elements segregate in a non-Mendelian fashion and can be lost both mitotically and meiotically during gametogenesis and somatic cell divisions (Herman *et al.*, 1976). The exceptions are those that carry the meiotic pairing regions (HRRs). Two such duplications have been characterized, *sDp1* (I;f) (Rose *et al.*, 1984) and *mDp1* (IV;f) (Rogalski and Riddle, 1988). The stability of nonpairing free duplications varies with size and other characteristics.

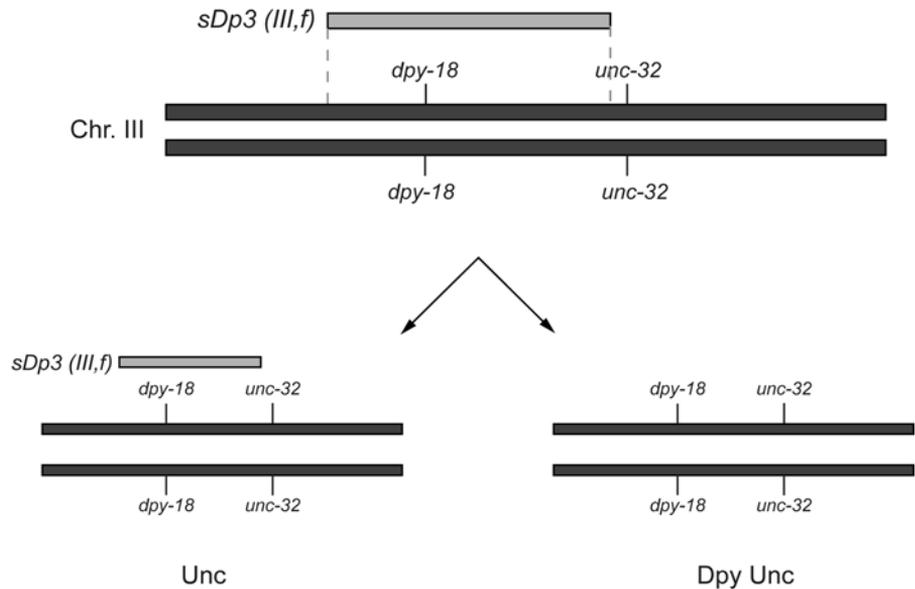
Some duplications have been recovered as exceptional segregants from translocation strains (Edgley *et al.*, 2006), or as products of a rare recombination event in an inversion strain (Zetka and Rose, 1992). Additionally, duplication derivatives produced by the spontaneous shortening of *sDp2* (I;f) have been characterized (McKim and Rose, 1990). This shortening occurred at variable rates, and produced duplications that were relatively stable in mitotic cell division (e.g., *sDp2*, *hDp5*, and *hDp20*) and duplications that were quite unstable, and continued to shorten more frequently than the stable ones (e.g., *hDp2* and *hDp23*). In the germ line, larger duplications were more stable and transmitted at a higher frequency. A more recent analysis using aCGH also found considerable variability in the length of *sDp3* (III;f) (Jones *et al.*, 2007), which was used to isolate and balance lethals on chromosome III (Stewart *et al.*, 1998). In addition to the length variability, discrete duplications internal to *sDp3* were observed, revealing the genetic complexity of some duplications, which can compromise the usefulness of duplications for accurate gene mapping.

### 1. Duplications as Balancers

Genomic duplications that do not crossover with the wild-type chromosomes make effective balancers. Mutations can be maintained as homozygotes while the duplication provides a wild-type copy of the mutated gene (Fig. 6). Unlike with translocation balancers aneuploid progeny do not result from mutations balanced in this manner. This can be advantageous for certain applications such as phenotypic characterization of lethal mutations that may be difficult to distinguish from the aneuploid progeny.

### 2. Duplications as Tools for Mosaic Analysis

Another application of duplication balancers is mosaic analysis (Hedgecock and Herman, 1995; Hunter and Wood, 1992; Yochem *et al.*, 2000). The first mosaicism studies in *C. elegans* employed X-irradiation of embryos heterozygous for the *flu-3* mutation. However, the frequency of mosaicism was very low (less than 0.1% mosaic worms) (Siddiqui and Babu, 1980). An approach that took advantage of the spontaneous loss of free duplications during somatic cell division was proposed by Herman (1989). In this approach, a free duplication that rescues a mutation will produce mosaic animals that comprise of cells retaining the duplication that will be phenotypically normal in addition to cells having lost the duplication displaying the



**Fig. 6** Schematic showing the genotype of the reference *sDp3* strain. The strain is homozygous for *dpy-18* and *unc-32* mutations and carries one copy of *sDp3*. *sDp3* provides a WT copy of *dpy-18*. Animals that retain the duplication display the *unc-32* phenotype while progeny that have lost *sDp3* are Dpy Unc.

mutant phenotype. The approach also applies to extrachromosomal SCs containing transgenic DNA sequence. In *C. elegans*, cell mosaics have been useful for examining cell autonomy and determining the timing of gene expression. For example Hunter *et al.* utilized *eDp6* (III;f) to demonstrate the cell autonomy of *tra-1*, the gene that determines somatic sex, indicating that the gene product is a cellular signaling factor (Hunter and Wood, 1990). Yuan *et al.* used the free duplication *nDp3* (IV;V;f) to show that the apoptotic genes, *ced-3* and *ced-4* function within dying cells to cause cell death (Yuan and Horvitz, 1990). Bucher *et al.* used *qDp3* (III;f) to combine lethal screens with mosaic analysis to investigate the role of essential genes in development, and found that many essential zygotic genes encoded specific developmental functions (Bucher and Greenwald, 1991).

### 3. Advantages and Disadvantages

An advantage of using a duplication to balance a mutation is that once the mutation is balanced, it can be maintained as a homozygote. This has advantages both for genetic crosses as every outcross progeny will carry the mutation, and for preparation of DNA as the yield is 2:1 from the balanced strain. The lethal homozygote can be readily observed in the balancer background without requiring additional outcrossing as for translocations. The disadvantage is that they may

spontaneously shorten and possibly rearrange, complicating their use for mapping mutations. In addition, mosaicism can complicate the interpretation of complementation experiments.

---

---

---

### III. Characterization of Chromosomal Rearrangements

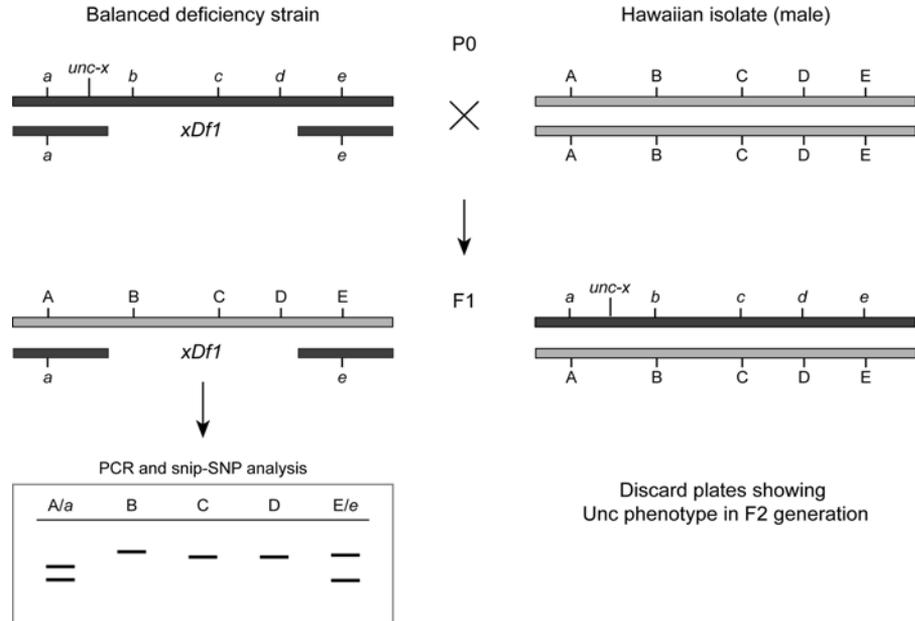
The majority of genomic rearrangements currently used in *C. elegans* research have been generated and captured in broad mutagenesis screens (Girard *et al.*, 2007). A consequence of this is that mutated genomes carry a significant number of single nucleotide differences. Not all of these mutational events will be eliminated by outcrossing and recombination. In many cases, accrued mutations will be retained because of suppressed crossing over in the region of the genomic rearrangement. Thus, the more fully characterized the rearrangement is the better it is for experimental interpretation. Many rearrangements have not yet been characterized at the molecular level. Analysis has been limited to general identification of the nature and extent of the rearrangement using relatively low-resolution labor-intensive genetic techniques. High-resolution characterization at the nucleotide level is now a practical option and a desirable approach to prevent problems arising from inconsistent or incorrect results.

#### A. Restriction Enzyme Digestion of Single Nucleotide Polymorphisms (snip-SNP)

Single nucleotide differences between strains (SNPs) provide a high-resolution mapping tool and a significant advance over traditional mapping methods (Wicks *et al.*, 2001). SNP mapping relies upon the availability of small changes at the genomic level between two closely related, yet distinct isolates of the same species. In *C. elegans* there is a sufficiently distant, though still compatible isolate, CB4856 (Hawaiian) (Flibotte *et al.*, 2009; Jakubowski and Kornfeld, 1999; Koch *et al.*, 2000; Swan *et al.*, 2002; Wicks *et al.*, 2001). Several million years of evolutionary drift between the wild-types Bristol N2 and Hawaiian CB4856 have resulted in a large number of base pair changes at intervals of 1000 bp on average.

Sequence alterations of various types are substrate for this type of analysis, including single-nucleotide changes, small deletions or insertions. Small deletions and insertions can be detected using a simple PCR assay. Nucleotide changes that alter a restriction site result in fragment length difference that can readily be observed by gel electrophoresis (snip-SNP).

Using SNPs to map deletions is relatively straightforward and has been applied to several deficiency strains (Kadandale *et al.*, 2005; Zhao *et al.*, 2008). A strain carrying the deletion to be characterized is mated to the Hawaiian isolate. Suitable snip-SNPs are then analyzed from DNA extracted from the resulting heterozygous progeny (Fig. 7). The deleted region of the genome will effectively be homozygous for the Hawaiian snip-SNPs while the rest of the genome will show a heterozygous pattern.



**Fig. 7** SNP mapping strategy for deficiency mapping. A strain heterozygous for the deficiency (in this hypothetical example *xDf1*) to be characterized and a balanced visible phenotypic marker is mated to the Hawaiian isolate strain. F1 progeny arising from this cross are individually picked and the resultant F2 animals scored for the presence of the visible phenotype (*mut-x*). F2 populations not displaying the *mut-x* phenotype (and therefore maintaining *xDf1*) are cultured, genomic DNA is extracted and analyzed using appropriate snip-SNPs.

A crossing scheme for SNP mapping a deficiency is shown in Fig. 7. The balanced *Df* strain is crossed to Hawaiian males. The F1 progeny will be heterozygous for Bristol and Hawaiian chromosomes and receive either the *Df* or the balancer chromosome. If the balancer is marked, it is possible to enrich for the heterozygotes. DNA from the F2 progeny is extracted and used as template in PCR amplification. It is important to prepare the DNA from the F2 generation to ensure that the viable Hawaiian chromosome does not outcompete the deficiency chromosome (Kadandale *et al.*, 2005). Since the DNA template can be used for multiple PCR reactions a strategy can be used to first roughly position the deficiency with broad markers and then use subsequent fine mapping in increments until SNPs are exhausted or breakpoints located to a sufficient resolution.

## 1. Advantages and Disadvantages

SNP analysis has several advantages over more traditional mapping techniques. The approach is rapid as a single genetic cross and DNA extraction followed by PCR restriction analysis can be done in a less than two weeks. No specific equipment or

reagents are required beyond those found in a standard molecular genetics laboratory (Kadandale *et al.*, 2005).

There are however some drawbacks. The mapping resolution is dependent on the availability of DNA sequence differences in the region of interest in the two strains. Negative results can be ambiguous, especially if rare or low-fidelity enzymes are used, which may require that additional control reactions be performed. This approach works well with simple deletions, but not with duplications, translocations, inversions, and complex deficiencies (indels).

## B. Oligonucleotide Array Comparative Genome Hybridization (aCGH)

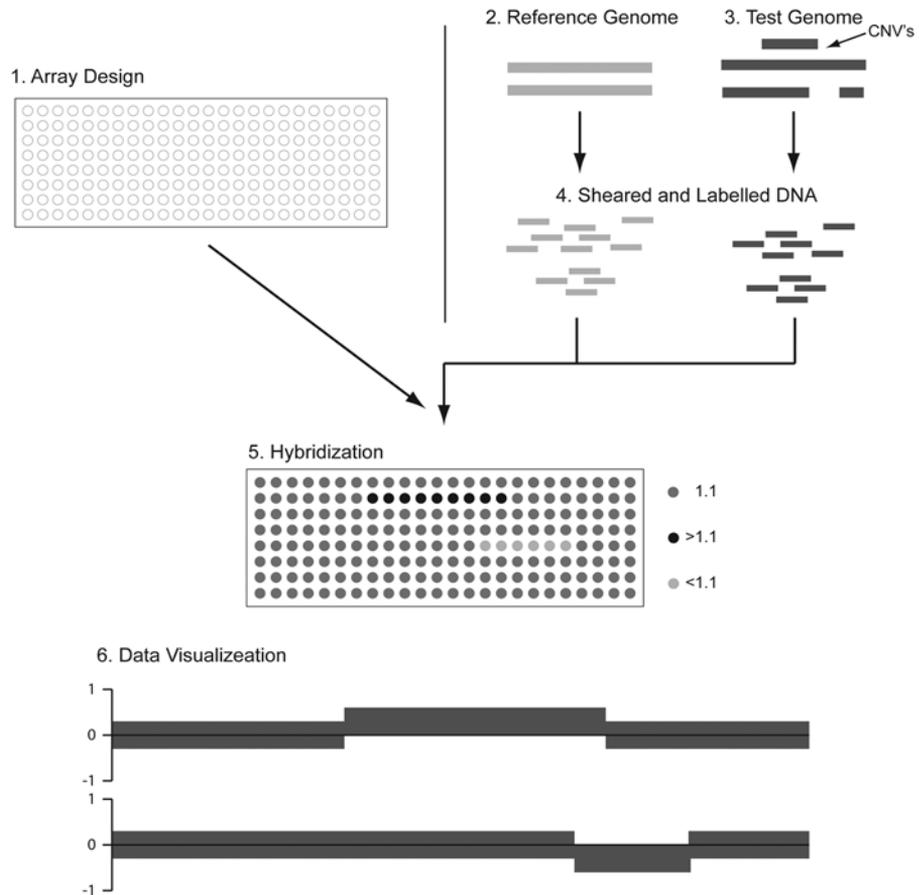
Complex rearrangements can be characterized on a genome-wide scale using oligonucleotide array comparative genomic analysis (aCGH) (Gresham *et al.*, 2008). aCGH is a technology used for high-resolution mapping of chromosomal copy number variation (CNV) at a genome-wide scale by comparing the ratio of DNA between two samples from the same organism (Dhami *et al.*, 2005; Selzer *et al.*, 2005).

The development of a *C. elegans* specific aCGH platform for identification of novel single-gene deletions has represented a powerful technology that can be adapted to the rapid and precise characterization of deficiency mapping strains (Maydan *et al.*, 2007). Currently, aCGH has been used to analyze the structure of deficiencies and duplications on chromosome III (Jones *et al.*, 2007), chromosome V (Jones *et al.*, 2009; Zhao *et al.*, 2008), and chromosome IV (D. Baillie, pers. comm.).

Determining the extent of duplications or deletions using aCGH is straightforward requiring only a sample of genomic DNA from the deficiency strain and a reference strain. The sensitivity of this approach is sufficient for the analysis of heterozygous animals (Maydan *et al.*, 2007). This means that the relatively large amount of genomic DNA required for the hybridization can be easily isolated from balanced rearrangements strains (Jones *et al.*, 2007). Using available software to visualised hybridization data CNVs can normally be detected without the need for statistical analysis or trained personnel. Precise breakpoint positions can be validated using methods such as PCR (Jones *et al.*, 2007) (Fig. 8).

### 1. Advantages and Disadvantages

aCGH is relatively straightforward and at less than \$1000 per chip is cost-effective considering the quantity and resolution of the information attained. Very high-resolution chips with probes for every kilobase (kb) are available and new high-density arrays with overlapping probes giving bp resolution for the entire genome have been developed (NimbleGenSystemsInc.). The results can be readily visualized without the need for statistical analysis. It is fast and breakpoint confirmation takes only a few weeks to obtain. Finally, little or no previous knowledge about the



**Fig. 8** Overview of aCGH analysis. (1) Array synthesis: Target genome sequence is used to select suitable probe sequences that are synthesized on an array. (2–4) DNA preparation: Genomic DNA from test (2) and reference (3) samples is isolated and differentially labeled with fluorescent dyes (4). DNA is sheared to facilitate hybridization to the array. (5–6) Data acquisition and analysis: samples are hybridized to the array together (5) and the array is processed and scanned. Signal output is processed to create a ratio of fluorescence, which is proportional to the ratio of reference to test samples. This data is visualized using graphical user interface applications (6).

rearrangement is required prior to aCGH analysis. Novel rearrangements can therefore be generated and characterized immediately.

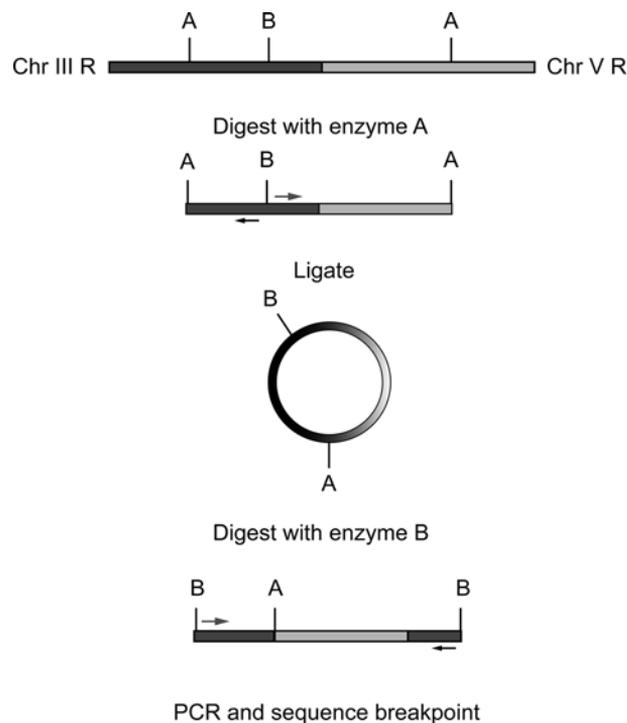
aCGH analysis is however not suitable for the analysis of very complex rearrangements. Duplications present in the background of a deficiency strain can be difficult to interpret as the aCGH approach does not give any positional information about where possible insertions have occurred. Additionally rearrangements that do not result in CNV, such as inversions, cannot be analyzed using this method.

### C. Inverse PCR

In the case of a balancer in which the location of one breakpoint has been mapped to a small region of the genome that information can be used to perform inverse PCR of the rearranged chromosomes. This approach will identify the remaining breakpoint position. In the case for *eTl*, which breaks in *unc-36* (III), PCR across the 7 kb long gene was used to locate the chromosome III breakpoint to a 450 bp interval. Inverse PCR was subsequently used to identify the sequence of the fusion between chromosome III and chromosome V (Zhao *et al.*, 2006) (Fig. 9). This information was useful for developing a PCR assay to rapidly identify the *eTl* chromosomes in genetic crosses (Zhao *et al.*, 2006).

### D. Whole Genome Sequencing

Advances in sequencing technology have made whole genome sequencing experiments feasible. There are a growing number of *C. elegans* whole genome sequences available (Hillier *et al.*, 2007, 2008; Rose *et al.*, 2010), in addition to the canonical assembled and aligned sequence in WormBase (Harris *et al.*, 2010) ([www.wormbase.org](http://www.wormbase.org)).



**Fig. 9** iPCR scheme is used to isolate *eTl* breakpoint. See text for details.

[wormbase.org](http://wormbase.org)). Genome sequencing can identify CNV as well as base pair differences and can therefore potentially identify all genomic alterations in a rearranged chromosome regardless of type or complexity.

## 1. Advantages and Disadvantages

The major advantage of genome sequencing is the ability to rapidly and accurately accumulate a complete sequence analysis of the entire spectrum of alterations in the genome, something that has not previously been possible. The technology is evolving rapidly making this approach much more accessible to individual researchers. However, one still needs to have access to high-level bioinformatic analysis in order to interpret the data. This is though a very powerful approach, providing a comprehensive view of the genome and may become increasingly accessible for analysis of SCs in the near future.

---

---

---

## IV. Engineered Constructs: Transgenic Arrays

### A. Experimental Applications of Transgenic Arrays

Transgenic arrays have been used for a range of research purposes including, but not limited to, discovery of the molecular identity of a mutated gene and cloning it by rescue of the mutant phenotype, study of overexpression or ectopic expression of genes, manipulation and investigation of gene structure and function, discovery and analysis of *cis*-acting regulatory motifs and their *trans*-acting control factors, and creation of tagged proteins for specific manipulation or study. In addition, they can be used in genetic screens and for mosaic analysis. In this section, we will highlight how transgenic arrays have been used for *C. elegans* research as well as the various techniques and approaches available for the generation and manipulation of these arrays.

## 1. Mutant Rescue

Rescuing a mutant phenotype is one of the most common uses of transgenic arrays in *C. elegans* research. Rescuing arrays contain a wild-type copy of the gene of interest. Libraries of both cosmid (<http://www.sanger.ac.uk/technology/clonerequests/>) and fosmid (<http://elegans.bcgsc.bc.ca/>) clones exist and these resources can be mined for a suitable clone containing the genomic fragment of the gene of interest including its surrounding promoter elements. Cosmids and Fosmids generally work well for mutant rescue because they contain a relatively large amount of DNA (35 kb) and are therefore likely to contain all the enhancer elements required for faithful expression of the gene (Tursun *et al.*, 2009). Full-length cDNAs can also be used for rescue experiment, though the cDNA will need to be cloned into a vector containing a suitable promoter. This type of construct is useful when the gene spans a large genomic region or for genes for which no other genomic clone is available. Alternatively, a

straightforward approach for generating a wild-type clone is to PCR the sequence directly from genomic DNA. This approach requires the use of a high-fidelity polymerase and that the region to be amplified is of a suitable size for reliable PCR amplification.

## 2. Expression Pattern Analysis

The transparency and size of *C. elegans* allows for visualization of gene expression patterns using the reporter genes *in vivo* without animal dissection. The *E. coli* gene *LacZ*, encoding  $\beta$ -galactosidase, was first used for expression analysis in *C. elegans* (Fire *et al.*, 1990) but has since been superseded with the introduction of GFP, the use of which was pioneered in this organism (Chalfie *et al.*, 1994). Expression reporters can be transcriptional or translational. Transcriptional reporter constructs contain the native promoter elements of the gene of interest which are used to drive expression of the reporter. Generally, making transcriptional reporters is straightforward and uses cloning methods or simple and efficient PCR-based methods (Hobert, 2002; Hunt-Newbury *et al.*, 2007; McKay *et al.*, 2003). Alternatively, translational reporters can be constructed where the reporter is fused into the protein-coding sequence providing a hybrid protein. Translational reporters often faithfully replicate the native gene's function though using this approach may require that several constructs be created to find a position in the gene where the insertion of GFP does not inhibit the function of the native protein (Merritt and Seydoux, 2010). Generally, using the translational reporter to rescue the phenotype of a mutation is sufficient to confirm that all promoter elements are present and that the hybrid protein functions in the same way as the native protein.

## 3. Ectopic Expression Reporters

For some purposes it is not necessary, or even desirable, to use the endogenous promoter. A variety of endogenous promoter elements have been characterized in *C. elegans* (Davis *et al.*, 2008). Such promoters can be used to drive ectopic expression of a gene of interest in a known tissue type or temporal pattern (Kalb *et al.*, 1998; Mah *et al.*, 2007; McGhee *et al.*, 2009). As more promoter elements are discerned more become available and an effort to fully characterize tissue-specific promoters is underway (M. Chalfie pers. comm.). Additionally, conditional promoters are available for inducing gene expression under a variety of conditions including various heat-shock elements (Seydoux *et al.*, 1996).

## 4. Other Uses

Transgenic arrays are not limited to gene expression studies but have also been used to create a plethora of experimental tools. Some of the many examples include: tissue-specific RNAi analysis (Briese *et al.*, 2006; Qadota *et al.*, 2007; Voutev and Hubbard, 2008); identification of synaptic partners in the nervous system

(Feinberg *et al.*, 2008); investigation of meiotic processes (Mets and Meyer, 2009); and investigation of endosomal trafficking (Poteryaev *et al.*, 2007).

## B. Methods for Construction of Transgenic Arrays Cloning

Fire *et al.* developed a useful tool kit which includes a comprehensive set of commercially available expression vectors (<http://www.addgene.org/>). These vectors allow for the fusion of gene sequences with a variety of fluorescent variants and ectopic and inducible promoters using standard cloning techniques. Alternative cloning systems based on recombination cloning are also available and replace traditional restriction-enzyme cloning (Hope *et al.*, 2004; Lamesch *et al.*, 2004; Luan *et al.*, 2004; Rual *et al.*, 2004). These techniques allow for modification and assembly of multiple components or “modules” enabling the user to rapidly recombine different parts of the transgene (Merritt and Seydoux, 2010). With this method, a library of different experimental reagents, such as tissue-specific constructs can be assembled relatively easily.

### 1. Advantages and Disadvantages

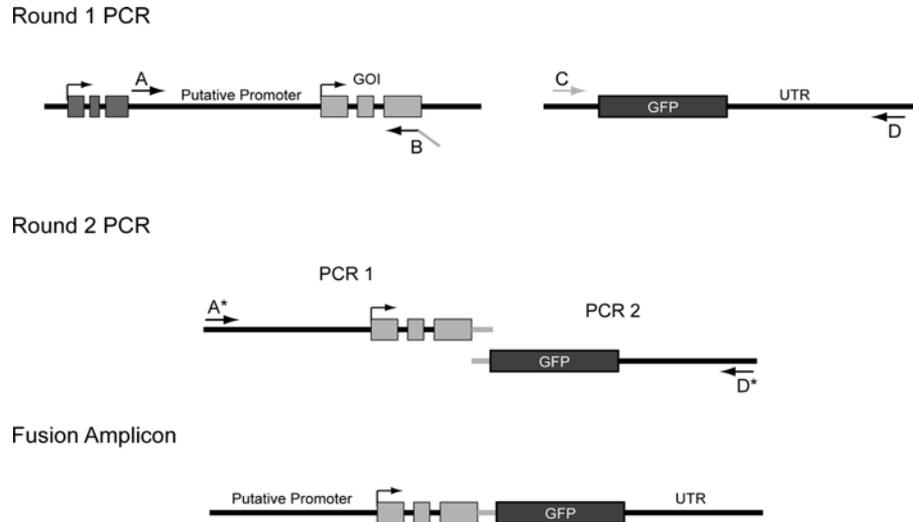
Cloning is a standard laboratory practice and results in the production of a construct that can be maintained and amplified as a renewable resource. Such clones can be used in additional engineering steps to quickly develop a suit on constructs for in-depth analysis. A drawback to this approach is the technical limitations on insert size precluding the use for large genes or genomic elements.

### 2. PCR Stitching

A simple and effective method for engineering reporter constructs is to use the PCR “stitching” method. PCR stitching is a method for joining two or more separate DNA sequences in a single PCR reaction (for an in-depth description of the technique see Boulin *et al.*, 2006). This technique relies on a short region of homology between the separate sequences acting on internal primer site (Fig. 10). The resulting fused sequence can be purified or often injected directly into animals. Since its development, this technique has been used extensively in *C. elegans* for both small and large-scale expression analysis (Dupuy *et al.*, 2007; Gallo *et al.*, 2006; Huang *et al.*, 2007; Zhao *et al.*, 2004a, 2004b).

### 3. Advantages and Disadvantages

PCR stitching is a straightforward approach that does not require any specialized reagents or techniques. Constructs can be rapidly generated and are easily modified to suit specific purposes. There are however limitations on the size of fusion sequences that can be easily produced and this approach results in a finite amount



**Fig. 10** PCR stitching. The approach requires two rounds of PCR amplification. In the first round, the genomic region of interest (a promoter element or full-length gene) is amplified while in a separate reaction the reporter element is also amplified. Primer B is designed with a 20 bp “tail” with homology to the 5’ end of the reporter element. In the second round, DNA from the two PCRs are combined into a single reaction. Subsequent amplification is performed with nested external primers. The presence of the 20 bp region of homology acts as an internal priming site for the two DNA fragments facilitating the amplification of a fusion product.

of material that may require additional cloning steps if the construct is required in large amounts.

#### 4. Recombineering

Since regulatory elements that affect gene expression can reside within introns or downstream of a gene (Conradt and Horvitz, 1999) and many genes are regulated at the post-transcriptional level (Ambros, 2004), it is essential to include these elements in an expression construct if an accurate representation of the expression pattern is to be achieved. Recombination-mediated methods to engineer fluorescent protein fusions in the context of genomic clones are therefore an attractive approach to use (Bamps and Hope, 2008; Dolphin and Hope, 2006; Sarov *et al.*, 2006; Tursun *et al.*, 2009). The availability of a *C. elegans* fosmid library (<http://elegans.bcgsc.bc.ca/>) that covers ~80% of the genome (D. Moerman, personal communication) is an essential resource that has made this approach feasible. The relatively large genomic fragments containing fosmids (~35 kb) combined with the redundant coverage of the fosmid library means that a suitable genomic clone containing the gene of interest and surrounding *cis*-regulatory control elements can be identified in most cases.

Recombineering relies on homologous recombination between a reporter construct and genomic clone mediated by bacteria expressing the  $\lambda$  Red recombinase

(Yu *et al.*, 2000). In brief, a linear DNA construct containing the tag of interest is generated with flanking sequence of homology to the insertion site (around 50 bp in length). This reporter “cassette” is then introduced into a  $\lambda$  Red recombinase compatible bacterial strain containing the fosmid to be manipulated. Subsequent activation of the  $\lambda$  Red recombinase results in homologous recombination between the reporter cassette and the fosmid. The resulting modified fosmid can be recovered, amplified, and used to generate transgenic animals. Several recombineering protocols have been described in detail (Dolphin and Hope, 2006; Sarov *et al.*, 2006) though a recent modification of the protocol by Hobert *et al.* has streamlined the process significantly (Tursun *et al.*, 2009). In addition, this group had developed a series of recombineering cassettes that allow virtually seamless insertion of several GFP variants and other useful elements.

## 5. Advantages and Disadvantages

The obvious advantage of this technique is that the ability to manipulate large genomic clones virtually eliminates the restriction of gene size. The protocol does however require a suitable fosmid covering the gene of interest.

## C. Properties for Construction of Transgenic Arrays

There are certain criteria that define a usable and maintainable transgenic array. Such arrays contain the DNA element specific to its intended purpose. For example, if the purpose is for rescue the array will contain the wild-type copy of the gene to be rescued, or if the purpose is to study gene expression, the array will contain a fluorescent reporter. In addition, an element that allows for selection of those animals that have successfully taken up and assembled the array is required. This can be either a genetic marker such as a wild-type copy of a gene that rescues an obvious phenotype (Granato *et al.*, 1994), or a molecular marker such as a visible fluorescent reporter (Gu *et al.*, 1998).

### 1. Selectable Transformation Markers

In order to maintain and follow the arrays through genetic crosses selectable markers are used. A frequently used approach is the rescue of a recessive phenotype by integration of the wild-type gene into the array. For an in-depth review of commonly used genes see Evans *et al.* (Girard *et al.*, 2007). Rescuing a detrimental phenotype has the advantage of creating strains that are relatively easy to maintain. *pha-1(e2123)*, for example, is a temperature-sensitive allele, the use of which allows animals to be raised at permissive temperature prior to transformation (Granato *et al.*, 1994). After injection of a rescuing construct, growth at the restrictive temperature allows for selection of animals maintaining the transgenic array. A disadvantage, however, is that arrays can only be easily selected in the original genetic background, precluding their use in genetic crosses.

Dominant transgenic markers are useful for following an array during crosses to other strains. An example is the rescue of the *rol-6(su106)* mutation (Mello *et al.*, 1991), which causes the worms to roll around its long axis. However, the morphological phenotype can interfere with observations of phenotypes caused by the gene of interest. Another popular dominant marker is *suf-5::GFP* which is benign morphologically (Gu *et al.*, 1998). Transgenic animals carrying the GFP marked array can be followed using a fluorescent dissecting microscope.

#### D. Methods for Creating Transgenic Animals

##### 1. Microinjection

The standard approach to making transgenic strains is to coinject two or more DNAs into the distal gonad syncytium of an animal. One of the DNAs carries the DNA construct to be studied (the transgene) and the other a plasmid that contains a marker to select for successful transformation. Both linear and circular DNAs can form into arrays. Injected DNAs undergo homologous recombination and in this way become assembled into an array (Mello *et al.*, 1991). Using injected DNAs that share sequence homology can facilitate efficient homologous recombination, although nonhomologous recombination has also been observed (Mello and Fire, 1995). It is possible to introduce several DNA constructs, although independent verification of each component should be performed (usually this can be done by PCR from DNA extracted from transgenic animals) to ascertain that the recovered array contains all coinjected components.

Transgenic animals produced by injection typically have large extrachromosomal arrays that contain many copies of the coinjected DNAs (Mello *et al.*, 1991). A fraction of these repetitive arrays become heritably stable and some of the first-generation progeny (F1) will transmit the array through subsequent generations. Once the strain is established, transmission is often reproducible with regard to heritability or in the case of somatic promoters, expression. However, some degree of mitotic instability will occur and these are incompletely inherited. If the arrays are integrated into a chromosome, usually by treatment with radiation, it is possible to create a strain that transmits the array in a Mendelian manner (Mello *et al.*, 1991).

##### 2. Advantages and Disadvantages

Transgenic constructs do not always reproduce the expression patterns of the endogenous genes. Genes expressed in the germ line are problematic. Transgenes in repetitive arrays are strongly silenced in germ cell nuclei (Kelly *et al.*, 1997). It is however possible to circumvent this issue by using mutant animals that are deficient for genes that elicit germ line silencing (Kelly and Fire, 1998). Additionally, it is difficult to predict and control the level of expression among different arrays, and expression can be variable among siblings of a single strain even when using integrated arrays showing expression variability (Mello and Fire, 1995); this can

be detrimental for experiments where native or low-level expression of the transgene is critical. RNAi-like effects such as cosuppression can also result in the suppression of endogenous gene function, complicating analyses (Dernburg *et al.*, 2000).

### 3. Complex Arrays

It is the repetitive nature of transgenic arrays that is the most likely cause of their preferential silencing in the germ line (Kelly *et al.*, 1997). To circumvent this, arrays can be constructed in such a manner as to limit the number of repetitive elements. One way this can be done is with the addition of fragmented genomic DNA into injection mixes (Kelly *et al.*, 1997). The incorporation of the genomic fragments acts as a buffer to limit repeat sequence formation. These “complex” arrays transmit as heritable extrachromosomal complexes in the same manner as standard arrays.

### 4. Advantages and Disadvantages

The major advantage of complex arrays is for analysis of germ line-expressed genes. For unknown reasons however, germ line expression from complex arrays is not as stable as standard arrays and expression can disappear after the first few generations even in strains that retain the array. Maintaining animals at 25°C suppresses transgene inactivation (Reese *et al.*, 2000; Strome *et al.*, 2001).

### 5. Biolistic Transformation

Microparticle bombardment is another method commonly used for transgenic array transformation (Jackstadt *et al.*, 1999; Praitis *et al.*, 2001; Wilm *et al.*, 1999; Zhao *et al.*, 2009). In this approach, DNA is bound onto gold particles that are then “shot” into worms using a ballistic bombardment instrument or “gene gun” (Wilm *et al.*, 1999). Particle bombardment generates extrachromosomal arrays with a significant number of transformants having only a few copies of the transgene integrated at various nonhomologous sites in the genome (Jackstadt *et al.*, 1999; Zhao *et al.*, 2009). The mechanism of integration is unclear (Fire, 1986; Girard *et al.*, 2007; Praitis *et al.*, 2001). For this technique, the transgenes must be subcloned into the same plasmid containing the transformation marker. This is to ensure that all required transformation and reporter components are delivered and integrated together (Jackstadt *et al.*, 1999).

### 6. Advantages and Disadvantages

This method while technically more difficult has several advantages over standard injection methods. An important advantage of this approach is that stable integrated transgenic strains can be isolated directly (Praitis *et al.*, 2001). Moreover, many integrated transgenes do not undergo germ line silencing (Praitis *et al.*, 2001). Finally, a scaled-up bombardment protocol has been developed to allow consistent isolation of strains with homologous gene replacements (Berezikov *et al.*, 2004).

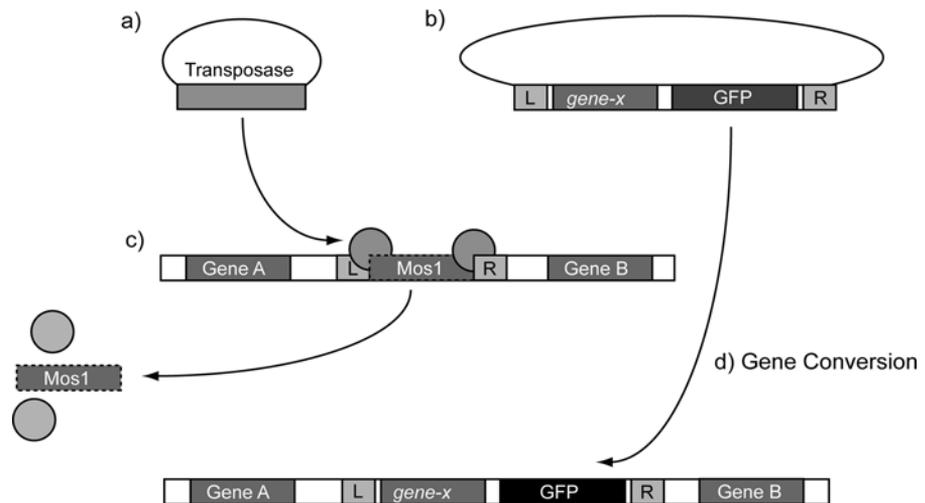
However, preparation of nematodes and materials for gene bombardment is significantly more labor and material intensive than for microinjection, and it takes longer to generate transgenic strains.

### 7. Single-Copy Insertion by MosSCI

Recently, an approach for obtaining integration of transgenes as single copies at a defined genomic site had been developed (Frokjaer-Jensen *et al.*, 2008). Mos1-mediated single-copy insertion (MosSCI) takes advantage of the Mos transposable element (Granger *et al.*, 2004; Robert and Bessereau, 2009; Robert *et al.*, 2009) to introduce modified reporter constructs directly into the genome (Fig. 11).

In this approach, the construct of interest is inserted into a vector containing flanking homology to the genomic region surrounding a known Mos element. This vector includes a WT copy of *unc-119* gene for use as a positive selection marker. This insertion vector, along with a clone carrying the transposase gene under the control of a heat-shock promoter and fluorescent reporters are injected into an *unc-119* mutant strain that harbors a single Mos1 element in an intergenic region of the genome. (Vector and strain details are available at <http://sites.google.com/site/jorgensenmossci/Home>).

Animals containing the assembled transgenic array express the coinjected fluorescent markers. Heat-shock activation of the transposase in these animals leads to excision of the Mos1 element. In rare instances, the resulting double strand break is



**Fig. 11** MosSci single-copy insertion of transgenes. In this approach, a reporter construct is engineered into a modified Mos1 transposon. This construct, along with a clone carrying the activating transposase gene under the control of a heat-shock promoter, is injected into a carrier strain that harbors a single Mos1 integration site in an intergenic region of the genome. Subsequent induction of the transposase results in insertion of the hybrid transposon into the genome.

repaired using the homologous DNA engineered into the insertion vector as template. Integration events can be isolated by screening for WT animals (*unc-119* positive) that have lost the transgenic array (do not express the fluorescent reporters).

## 8. Advantages and Disadvantage

MosSCI is a potentially powerful approach that eliminates many of the problems associated with generating transgenic arrays. The insertion of a single copy of the construct circumvents the issues of germ line silencing and overexpression, leading to more stable native gene expression. The technique is however relatively labor intensive, requiring the generation of several potentially complex DNA constructs and the use of animals containing specific *Mos1* insertions (Frokjaer-Jensen *et al.*, 2008). Screening for rare integration events also requires the use of a fluorescent microscope. Additionally, this technique only allows for efficient insertion constructs of around 15 kb and up to 30% of the integrants will not be full length. This method is however currently the only reliable method available for creating single copy integrants.

## E. Use of Transgenic Arrays as Balancers

Under certain conditions, a transgenic array can be considered a specialized form of duplication. In this context, they have been used to rescue and maintain specific mutations. However, their rescuing ability has also been utilized as part of a system for isolating suppressors of mutations that result in low-viability phenotypes (Fay and Han, 2000; Fay *et al.*, 2002). In these screens, wild-type copies of the gene are carried in the array, rescuing the detrimental phenotype of the mutated animals. An advantage of using a transgenic array is that it can be engineered to meet the needs of the screen, for example, Fay *et al.* created rescuing arrays containing wild-type copies of the *lin-35* gene with a GFP reporter to facilitate ease of screening. Transgenic lines made in this manner transmit the extrachromosomal array only to a fraction of their progeny. The screen exploits this characteristic by screening for viable animals that have lost the array after a mutagenesis treatment. Such animals are likely to have acquired a second site suppressor mutation.

## 1. Advantages and Disadvantages

Transgenic arrays can be used to great effect for the rescue of specific single-gene mutations. Transgenic arrays, unlike large rearrangements, can also be tailored to specific requirements such as the inclusion of specific promoter elements or the use of fluorescent marker for rapid identification. It is however difficult to rescue mutations in genes that are required in the germ line due to silencing of the array. The use of transgenes for mutagenesis screens can also create complexity when screening for suppressor mutations as integration of the transgene sequence into the genome or disruption of the fluorescent markers can lead to false positives that are difficult to identify and eliminate.

## Acknowledgments

The authors thank the following people for their valuable help and comments on this manuscript: D.L. Baillie, D. Moerman, D. Riddle, R.C. Johnsen, N.J. O'Neil, M. Edgley, J. McLellan, and H. Shin, with a special thanks to our reviewers for their thoughtful comments.

## References

- Adames, K. A., Gawne, J., Wicky, C., Muller, F., and Rose, A. M. (1998). Mapping a telomere using the translocation eT1(III;V) in *Caenorhabditis elegans*. *Genetics* **150**, 1059–1066.
- Ahmed, S., and Hodgkin, J. (2000). MRT-2 checkpoint protein is required for germ line immortality and telomere replication in *C. elegans*. *Nature* **403**, 159–164.
- Ahnn, J., and Fire, A. (1994). A screen for genetic loci required for body-wall muscle development during embryogenesis in *Caenorhabditis elegans*. *Genetics* **137**, 483–498.
- Ambros, V. (2004). The functions of animal microRNAs. *Nature* **431**, 350–355.
- Austin, J., and Kimble, J. (1989). Transcript analysis of *glp-1* and *lin-12*, homologous genes required for cell interactions during development of *C. elegans*. *Cell* **58**, 565–571.
- Bamps, S., and Hope, I. A. (2008). Large-scale gene expression pattern analysis, in situ, in *Caenorhabditis elegans*. *Brief Funct. Genomic Proteomic* **7**, 175–183.
- Barstead, R. J., and Moerman, D. G. (2006). *C. elegans* deletion mutant screening. *Methods Mol. Biol.* **351**, 51–58.
- Belfiore, M., Mathies, L. D., Pugnale, P., Moulder, G., and Barstead, R., *et al.* (2002). The MEP-1 zinc-finger protein acts with MOG DEAH box proteins to control gene expression via the *fem-3* 3' untranslated region in *Caenorhabditis elegans*. *RNA* **8**, 725–739.
- Berezikov, E., Bargmann, C. I., and Plasterk, R. H. (2004). Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40.
- Bosher, J. M., Hahn, B. S., Legouis, R., Sookhareea, S., and Weimer, R. M., *et al.* (2003). The *Caenorhabditis elegans* *vab-10* spectraplaklin isoforms protect the epidermis against internal and external forces. *J. Cell Biol.* **161**, 757–768.
- Boulin, T., Etchberger, J. F., and Hobert, O. (2006). Reporter Gene Fusions. *WormBook* 5:1-23
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- Briese, M., Esmacili, B., Johnson, N. M., and Sattelle, D. B. (2006). pWormgatePro enables promoter-driven knockdown by hairpin RNA interference of muscle and neuronal gene products in *Caenorhabditis elegans*. *Invert Neurosci.* **6**, 5–12.
- Broverman, S. A., and Meneely, P. M. (1994). Meiotic mutants that cause a polar decrease in recombination on the X chromosome in *Caenorhabditis elegans*. *Genetics* **136**, 119–127.
- Bucher, E. A., and Greenwald, I. (1991). A genetic mosaic screen of essential zygotic genes in *Caenorhabditis elegans*. *Genetics* **128**, 281–292.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Chanal, P., and Labouesse, M. (1997). A screen for genetic loci required for hypodermal cell and glial-like cell development during *Caenorhabditis elegans* embryogenesis. *Genetics* **146**, 207–226.
- Clark, D. V., Johnsen, R. C., McKim, K. S., and Baillie, D. L. (1990). Analysis of lethal mutations induced in a mutator strain that activates transposable elements in *Caenorhabditis elegans*. *Genome* **33**, 109–114.
- Clark, D. V., Rogalski, T. M., Donati, L. M., and Baillie, D. L. (1988). The *unc-22(IV)* region of *Caenorhabditis elegans*: Genetic analysis of lethal mutations. *Genetics* **119**, 345–353.
- Conradt, B., and Horvitz, H. R. (1999). The TRA-1A sex determination protein of *C. elegans* regulates sexually dimorphic cell deaths by repressing the *egl-1* cell death activator gene. *Cell* **98**, 317–327.
- Cui, M., Kim, E. B., and Han, M. (2006). Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in *C. elegans*. *PLoS Genet.* **2**, e74.
- Davis, M. W., Morton, J. J., Carroll, D., and Jorgensen, E. M. (2008). Gene activation using FLP recombinase in *C. elegans*. *PLoS Genet.* **4**, e1000028.

- DeLong, L., Casson, L. P., and Meyer, B. J. (1987). Assessment of X chromosome dosage compensation in *Caenorhabditis elegans* by phenotypic analysis of *lin-14*. *Genetics* **117**, 657–670.
- Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the *C. elegans* germ line. *Genes Dev.* **14**, 1578–1583.
- Dhami, P., Coffey, A. J., Abbs, S., Vermeesch, J. R., and Dumanski, J. P., *et al.* (2005). Exon array CGH: Detection of copy-number changes at the resolution of individual exons in the human genome. *Am. J. Hum. Genet.* **76**, 750–762.
- Dolphin, C. T., and Hope, I. A. (2006). *Caenorhabditis elegans* reporter fusion genes generated by seamless modification of large genomic DNA clones. *Nucleic Acids Res.* **34**, e72.
- Dupuy, D., Bertin, N., Hidalgo, C. A., Venkatesan, K., and Tu, D., *et al.* (2007). Genome-scale analysis of *in vivo* spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat. Biotechnol.* **25**, 663–668.
- Edgley, M. L., Baillie, D. L., Riddle, D. L., and Rose, A. M. (2006). Genetic Balancers. *WormBook* **6**, 1–32.
- Edgley, M. L., and Riddle, D. L. (2001). LG II balancer chromosomes in *Caenorhabditis elegans*: mT1 (II;III) and the mIn1 set of dominantly and recessively marked inversions. *Mol. Genet. Genomics* **266**, 385–395.
- Fay, D. (2006). Genetic Mapping and Manipulation: Chapter 6 – Mapping with Deficiencies and Duplications. *WormBook* **17**, 1–12.
- Fay, D. S., and Han, M. (2000). The synthetic multivulval genes of *C. elegans*: Functional redundancy Ras-antagonism, and cell fate determination. *Genesis* **26**, 279–284.
- Fay, D. S., Keenan, S., and Han, M. (2002). *fzr-1* and *lin-35/Rb* function redundantly to control cell proliferation in *C. elegans* as revealed by a nonbiased synthetic screen. *Genes Dev.* **16**, 503–517.
- Feinberg, E. H., Vanhoven, M. K., Bendesky, A., Wang, G., and Fetter, R. D., *et al.* (2008). GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron.* **57**, 353–363.
- Ferguson, E. L., and Horvitz, H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. *Genetics* **110**, 17–72.
- Fire, A. (1986). Integrative transformation of *Caenorhabditis elegans*. *EMBO J.* **5**, 2673–2680.
- Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**, 189–198.
- Flibotte, S., Edgley, M. L., Maydan, J., Taylor, J., and Zapf, R., *et al.* (2009). Rapid high resolution single nucleotide polymorphism-comparative genome hybridization mapping in *Caenorhabditis elegans*. *Genetics* **181**, 33–37.
- Fodor, A., and Deák, P. (1985). The isolation and genetic analysis of a *Caenorhabditis elegans* translocation (szT1) strain bearing an X- chromosome balancer. *J. Genet.* **64**, 143–157.
- Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., and Thummel, J. M., *et al.* (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383.
- Gallo, M., Mah, A. K., Johnsen, R. C., Rose, A. M., and Baillie, D. L. (2006). *Caenorhabditis elegans* *dpy-14*: An essential collagen gene with unique expression profile and physiological roles in early development. *Mol. Genet. Genomics* **275**, 527–539.
- Girard, L. R., Fiedler, T. J., Harris, T. W., Carvalho, F., and Antoshechkin, I., *et al.* (2007). WormBook: The online review of *Caenorhabditis elegans* biology. *Nucleic Acids Res.* **35**, D472–D475.
- Gladden, J. M., Farboud, B., and Meyer, B. J. (2007). Revisiting the X:A signal that specifies *Caenorhabditis elegans* sexual fate. *Genetics* **177**, 1639–1654.
- Gladden, J. M., and Meyer, B. J. (2007). A ONECUT homeodomain protein communicates X chromosome dose to specify *Caenorhabditis elegans* sexual fate by repressing a sex switch gene. *Genetics* **177**, 1621–1637.
- Granato, M., Schnabel, H., and Schnabel, R. (1994). *pha-1*, a selectable marker for gene transfer in *C. elegans*. *Nucleic Acids Res.* **22**, 1762–1763.
- Granger, L., Martin, E., and Segalat, L. (2004). Mos as a tool for genome-wide insertional mutagenesis in *Caenorhabditis elegans*: Results of a pilot study. *Nucleic Acids Res.* **32**, e117.

- Gresham, D., Dunham, M. J., and Botstein, D. (2008). Comparing whole genomes using DNA microarrays. *Nat. Rev. Genet.* **9**, 291–302.
- Gu, T., Orita, S., and Han, M. (1998). *Caenorhabditis elegans* SUR-5, a novel but conserved protein, negatively regulates LET-60 Ras activity during vulval induction. *Mol. Cell. Biol.* **18**, 4556–4564.
- Harris, T. W., Antoshechkin, I., Bieri, T., Blasiar, D., and Chan, J., *et al.* (2010). WormBase: A comprehensive resource for nematode research. *Nucleic Acids Res.* **38**, D463–D467.
- Hedgecock, E. M., and Herman, R. K. (1995). The *ncl-1* gene and genetic mosaics of *Caenorhabditis elegans*. *Genetics* **141**, 989–1006.
- Herman, R. K. (1978). Crossover suppressors and balanced recessive lethals in *Caenorhabditis elegans*. *Genetics* **88**, 49–65.
- Herman, R. K. (1989). Mosaic analysis in the nematode *Caenorhabditis elegans*. *J. Neurogenet.* **5**, 1–24.
- Herman, R. K., Albertson, D. G., and Brenner, S. (1976). Chromosome rearrangements in *Caenorhabditis elegans*. *Genetics* **83**, 91–105.
- Herman, R. K., Kari, C. K., and Hartman, P. S. (1982). Dominant X-chromosome nondisjunction mutants of *Caenorhabditis elegans*. *Genetics* **102**, 379–400.
- Herman, R. K., Madl, J. E., and Kari, C. K. (1979). Duplications in *Caenorhabditis elegans*. *Genetics* **92**, 419–435.
- Hillers, K. J., and Villeneuve, A. M. (2003). Chromosome-wide control of meiotic crossing over in *C. elegans*. *Curr. Biol.* **13**, 1641–1647.
- Hillier, L. W., Marth, G. T., Quinlan, A. R., Dooling, D., and Fewell, G., *et al.* (2008). Whole-genome sequencing and variant discovery in *C. elegans*. *Nat. Methods* **5**, 183–188.
- Hillier, L. W., Miller, R. D., Baird, S. E., Chinwalla, A., and Fulton, L. A., *et al.* (2007). Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. *PLoS Biol.* **5**, e167.
- Hobert, O. (2002). PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques* **32**, 728–730.
- Hodgkin, J. (1980). More sex-determination mutants of *Caenorhabditis elegans*. *Genetics* **96**, 649–664.
- Hodgkin, J. (1986). Sex determination in the nematode *C. elegans*: Analysis of *tra-3* suppressors and characterization of *fem* genes. *Genetics* **114**, 15–52.
- Hope, I. A., Stevens, J., Garner, A., Hayes, J., and Cheo, D. L., *et al.* (2004). Feasibility of genome-scale construction of promoter::reporter gene fusions for expression in *Caenorhabditis elegans* using a multisite gateway recombination system. *Genome Res.* **14**, 2070–2075.
- Horvitz, H. R., Brenner, S., Hodgkin, J., and Herman, R. K. (1979). A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. *Mol. Gen. Genet.* **175**, 129–133.
- Howell, A. M., and Rose, A. M. (1990). Essential genes in the *hDf6* region of chromosome I in *Caenorhabditis elegans*. *Genetics* **126**, 583–592.
- Huang, P., Pleasance, E. D., Maydan, J. S., Hunt-Newbury, R., and O’Neil, N. J., *et al.* (2007). Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. *Genome Res.* **17**, 1478–1485.
- Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., and Anastas, D., *et al.* (2007). High-throughput *in vivo* analysis of gene expression in *Caenorhabditis elegans*. *PLoS Biol.* **5**, e237.
- Hunter, C. P., and Wood, W. B. (1990). The *tra-1* gene determines sexual phenotype cell-autonomously in *C. elegans*. *Cell* **63**, 1193–1204.
- Hunter, C. P., and Wood, W. B. (1992). Evidence from mosaic analysis of the masculinizing gene *her-1* for cell interactions in *C. elegans* sex determination. *Nature* **355**, 551–555.
- Jackstadt, P., Wilm, T. P., Zahner, H., and Hobom, G. (1999). Transformation of nematodes via ballistic DNA transfer. *Mol. Biochem. Parasitol.* **103**, 261–266.
- Jakubowski, J., and Kornfeld, K. (1999). A local, high-density, single-nucleotide polymorphism map used to clone *Caenorhabditis elegans* *cdf-1*. *Genetics* **153**, 743–752.
- Johnsen, R. C., and Baillie, D. L. (1988). Formaldehyde mutagenesis of the eT1 balanced region in *Caenorhabditis elegans*: Dose–response curve and the analysis of mutational events. *Mutat. Res.* **201**, 137–147.

- Johnsen, R. C., and Baillie, D. L. (1991). Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans*. *Genetics* **129**, 735–752.
- Jones, M. R., Chua, S. Y., O'Neil, N. J., Johnsen, R. C., and Rose, A. M., *et al.* (2009). High-resolution array comparative genomic hybridization analysis reveals unanticipated complexity of genetic deficiencies on chromosome V in *Caenorhabditis elegans*. *Mol. Genet. Genomics* **282**, 37–46.
- Jones, M. R., Maydan, J. S., Flibotte, S., Moerman, D. G., and Baillie, D. L. (2007). Oligonucleotide array comparative genomic hybridization (oaCGH) based characterization of genetic deficiencies as an aid to gene mapping in *Caenorhabditis elegans*. *BMC Genomics* **8**, 402.
- Kadandale, P., Geldziler, B., Hoffmann, M., and Singson, A. (2005). Use of SNPs to determine the break-points of complex deficiencies, facilitating gene mapping in *Caenorhabditis elegans*. *BMC Genet.* **6**, 28.
- Kalb, J. M., Lau, K. K., Goszczynski, B., Fukushige, T., and Moons, D., *et al.* (1998). pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the *C. elegans* pharynx. *Development* **125**, 2171–2180.
- Kelly, W. G., and Fire, A. (1998). Chromatin silencing and the maintenance of a functional germ line in *Caenorhabditis elegans*. *Development* **125**, 2451–2456.
- Kelly, W. G., Xu, S., Montgomery, M. K., and Fire, A. (1997). Distinct requirements for somatic and germ line expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238.
- Koch, R., van Luenen, H. G., van der Horst, M., Thijssen, K. L., and Plasterk, R. H. (2000). Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res.* **10**, 1690–1696.
- Labouesse, M. (1997). Deficiency screen based on the monoclonal antibody MH27 to identify genetic loci required for morphogenesis of the *Caenorhabditis elegans* embryo. *Dev. Dyn.* **210**, 19–32.
- Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., and Li, N., *et al.* (2004). *C. elegans* ORFeome version 3.1: Increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res.* **14**, 2064–2069.
- Lee, J., Jongeward, G. D., and Sternberg, P. W. (1994). unc-101, a gene required for many aspects of *Caenorhabditis elegans* development and behavior, encodes a clathrin-associated protein. *Genes Dev.* **8**, 60–73.
- Lieb, J. D., de Solorzano, C. O., Rodriguez, E. G., Jones, A., and Angelo, M., *et al.* (2000). The *Caenorhabditis elegans* dosage compensation machinery is recruited to X chromosome DNA attached to an autosome. *Genetics* **156**, 1603–1621.
- Luan, C. H., Qiu, S., Finley, J. B., Carson, M., and Gray, R. J., *et al.* (2004). High-throughput expression of *C. elegans* proteins. *Genome Res.* **14**, 2102–2110.
- Mah, A. K., Armstrong, K. R., Chew, D. S., Chu, J. S., and Tu, D. K., *et al.* (2007). Transcriptional regulation of AQP-8, a *Caenorhabditis elegans* aquaporin exclusively expressed in the excretory system, by the POU homeobox transcription factor CEH-6. *J. Biol. Chem.* **282**, 28074–28086.
- Maydan, J. S., Flibotte, S., Edgley, M. L., Lau, J., and Selzer, R. R., *et al.* (2007). Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res.* **17**, 337–347.
- McGhee, J. D., Fukushige, T., Krause, M. W., Minnema, S. E., and Goszczynski, B., *et al.* (2009). ELT-2 is the predominant transcription factor controlling differentiation and function of the *C. elegans* intestine, from embryo to adult. *Dev. Biol.* **327**, 551–565.
- McKay, S. J., Johnsen, R., Khattri, J., Asano, J., and Baillie, D. L., *et al.* (2003). Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 159–169.
- McKim, K. S., Heschl, M. F., Rosenbluth, R. E., and Baillie, D. L. (1988a). Genetic organization of the unc-60 region in *Caenorhabditis elegans*. *Genetics* **118**, 49–59.
- McKim, K. S., Howell, A. M., and Rose, A. M. (1988b). The effects of translocations on recombination frequency in *Caenorhabditis elegans*. *Genetics* **120**, 987–1001.
- McKim, K. S., Peters, K., and Rose, A. M. (1993). Two types of sites required for meiotic chromosome pairing in *Caenorhabditis elegans*. *Genetics* **134**, 749–768.
- McKim, K. S., and Rose, A. M. (1990). Chromosome I duplications in *Caenorhabditis elegans*. *Genetics* **124**, 115–132.

- McKim, K. S., Starr, T., and Rose, A. M. (1992). Genetic and molecular analysis of the dpy-14 region in *Caenorhabditis elegans*. *Mol. Gen. Genet.* **233**, 241–251.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Meneely, P. M., and Herman, R. K. (1979). Lethals, steriles and deficiencies in a region of the X chromosome of *Caenorhabditis elegans*. *Genetics* **92**, 99–115.
- Meneely, P. M., and Herman, R. K. (1981). Suppression and function of X-linked lethal and sterile mutations in *Caenorhabditis elegans*. *Genetics* **97**, 65–84.
- Meneely, P. M., and Wood, W. B. (1984). An autosomal gene that affects X chromosome expression and sex determination in *Caenorhabditis elegans*. *Genetics* **106**, 29–44.
- Merritt, C., and Seydoux, G. (2010). Transgenic Solutions for the Germ Line. *WormBook* **8**, 1–21.
- Mets, D. G., and Meyer, B. J. (2009). Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. *Cell* **139**, 73–86.
- Mitani, S. (2009). Nematode, an experimental animal in the national BioResource project. *Exp. Anim.* **58**, 351–356.
- Moerman, D. G., and Baillie, D. L. (1979). Genetic organization in *Caenorhabditis elegans*: Fine-structure analysis of the unc-22 gene. *Genetics* **91**, 95–103.
- Moerman, D. G., and Baillie, D. L. (1981). Formaldehyde mutagenesis in the nematode *Caenorhabditis elegans*. *Mutat. Res.* **80**, 273–279.
- Moerman, D. G., and Barstead, R. J. (2008). Towards a mutation in every gene in *Caenorhabditis elegans*. *Brief Funct. Genomic Proteomic* **7**, 195–204.
- Muller, H. J. (1918). Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. *Genetics* **3**, 422–499.
- NimbleGenSystemsInc., NimbleGenSystemsInc. [www.nimblegen.com]
- Phillips, C. M., Meng, X., Zhang, L., Chretien, J. H., and Urnov, F. D., *et al.* (2009). Identification of chromosome sequence motifs that mediate meiotic pairing and synapsis in *C. elegans*. *Nat. Cell Biol.* **11**, 934–942.
- Plenefisch, J. D., DeLong, L., and Meyer, B. J. (1989). Genes that implement the hermaphrodite mode of dosage compensation in *Caenorhabditis elegans*. *Genetics* **121**, 57–76.
- Poteryaev, D., Fares, H., Bowerman, B., and Spang, A. (2007). *Caenorhabditis elegans* SAND-1 is essential for RAB-7 function in endosomal traffic. *EMBO J.* **26**, 301–312.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Qadota, H., Inoue, M., Hikita, T., Koppen, M., and Hardin, J. D., *et al.* (2007). Establishment of a tissue-specific RNAi system in *C. elegans*. *Gene* **400**, 166–173.
- Reese, K. J., Dunn, M. A., Waddle, J. A., and Seydoux, G. (2000). Asymmetric segregation of PIE-1 in *C. elegans* is mediated by two complementary mechanisms that act through separate PIE-1 protein domains. *Mol. Cell* **6**, 445–455.
- Robert, V. J., and Bessereau, J. L. (2010). Manipulating the *Caenorhabditis elegans* genome using mariner transposons. *Genetica* **138**(5), 541–549.
- Robert, V. J., Katic, I., and Bessereau, J. L. (2009). Mos1 transposition as a tool to engineer the *Caenorhabditis elegans* genome by homologous recombination. *Methods* **49**, 263–269.
- Rogalski, T. M., Moerman, D. G., and Baillie, D. L. (1982). Essential genes and deficiencies in the unc-22 IV region of *Caenorhabditis elegans*. *Genetics* **102**, 725–736.
- Rogalski, T. M., and Riddle, D. L. (1988). A *Caenorhabditis elegans* RNA polymerase II gene, ama-1 IV, and nearby essential genes. *Genetics* **118**, 61–74.
- Rose, A. M., and Baillie, D. L. (1979). A mutation in *Caenorhabditis elegans* that increases recombination frequency more than threefold. *Nature* **281**, 599–600.
- Rose, A. M., Baillie, D. L., and Curran, J. (1984). Meiotic pairing behavior of two free duplications of linkage group I in *Caenorhabditis elegans*. *Mol. Gen. Genet.* **195**, 52–56.

- Rose, A. M., O'Neil, N. J., Bilenky, M., Butterfield, Y. S., and Malhis, N., *et al.* (2010). Genomic sequence of a mutant strain of *Caenorhabditis elegans* with an altered recombination pattern. *BMC Genomics* **11**, 131.
- Rosenbluth, R. E., and Baillie, D. L. (1981). The genetic analysis of a reciprocal translocation, eTI(III;V), in *Caenorhabditis elegans*. *Genetics* **99**, 415–428.
- Rosenbluth, Raja E., Cuddeford, C., and Baillie, D. L. (1983). Mutagenesis in *Caenorhabditis elegans*: I. A rapid eukaryotic mutagen test system using the reciprocal translocation, eTI(III;V). *Mut. Res./Fundam. Molec. Mech. Mutagenesis* **110**, 39–48.
- Rosenbluth, R. E., Cuddeford, C., and Baillie, D. L. (1985). Mutagenesis in *Caenorhabditis elegans* II. A spectrum of mutational events induced with 1500 r of gamma-radiation. *Genetics* **109**, 493–511.
- Rosenbluth, R. E., Johnsen, R. C., and Baillie, D. L. (1990). Pairing for recombination in LGV of *Caenorhabditis elegans*: A model based on recombination in deficiency heterozygotes. *Genetics* **124**, 615–625.
- Rual, J. F., Hill, D. E., and Vidal, M. (2004). ORFeome projects: Gateway between genomics and omics. *Curr. Opin. Chem. Biol.* **8**, 20–25.
- Sarov, M., Schneider, S., Pozniakovski, A., Roguev, A., and Ernst, S., *et al.* (2006). A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nat. Methods* **3**, 839–844.
- Selzer, R. R., Richmond, T. A., Pofahl, N. J., Green, R. D., and Eis, P. S., *et al.* (2005). Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosomes Cancer* **44**, 305–319.
- Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., and Priess, J. R., *et al.* (1996). Repression of gene expression in the embryonic germ lineage of *C. elegans*. *Nature* **382**, 713–716.
- Siddiqui, S. S., and Babu, P. (1980). Genetic mosaics of *Caenorhabditis elegans*: A tissue-specific fluorescent mutant. *Science* **210**, 330–332.
- Sigurdson, D. C., Spanier, G. J., and Herman, R. K. (1984). *Caenorhabditis elegans* deficiency mapping. *Genetics* **108**, 331–345.
- Stewart, H. I., O'Neil, N. J., Janke, D. L., Franz, N. W., and Chamberlin, H. M., *et al.* (1998). Lethal mutations defining 112 complementation groups in a 4.5 Mb sequenced region of *Caenorhabditis elegans* chromosome III. *Mol. Gen. Genet.* **260**, 280–288.
- Stewart, H. I., Rosenbluth, R. E., and Baillie, D. L. (1991). Most ultraviolet irradiation induced mutations in the nematode *Caenorhabditis elegans* are chromosomal rearrangements. *Mutat. Res.* **249**, 37–54.
- Strome, S., Powers, J., Dunn, M., Reese, K., and Malone, C. J., *et al.* (2001). Spindle dynamics and the role of gamma-tubulin in early *Caenorhabditis elegans* embryos. *Mol. Biol. Cell* **12**, 1751–1764.
- Swan, K. A., Curtis, D. E., McKusick, K. B., Voinov, A. V., and Mapa, F. A., *et al.* (2002). High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res.* **12**, 1100–1105.
- Terns, R. M., Kroll-Conner, P., Zhu, J., Chung, S., and Rothman, J. H. (1997). A deficiency screen for zygotic loci required for establishment and patterning of the epidermis in *Caenorhabditis elegans*. *Genetics* **146**, 185–206.
- Tursun, B., Cochella, L., Carrera, I., and Hobert, O. (2009). A toolkit and robust pipeline for the generation of fosmid-based reporter genes in *C. elegans*. *PLoS One* **4**, e4625.
- Villeneuve, A. M. (1994). A *cis*-acting locus that promotes crossing over between X chromosomes in *Caenorhabditis elegans*. *Genetics* **136**, 887–902.
- Voutev, R., and Hubbard, E. J. (2008). A “FLP-Out” system for controlled gene expression in *Caenorhabditis elegans*. *Genetics* **180**, 103–119.
- Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H., and Plasterk, R. H. (2001). Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* **28**, 160–164.
- Wilm, T., Demel, P., Koop, H. U., Schnabel, H., and Schnabel, R. (1999). Ballistic transformation of *Caenorhabditis elegans*. *Gene* **229**, 31–35.
- Yochem, J., Sundaram, M., and Bucher, E. A. (2000). Mosaic analysis in *Caenorhabditis elegans*. *Methods Mol. Biol.* **135**, 447–462.
- Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., and Copeland, N. G., *et al.* (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**, 5978–5983.

- Yuan, J. Y., and Horvitz, H. R. (1990). The *Caenorhabditis elegans* genes *ced-3* and *ced-4* act cell autonomously to cause programmed cell death. *Dev. Biol.* **138**, 33–41.
- Zetka, M. C., and Rose, A. M. (1992). The meiotic behavior of an inversion in *Caenorhabditis elegans*. *Genetics* **131**, 321–332.
- Zhao, Y., Lai, K., Cheung, I., Youds, J., and Tarailo, M., *et al.* (2006). A mutational analysis of *Caenorhabditis elegans* in space. *Mutat. Res.* **601**, 19–29.
- Zhao, Y., Tarailo-Graovac, M., O’Neil, N. J., and Rose, A. M. (2008). Spectrum of mutational events in the absence of DOG-1/FANCI in *Caenorhabditis elegans*. *DNA Repair (Amst.)* **7**, 1846–1854.
- Zhao, Z., Fang, L. L., Johnsen, R., and Baillie, D. L. (2004a). ATP-binding cassette protein E is involved in gene transcription and translation in *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **323**, 104–111.
- Zhao, Z., Flibotte, S., Murray, J. I., Blick, D., and Boyle, T. J., *et al.* (2010). New tools for investigating the comparative biology of *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Genetics* **184**(3), 853–863.
- Zhao, Z., Sheps, J. A., Ling, V., Fang, L. L., and Baillie, D. L. (2004b). Expression analysis of ABC transporters reveals differential functions of tandemly duplicated genes in *Caenorhabditis elegans*. *J. Mol. Biol.* **344**, 409–417.

## Glossary

- Allele:** a variant of a gene
- aCGH:** oligonucleotide array comparative genome hybridization, aka oaCGH or arrayCGH
- Alignment:** positioning of two homologs in close proximity of each other
- Aneuploid:** alteration in copy number, more or less than two copies, or major portion of, a chromosome
- Backcross:** crossing a mutant phenotype to wild-type
- Balancer:** genetic construct that reduces crossing over allowing maintenance of the heterozygote
- CGC:** *Caenorhabditis* Genetics Centre, in Minneapolis, MN
- Cis-heterozygote:** mutations on the same homolog
- Conditional:** phenotype is dependent on certain conditions, for example, temperature
- Complex rearrangement:** more than one DNA rearrangement
- Crossing Over:** physical breakage and reunion of homologous DNA
- Crossover suppressor:** prevents crossing over between homologs
- Deficiency:** a deletion that removes a copy of one or more adjacent genes
- Deletion:** one copy of a portion of the DNA
- Duplication:** an extra copy of a portion of DNA
- EMS:** ethyl methane sulfonate, a commonly used chemical mutagen
- Enhancer:** strengthening of the mutant phenotype
- Essential gene:** the gene whose function is required for survival
- FISH:** Fluorescent in situ hybridization
- Gene:** a unit of DNA encoding an RNA
- Gene conversion:** biological substitution of one allelic form for another in a heterozygote
- Gene interaction:** phenotype produced by two or more mutant phenotypes
- Genetic map:** map based on crossover distances between phenotypic markers
- Genetic nomenclature:** rules for writing phenotypes, genotypes, and gene products
- Genotype:** description of a gene
- GFP:** green fluorescent protein
- Hermaphrodite:** individual producing both sperm and oocytes
- Heterozygote:** two different allelic copies of a gene
- Homolog:** copy of a chromosome
- Homozygote:** two identical copies of a gene
- Him:** high incidence of males

**HRR:** homolog recognition region

**Integration:** extraneous DNA that inserts into a chromosome

**Inversion:** a region of DNA that is “inverted”, that is, turned end for end

**Isolate:** single individual used to establish a strain

**Lab designations:** Two letter registered descriptor of each *C. elegans* lab, available at WormBase

**Lethal:** a mutation that cannot be propagated as a homozygote over successive generations

**Linkage:** on the same DNA molecule

**Male sperm:** sperm from male *C. elegans* as opposed to hermaphroditic sperm

**Maternal Effect Lethal:** an adult that produces infertile progeny

**Mutation:** a rare alteration in the DNA sequence

**Mutagen:** chemical or physical source that causes changes in the DNA sequence

**Mutagen dosage:** amount of mutagen applied and duration of exposure

**Morphological Marker:** a phenotype that alters the body structure in a visible way

**Nonconditional:** phenotype is independent of external conditions

**Nondisjunction:** failure of two homologs to separate two different cells during meiosis

**Ortholog:** shared DNA sequence identity or similarity between species

**Out-cross:** fertilization of hermaphrodite oocytes by male sperm

**Paralog:** shared DNA sequence identity or similarity within a species

**PCR:** polymerized chain reaction

**Phenotype:** recognizable feature of the animal resulting from its genetic composition

**Physical map:** map based on the DNA sequence

**Primers:** DNA oligos used to start the PCR amplification process

**Punnet Square:** representation of haploid gametic forms and the resulting diploid progeny

**Promoter:** binding site for DNA polymerase and accessory proteins

**Promoter::GFP:** promoter sequence linked to a green fluorescence protein reporter

**Radiation:** physical mutagen in the form of energy causing changes in the DNA, for example, gamma, X-ray, UV

**Rearrangement:** a reorganization of the genetic material

**Reciprocal translocation:** rearrangements of two chromosomes such that no genetic material is lost

**Recombination:** biological reorganization of homologous chromosomes

**Rescue:** return to wild-type phenotype

**RNAi:** interference of gene expression by double-stranded RNA

**Screening:** process of identifying mutant phenotypes

**Self-fertilization:** fertilization, in a hermaphrodite, of oocytes by its own sperm

**Semidominant:** phenotype recognizable as a heterozygote

**SNP:** single nucleotide polymorphism (between two populations)

**Snip-SNP:** single nucleotide difference resulting in alteration of a restriction digest cut site

**Spontaneous mutation:** alteration of the DNA sequence in the absence of a known mutagen

**Stability:** frequency with which a genotype alters

**Strain construction:** crosses to bring together genotypes to be maintained as a heritable strain

**Sterile Adult:** an adult that does not produce fertilized oocytes

**Synapsis:** formation of the synaptonemal complex during meiotic prophase

**Suppressor:** lessening of the mutant phenotype

**Telomere:** specialized sequence at the end of a linear chromosome

**Terminal deletion:** deletion of DNA-removing sequences adjacent to the telomere

**Transgenic array:** introduced concatenated DNA that is transmitted heritably

**Transposon:** region of DNA capable of movement from one site to another

**Trans-heterozygote:** mutations on each of the homologs

**Translocation:** rearrangement of DNA from one region of the genome to another

**Transgenes:** introduced DNA transferred from outside the animal

**Wild type:** accepted nonmutant form of the animal or a gene

**WormBase:** database of information about *C. elegans*, [www.wormbase.org](http://www.wormbase.org)

---

---

## CHAPTER 3

# Genome Engineering by Transgene- Instructed Gene Conversion in *C. elegans*

**Valérie J.P. Robert**<sup>\*,†,‡,¶</sup> and **Jean-Louis Bessereau**<sup>\*,†,‡</sup>

<sup>\*</sup>Ecole Normale Supérieure, Institut de Biologie de l'ENS, IBENS, Paris, France

<sup>†</sup>INSERM U1024, Paris, France

<sup>‡</sup>CNRS, UMR 8197, Paris, France

<sup>¶</sup>Present address: Laboratory of Molecular and Cellular Biology, UMR5239 CNRS, Ecole Normale Supérieure de Lyon, Lyon, France

---

### Abstract

- I. Introduction
    - A. Transgenesis in *C. elegans*
    - B. Genome Engineering by Homologous Recombination in *C. elegans*
  - II. General Features of the *Mos*TIC Technique and Experimental Outline
  - III. Methods
    - A. Identification of a *Mos1* Insertion in the Target Region
    - B. Construction of Transgenic Lines
    - C. Heat-Shock Induction of *Mos1* Excision and Pooling of Heat-Shocked Animals
    - D. Screening
    - E. An Alternative Protocol Based on the Constitutive Expression of the *Mos* Transposase
    - F. An Alternative Protocol for *Mos*TIC-Engineered KO/del Alleles Selection
    - G. Frequently Asked Questions
  - IV. Materials
    - A. Equipment
    - B. Solutions and Reagents
    - C. Plasmids
  - V. Discussion: *Mos*TIC Applications and Derivatives
    - A. Gene Function Analysis Using *Mos*TIC-Engineered Alleles
    - B. Single-Copy Transgene Genomic Integration by *Mos*SCI
  - VI. Summary
- Acknowledgment  
References

---

---

---

## Abstract

The nematode *Caenorhabditis elegans* is an anatomically simple metazoan that has been used over the last 40 years to address an extremely wide range of biological questions. One major advantage of the *C. elegans* system is the possibility to conduct large-scale genetic screens on randomly mutagenized animals, either looking for a phenotype of interest and subsequently relate the mutated gene to the biological process under study (“forward genetics”), or screening for molecular lesions impairing the function of a specific gene and later analyze the phenotype of the mutant (“reverse genetics”). However, the nature of the genomic lesion is not controlled in either strategy. Here we describe a technique to engineer customized mutations in the *C. elegans* genome by homologous recombination.

This technique, called *MosTIC* (for *Mos1* excision induced transgene-instructed gene conversion), requires a *C. elegans* strain containing an insertion of the *Drosophila* transposon *Mos1* within the locus to modify. Expression of the *Mos* transposase in the germ line triggers *Mos1* excision, which causes a DNA double strand break (DSB) in the chromosome at the excision site. The DSB locally stimulates DNA repair by homologous recombination, which can sometimes occur between the chromosome and a transgene containing sequence homologous to the broken locus. In that case, sequence variations contained in the repair template will be copied by gene conversion into the genome. Here we provide a detailed protocol of the *MosTIC* technique, which can be used to introduce point mutations and generate knockout and knock-in alleles.

---

---

---

## I. Introduction

Gene knockout (KO/del) and knock-in (KI) have emerged as essential components of the genetic toolbox used to study gene function in model organisms. In yeast and mouse, techniques have been developed in the eighties to engineer chromosomal loci (Doetschman *et al.*, 1987; Scherer and Davis, 1979; Thomas and Capecchi, 1987); these rely on the recombination between the chromosome and a transgenic DNA fragment carrying the sequence to introduce into the genome flanked by sequences homologous to the targeted locus. Positive and negative selection markers are used to identify recombination events and select against random insertion into the genome. In the nematode *Caenorhabditis elegans*, however, the low frequency of spontaneous recombination between the genome and exogenous DNA has impeded for a long time the establishment of an equivalent strategy. Therefore, most of the data generated in *C. elegans* to characterize gene expression patterns or to define the subcellular localization of tagged proteins relied on the use of transgenes made of repetitive extrachromosomal arrays.

### A. Transgenesis in *C. elegans*

The initial technique of transgenesis developed in *C. elegans* is based on DNA microinjection in the syncytial part of the germ line (Evans, 2006; Mello and Fire, 1995; Mello *et al.*, 1991; Stinchcomb *et al.*, 1985). Injected DNA fragments are randomly concatemerized in the germ line, resulting in the formation of large arrays, which contain hundreds of copies of the injected DNA. These arrays remain extrachromosomal and are replicated during cell division. This transgenesis technique is extremely simple and efficient, yet it has some intrinsic limitations. First, genes contained in repetitive transgenic sequences are often overexpressed in the soma, sometimes generating toxic effects and experimental artifacts such as nonphysiological expression patterns. Second, repetitive sequences present in extrachromosomal arrays are, most of the time, silenced in the germ line (Kelly *et al.*, 1997). In addition, their presence triggers an RNA silencing process known as cosuppression that induces the silencing of the homologous endogenous locus (Dernburg *et al.*, 2000; Ketting and Plasterk, 2000; Robert *et al.*, 2005). Cosuppression phenocopies loss-of-function alleles of germ line-expressed genes, often resulting in animal sterility and preventing the use of extrachromosomal arrays to study germ line-expressed genes in *C. elegans*. Third, extrachromosomal arrays are not stably inherited during mitosis and meiosis. As a consequence, transgenic animals are mosaic and transgenic siblings can exhibit different phenotypes depending on how many cells contain the extrachromosomal array. This problem can be solved by randomly integrating the repetitive transgenes in the genome. Radiations or chemicals are used to cause random double strand breaks (DSBs), thereby stimulating DNA recombination between the transgene and the genome. The transgenic population is subsequently screened for animals that segregate the transgenes in a Mendelian fashion (Evans, 2006). Fourth, structural rearrangements can occur over time in the extrachromosomal array, hence causing variations of transgene expression patterns and levels among different generations of the same transgenic line.

A second technique of transgenesis was developed based on the bombardment of the germ line with DNA-coated gold particles (Evans, 2006; Green *et al.*, 2008; Praitis *et al.*, 2001; Wilm *et al.*, 1999). About 10% of the transformants obtained with this “biolistic” transformation technique are unique low-copy chromosomal insertions that can be readily identified by the use of a visible counterselection marker (Vazquez-Manrique *et al.*, 2010). Biolistic has been used to generate *C. elegans* strains stably expressing transgenes both in the germ line (see for examples Cheeseman *et al.*, 2004; Merritt *et al.*, 2008; Sijen and Plasterk, 2003) and the soma (see for examples Berset *et al.*, 2005; Praitis *et al.*, 2005). Results obtained with such transgenes are thought to be more physiologically relevant than the ones obtained with repetitive extrachromosomal arrays (Praitis *et al.*, 2001).

## B. Genome Engineering by Homologous Recombination in *C. elegans*

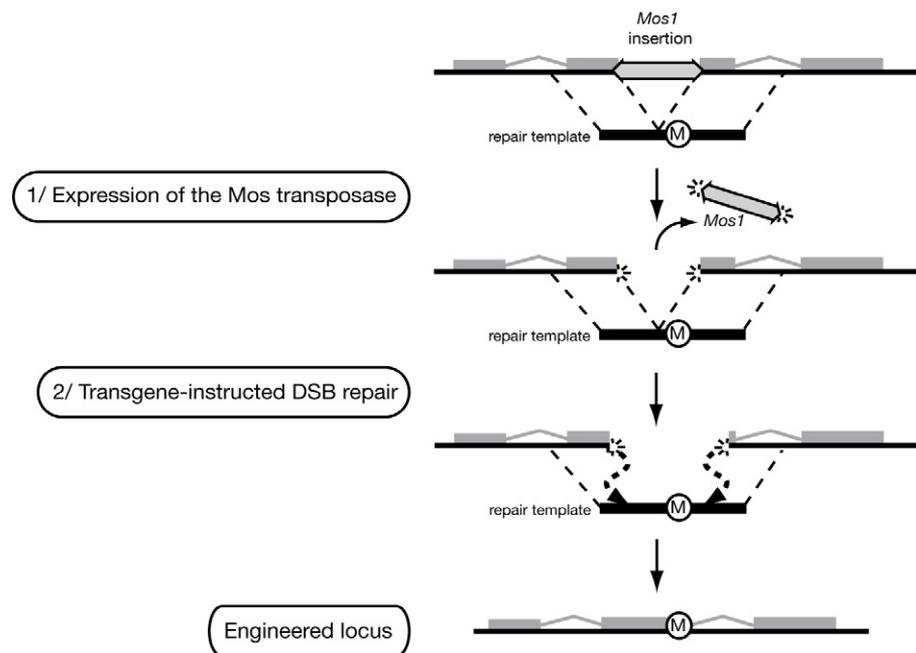
Over the last 20 years, several attempts were made to develop a genome engineering technique using different strategies to promote the recombination between the genome and a homologous transgene.

First, it was hypothesized that extrachromosomal arrays made by DNA microinjection in the germ line syncytium were bad recombination substrates because of their repetitive structure (Broverman *et al.*, 1993). Therefore alternative ways of introducing DNA in the germ line were tested. Engineered DNA fragments were directly injected into the nuclei of meiotic oocytes. Genomic integration of the injected DNA was identified in the progeny of the injected animals. Out of the recombination events, 3% (2 out of 63) resulted from homologous recombination between the injected DNA and the genome. However, this technique was not further developed, mostly because of the difficulty to inject DNA into meiotic oocyte nuclei. Similarly, DNA sequences introduced in *C. elegans* by biolistic were shown to recombine at low frequency with homologous genomic sequences (Berezikov *et al.*, 2004; Jantsch *et al.*, 2004). The use of a visible counterselection marker was later proved to be very useful to identify rare homologous recombination events among all transformants (Vazquez-Manrique *et al.*, 2010). However, the homologous recombination frequency obtained with this method is low. About one recombination event was obtained among 300 transformants derived from 30 independent bombardment experiments. In a laboratory where biolistic transformation is already established, performing and screening 30 independent bombardment experiments will take approximately 2 months of full-time bench work and additional efforts will still be required to identify *bona fide* recombinants.

A second set of strategies was using endogenous *C. elegans* DNA transposons of the *Tc* family. DNA transposons move by a “cut-and-paste” mechanisms: they encode a transposase that binds the terminal ends of the transposon and catalyzes the excision and reinsertion of the transposon DNA, leaving at the excision site a DSB that must be repaired by the cellular machinery. It was demonstrated that such DSB can sometimes be repaired by homology-dependent recombination between the broken chromosome and DNA provided in *trans* as a repair template (Gloor *et al.*, 1991; Plasterk, 1991; Robert *et al.*, 2008). This provided a means to copy sequence variations from the repair template into a genomic target region. The feasibility of such strategy was demonstrated in *C. elegans* by Plasterk and Groenen (Plasterk and Groenen, 1992) and later revisited to establish a genome engineering protocol (Barrett *et al.*, 2004). However, the use of endogenous transposons has a number of disadvantages. Because these elements are not mobile in the germ line of wild-type animals, transposon mobilization is achieved in genetic backgrounds, known as *mutators*, where germ line transposition is derepressed. As a result, transposons accumulate in the genome of these *mutator* strains, causing uncontrolled mutations resulting in a high morbidity of the strains. In addition, the frequency of homologous recombination events remains modest, necessitating the growth and screening of large populations of animals in which the insertion of interest is unstable.

To circumvent these problems, we used the heterologous DNA transposon *Mos1* to induce DSBs in the *C. elegans* genome (Robert and Bessereau, 2007; Robert *et al.*, 2008). *Mos1* was initially isolated in *D. mauritiana* (Hartl, 2001; Jacobson and Hartl, 1985; Jacobson *et al.*, 1986) and subsequently mobilized in the *C. elegans* germ line (Bessereau *et al.*, 2001). Briefly, it was shown that *Mos1* copies provided in an extrachromosomal transgene could insert into the genome when expressing the *Mos* transposase in the germ line. It generates a small number of insertions (on average 2.5 insertions per genome) that are stable in the absence of the *Mos* transposase (Williams *et al.*, 2005). Such insertions are efficiently remobilized in the presence of the *Mos* transposase, generating DSBs that are preferentially repaired by homologous recombination (Robert *et al.*, 2008). Based on these observations, we developed a genome engineering technique called *MosTIC* (for *Mos1*-excision transgene-instructed gene conversion; Fig. 1) (Robert and Bessereau, 2007) and optimized efficient protocols (Robert *et al.*, 2009).

Side-by-side comparison of results obtained with strains carrying repetitive extrachromosomal arrays and KI strains generated by *MosTIC* demonstrates that genomic



**Fig. 1** Transgene-instructed gene conversion in *C. elegans*. A targeted DNA double strand break (DSB) is created in the genome by triggering the excision of the DNA transposon *Mos1* (step 1). A transgene containing a mutation “M” flanked by DNA sequence homologous to the broken locus can be used as a repair template (step 2). Gene conversion results in the introduction of the mutation in the chromosome.

engineering avoids overexpression artifacts and provides more physiologically relevant results (see “Discussion” for an example). Here, we provide a detailed description of the *Mos*TIC protocol routinely used in our laboratory and we discuss an alternative protocol that we have established.

## II. General Features of the *Mos*TIC Technique and Experimental Outline

An overview of the *Mos*TIC procedure is depicted in Fig. 2A and can be summarized in five points:

- (i) A *MosI* insertion is identified in the region of interest.
- (ii) A repair template containing the desired modification is designed.
- (iii) Transgenic lines containing the selected *MosI* insertion are generated with the repair template and a vector providing the expression of the *Mos* transposase under the control of a heat-shock inducible promoter.
- (iv) *MosI* excision and *Mos*TIC are triggered by a heat-shock treatment.
- (v) *Mos*TIC-engineered alleles are identified using either PCR or phenotypic reversion to screen the progeny of heat-shocked animals.

Using this protocol, it is possible to engineer point mutation alleles, KO/del alleles and KI alleles expressing tagged protein versions (Fig. 2B). *Mos*TIC alleles engineered at the *MosI* insertion point are recovered with frequencies ranging from  $10^{-3}$  to  $10^{-5}$  events per offspring. This frequency decreases when the modification to introduce into the genome is located further away of the *MosI* insertion point. Characterization of a *Mos*TIC conversion tract was performed at the *unc-5* locus (Robert and Bessereau, 2007 and Fig. 2C) using a repair template containing multiple silent polymorphisms. Point mutations localized in a 1 kb-long fragment centered at the *MosI* insertion point were copied in at least 50% of the *Mos*TIC alleles. The recombination efficiency was decreased 20 times 1.5 kb away from the *MosI* insertion point.

Using this standard protocol, we have been able to generate seven different *Mos*TIC KI alleles at five independent loci of the *C. elegans* genome (Gendrel *et al.*, 2009; Robert and Bessereau, 2007; and V. Robert, T. Boulin, C. Stigloher, H. Tu and J.L. Bessereau, unpublished results). Tags have been inserted at distances varying from 20 to 800 bp of the *MosI* insertion point and estimated *Mos*TIC efficiencies were similar to the ones observed for point mutation engineering.

When it is possible to screen for an engineered allele based on a visible phenotype, it takes only 2 weeks after heat-shock induction of the *Mos* transposase expression to isolate a strain containing the *Mos*TIC allele. When using a PCR-based strategy, several rounds of sibling selection are required to isolate single modified animals and it takes about 1 month after the initial heat-shock to isolate the *Mos*TIC allele.



---

---

### III. Methods

#### A. Identification of a *Mos1* Insertion in the Target Region

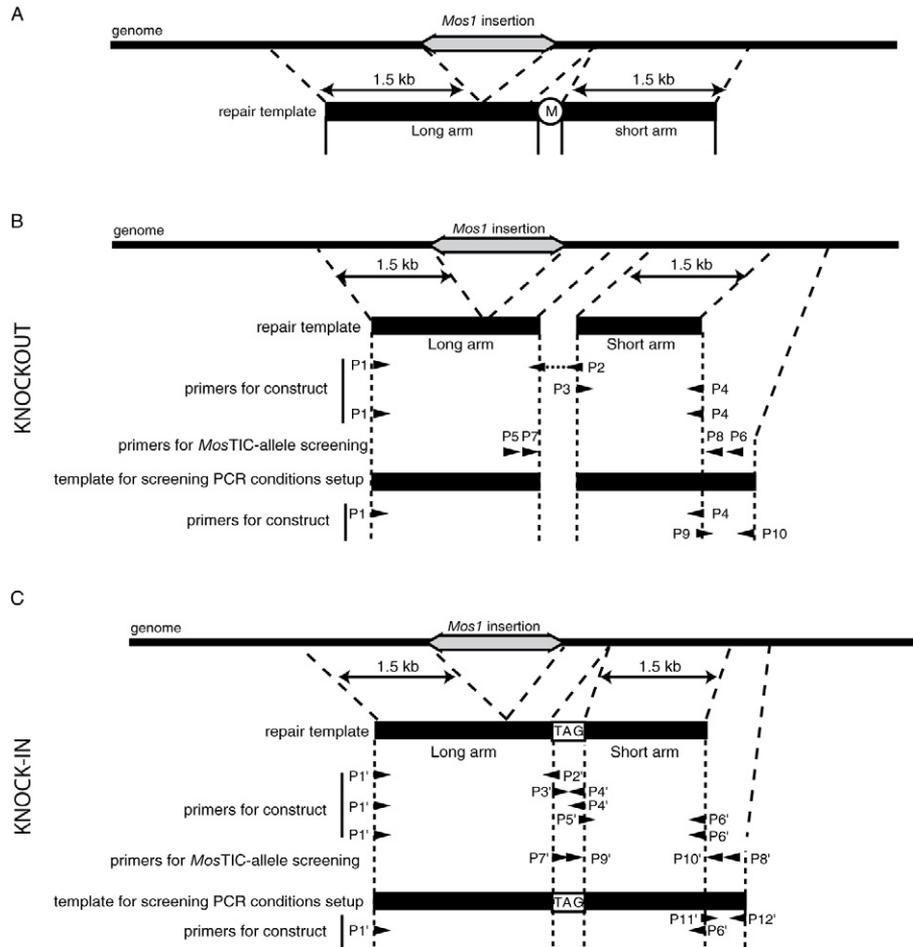
*MosTIC* requires an entry strain containing a *Mos1* insertion preferentially located at 500 bp or less from the target site. A *Mos1* insertion library has been generated by the Nemagenetag consortium (Bazopoulou and Tavernarakis, 2009; Duverger *et al.*, 2007). It contains 13,845 independent insertions, which provide entry points for manipulating about 10% of the genome. Mapped insertions are annotated in wormbase (<http://www.wormbase.org/>) and *C. elegans* stocks carrying these insertions are distributed upon request ([http://www.cgmc.univ-lyon1.fr/cgmc\\_info\\_celeganstp.php](http://www.cgmc.univ-lyon1.fr/cgmc_info_celeganstp.php)). Alternatively, strains containing *Mos1* insertions in a gene of interest can be generated by performing its own *Mos1* mutagenesis (Williams *et al.*, 2005) as described in the following protocols (Bessereau, 2006; Boulin and Bessereau, 2007). *Mos1* mutagenesis strategy can be extremely useful to identify *Mos1* insertions in several genes involved in the same biological process and further use the *Mos1* alleles as entry points for genomic engineering.

#### B. Construction of Transgenic Lines

##### 1. Building the Repair Template

Standard repair templates designed to engineer *MosTIC* KO/del or KI alleles are described in Fig. 3. A standard *MosTIC* repair template (Fig. 3A) contains the modifications to introduce into the genome flanked by one long arm and one short arm of genomic sequences. The long arm (left arm on Fig. 3) contains the genomic sequence located between the point where modifications have to be introduced in the locus and the *Mos1* insertion point and an additional 1.5-kb long genomic fragment flanking the left side of the *Mos1* insertion point. The short arm (right arm on Fig. 3) contains a 1.5-kb long genomic fragment at the right side of the genomic point where the modifications have to be introduced. The repair template is preferentially made using PCR fusion-based strategies (Hobert, 2002) (Fig. 3) but any molecular biology techniques can be used. Next, it is cloned into a standard plasmid and sequenced over its full-length before injection.

For the construction of “KO/del” repair templates, independent PCR are performed with primers P1–P2 and P3–P4 (Fig. 3B) on genomic DNA extracted from wild-type animals. P2 is 44 bp long and overlaps the region that has to be deleted. Its 3' end contains a 20 nucleotide-long sequence complementary to the segment that flanks the deletion on its left side and its 5' end contains a 24 nucleotide-long sequence complementary to the segment that flanks the deletion on its right side. This sequence will hybridize with the end of the P3–P4 product during the PCR fusion reaction, which is performed using primers P1 and P4 and an equimolar mix of P1–P2 and P3–P4 PCR products. Amplification can be performed using the Phusion high-fidelity DNA Polymerase (Finnzymes) or any other high-fidelity



**Fig. 3** (A) Description of a standard *Mos1* repair template. (B) Repair templates and control plasmids for the engineering and PCR screening of *Mos1* KO/del and KI alleles. Primers required for construction of the templates and PCR screening are described in the text.

polymerase. If necessary, a gradient of annealing temperature ranging from 55 to 68°C can be performed to optimize the amplification of individual and fused fragments. After amplification, PCR products are purified by gel extraction using the Qiaquick Gel Extraction Kit (Qiagen) according to the instructions of the manufacturer. The final PCR product corresponding to the full-length repair template is subcloned using the Zero Blunt PCR cloning kit (Invitrogen) or any other cloning kit and sequenced.

“KI” repair templates are synthesized in two successive rounds of PCR fusion. First, primers P1' and P2' (Fig. 3C) are used on wild-type genomic DNA to amplify the long arm of the repair template and P3' and P4' are used to amplify the tag sequence using an appropriate PCR template. P2' contains in its 5' end a tail of 24 nucleotides, which overlaps with the end of the tag sequence. PCR fusion is performed using primers P1' and P4' and an equimolar mix of P1'–P2' and P3'–P4' PCR products. The short arm of the repair template is amplified on wild-type genomic DNA using P5' and P6'. P5' contains in its 5' end, a tail of 24 nucleotides, which overlaps with the end of the tag sequence. PCR fusion is performed by mixing equimolar quantities of P1'–P4' and P5'–P6' PCR products and using primers P1' and P6'.

Repair templates for point mutations can be constructed either by PCR fusion-based strategy with primers containing the mutations or by site-directed mutagenesis of a subcloned genomic fragment using, for example, the QuickChange II kit (Stratagene).

Several mechanisms are at work in the *C. elegans* germ line to repair a *Mos1*-triggered DSB (Robert *et al.*, 2008) and sometimes regenerate sequences encoding a functional protein. When screening *Mos1*-engineered alleles by phenotypic reversion, it might be useful to be able to quickly distinguish between revertants generated by transgene-instructed gene conversion from those generated by other mechanisms. In that case, we recommend to introduce a silent restriction site in the repair template close to the DSB site (see Robert and Bessereau, 2007 for an example).

## 2. Establishment of Transgenic Lines

- (a) Use standard *C. elegans* germ line microinjection procedure (Evans, 2006; Mello and Fire, 1995) to generate extrachromosomal arrays carrying both the repair template and a heat-shock inducible *Mos* transposase expression vector in a genetic background homozygous for the *Mos1* insertion to mobilize. Into a strain homozygous for the *Mos1* insertion of interest, inject the following mix:

- the repair template at 50 ng/μL.
- *pJL44* (*Phsp-16.48::MosTase*) at 50 ng/μL.
- *pPD118.33* (*Pmyo-2::GFP*) at 5 ng/μL as a transformation marker.

Inject about 20 animals and keep them at 20°C prior to transgenic F1 screening.

- (b) Isolate individual F1 transgenic animals with expression of GFP in the pharynx and screen their progeny to identify transgenic lines. Usually, 10–30% of the F1 transgenic animals give transgenic lines. Select five healthy lines derived from different F1 transgenic animals, with an intermediate to high transgene transmission rate (>50% transmission of the transgenic array).
- (c) Amplify the transgenic populations by maintaining the transgenic lines at 25°C, if possible, in order to minimize potential transgene silencing.

### C. Heat-Shock Induction of *Mos1* Excision and Pooling of Heat-Shocked Animals

- (a) Select nematode growth medium (NGM) plates (Stiernagle, 2006) containing about 200 young transgenic adults from one to five transgenic lines, seal with parafilm and immerse for 1 h in a water bath setup at 33°C.
- (b) Let the worms recover for 1 h at 15°C.
- (c) Immerse for 1 h in a water bath setup at 33°C. Remove the parafilm.
- (d) After one night at 20°C, transfer heat-shocked animals to fresh plates. Depending on the fertility of the heat-shocked transgenic animals, put one to five animals on the same plate in order to obtain 100 F1 animals in each pool. To calibrate this step, we recommend to heat-shock few transgenic animals and estimate their brood size before starting the *MosTIC* experiment.
- (e) Before screening, roughly estimate the F1 population size to be able later to calculate the *MosTIC* efficiency.

Typically, 50–100 pools are screened for a single heat-shock induction experiment. If *MosTIC* occurs with a frequency of  $5 \times 10^{-4}$  events per offspring in the progeny of the heat-shocked animals, the probability of recovering at least one *MosTIC* allele when screening 10,000 progeny of heat-shocked animals is 99%. If *MosTIC* efficiency is lower, we recommend to repeat several heat-shock induction experiments of the same size rather than increasing the size of a single induction experiment.

### D. Screening

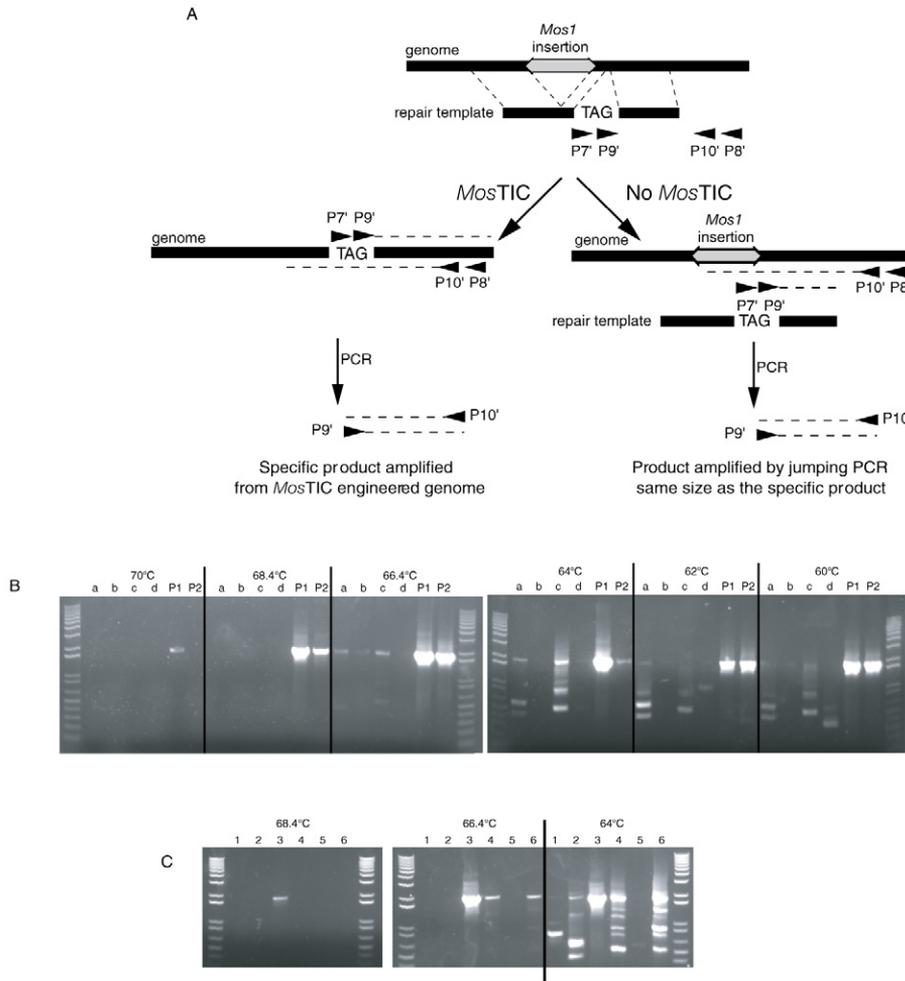
#### 1. PCR Screening

In most cases, identification of *MosTIC* KO/del and KI alleles relies on PCR screening, which is performed in two successive rounds of nested PCR. To facilitate the handling of a large amount of samples, we recommend to work with multichannel pipettes and repeating dispenser using 96-well PCR plates.

##### (a) *Primer Design and “Jumping PCR” (Fig. 4)*

For each primer pair, one primer is located inside the repair template and the other one outside. For KO/del alleles, the “inside” primers recognize the repair template on its left/long arm (primers P5 and P7 on Fig. 3). For KI alleles, the “inside” primers recognize the repair template in the tag sequence (primers P7' and P9' on Fig. 3). In both case, the “outside” primers recognize genomic sequences to the right of the short arm.

One common pitfall of this strategy arises from PCR fragments having the same size as the specific PCR product amplified in the presence of a *MosTIC*-engineered allele and sometimes amplified from transgenic animals that do not contain *MosTIC*-engineered alleles (Fig. 4A). This artifact arises by “jumping PCR” (Paabo *et al.*, 1990) between single-stranded DNA generated from the transgene on the one hand and the genome on the other hand (Fig. 4A and B). Jumping PCR is dependent on the nucleotidic composition of the amplified



**Fig. 4** Jumping PCR. (A) Example of jumping PCR occurring during the screening for a KI allele. A PCR product can be generated by “jumping” between two single-strand DNA products synthesized from the transgenic repair template and the genome respectively. This will generate a false-positive PCR signal even in the absence of a *MosTIC*-engineered allele (adapted from Robert *et al.*, 2009). For primers names, we use the same nomenclature as in Fig. 3. (B) Effect of annealing temperature on jumping PCR. In this experiment, mixed transgenic/nontransgenic worm populations derived from four non-heat-shocked independent lines (a, b, c, d) constructed to engineer by *MosTIC* a KI allele at the *unc-5* locus were analyzed by PCR using the primers designed for the *MosTIC* allele PCR screening. At annealing temperatures (indicated at the top of the gels) ranging from 60 to 66.4°C, false-positive PCR products are generated by jumping PCR. They disappear at 68.4 and 70°C. Two dilutions of a control plasmid (P1=1 ng/ $\mu$ L and P2=10<sup>-2</sup> ng/ $\mu$ L) are used to verify that the amplification is still efficient in the presence of a specific PCR template. (C) Example of a PCR screen performed on animals derived from the lines tested in (B). Mixed transgenic/nontransgenic populations derived from heat-shocked transgenic animals are tested by PCR. When primer annealing is performed at 64 and 66.4°C, many false-positive signals are detected. When annealing is performed at 68.4°C, only one signal is present (pool #3). Further sibling selection experiments performed on pool 3 identified worms carrying the desired *MosTIC*-engineered *unc-5* locus.

sequence and in order to minimize it, PCR conditions need to be set up for each *MosTIC* experiment.

(b) *PCR Conditions Setup*

We observed that jumping PCR could be minimized by reducing the annealing time, increasing the annealing temperature, and/or diluting 10–100 times the worm lysates before starting PCR.

To optimize the PCR conditions and minimize “jumping PCR,” proceed as follow:

1. Construct a positive control for your PCR (Fig. 3B and 3C). It must contain the full-length repair template and an additional 200–300 nucleotides long genomic fragment including the primers used for *MosTIC*-engineered allele screening. Using the Phusion high-fidelity DNA polymerase (Finnzymes) or any other high-fidelity polymerase, amplify (i) the full-length repair template on the previously constructed plasmid using P1 and P4 for KO/del repair template and P1' and P6' for KI repair template and (ii) the 200–300 nucleotide long genomic fragment adjacent to the short arm of the repair template. This latter is amplified with primers P9–P10 and P11' –P12' respectively. P9 and P11' contain in their 5' end a tail of 24 nucleotides that overlaps with the end of the short arm of the repair template and fusion PCR is performed using P1–P10 and P1' –P12', respectively. The PCR product is subcloned using the Zero Blunt PCR cloning kit (Invitrogen) or any other cloning kit.
2. For each transgenic line, pick three to five non-heat-shocked transgenic animals on a 6 cm fresh NGM plate. Grow these worms and their progeny for 1 week at 20°C. Collect the mixed nontransgenic/transgenic worm population with 1 mL of M9 1× buffer and transfer them to a 1.5 mL tube. Place on ice for 10 min for sedimentation. Collect 50 µL at the bottom of the tube and transfer to 0.2 mL PCR tubes. Add 50–100 µL of worm lysis buffer complemented with proteinase K. Perform lysis at 65°C for 2–3 h and inactivate proteinase K by incubating the lysate at 95°C for 20 min. Such lysates can be stored at –80 °C for at least 3 weeks.
3. To set up PCR conditions, use 1 µL of each lysate, a 1 ng/µL and a 10<sup>-2</sup> ng/µL dilution of the control plasmid preparation. For each sample, perform the following PCR program (using a standard *Taq* DNA polymerase):
  - step 1: 3 min at 95°C,
  - step 2: 30 s at 95°C,
  - step 3: 15 s at annealing temperature,
  - step 4: 2 min at 72°C temperature,
  - step 5: cycle 29 times from steps 2 to 4
  - step 6: 5 min at 72°C.

Make a gradient of annealing temperature ranging from 55 to 70°C. For the nested PCR, dilute the first PCR 100 times and perform the second PCR at the same annealing temperature.

An example of a setup experiment is shown Fig. 4B. It demonstrates that jumping PCR is detected when primer annealing is performed between 60°C and 66.4°C. There is no PCR jumping when primer annealing is performed at 68.4 °C. However, the control plasmid can still be amplified using this latter annealing temperature suggesting that a *Mos*TIC allele with a similar structure should be detected under these conditions. A similar experiment was performed on pools of heat-shocked transgenic animals derived from the lines tested on Fig. 4B (Fig. 4C), five pools out of six were positive when annealing the primers at 64°C. In contrast, only pool 3 was still positive when annealing the primers at 68.4°C. Further screening confirmed that pool 3 indeed contained a *Mos*TIC-engineered allele.

(c) *PCR Screening Using a Sibling-Selection Strategy*

1. Wait for the pools of heat-shocked animals (cf. Part III. B) to exhaust their food supply.
2. Wash half of the plate with 1 mL of M9 1 × buffer and transfer the animals to a 1.5 mL tube. Let the worms sediment on ice, collect 50 μL at the bottom of the tube, and transfer the worms to PCR plates.
3. Perform lysis and PCR as previously described (cf. Part III. C. 1. b 2 & 3) using the optimized PCR parameters.
4. Once a positive PCR signal is identified, transfer a chunk of the corresponding plate to a fresh plate. From the developing population, make 15 pools of 20 nontransgenic (non-GFP-positive) animals. Selecting nontransgenic animals at this step (i) prevents potential jumping PCR problems and (ii) does not affect the recovery frequency of *Mos*TIC events since they happened in the germ line of the heat-shocked animals few generations before. Wait for the pools to exhaust their food supply and analyze them as described above. At this step, *Mos*TIC-engineered alleles can be, most of the time, detected by a single PCR round. For no clear reason, we sometimes observe a significant decrease of the efficiency of the nested PCR. It might be due to the absence of single-strand DNA generated from the transgene and that is used as PCR template in the previous screening step.
5. From one positive subpool, clone 40–80 individuals to single plates to identify the *Mos*TIC-engineered strain.

## 2. Phenotypic Screening

When phenotypic screening strategies could be used, they turned out to be very efficient in some of our experiments (Robert and Bessereau, 2007 and unpublished data). They can be used, for instance, to generate a KO/del allele starting from a nonmutagenic *Mos*I insertion or to generate a functional KI allele starting from a mutagenic *Mos*I insertion. In both cases, it is recommended to work with nonrescuing repair templates. The presence of a repair template able to rescue by itself the

function of the target gene will make more difficult the identification of *Mos*TIC-engineered animals either because it will mask the presence of *Mos*TIC-engineered animals or because it will generate many false-positives.

It should also be kept in mind that DSB repair can sometimes regenerate functional alleles after excision of mutagenic *Mos1* insertions with the same efficiency as transgene-instructed gene conversion (Robert and Bessereau, 2007; Robert *et al.*, 2008 and unpublished data) and is also able to generate mutant alleles after excision of nonmutagenic *Mos1* insertions. Hence, it will always be necessary to analyze the molecular structure of the revertant or mutant alleles selected on individual phenotypes to identify *bona fide* *Mos*TIC-engineered strains.

### E. An Alternative Protocol Based on the Constitutive Expression of the *Mos* Transposase

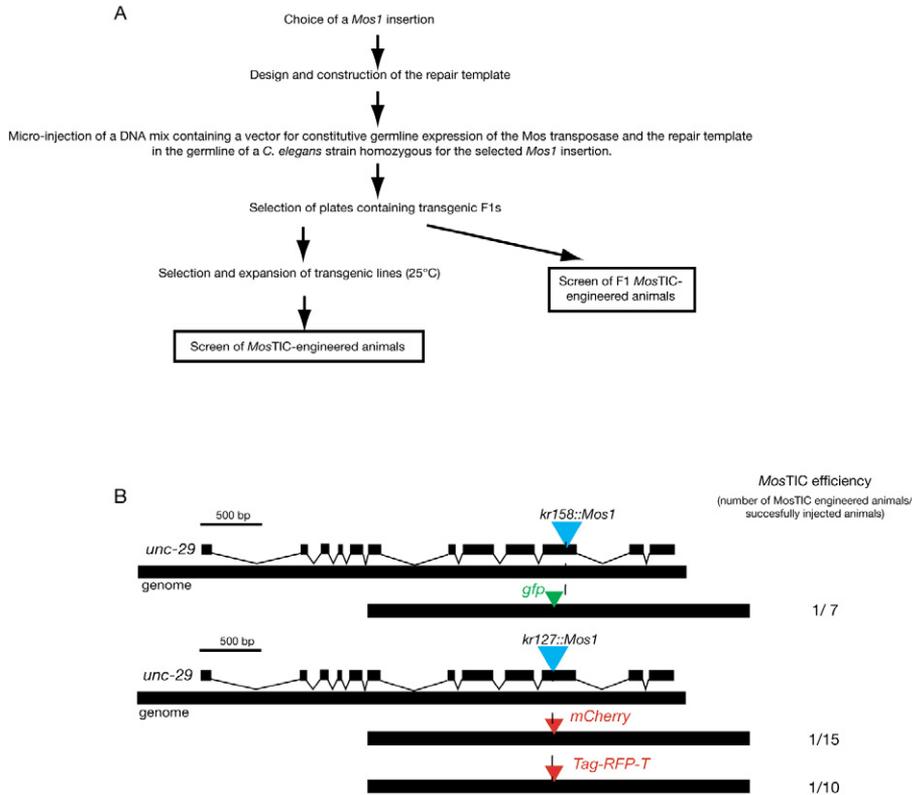
An alternative *Mos*TIC protocol is being developed in our laboratory (Fig. 5). It is based on the observation that *Mos1* insertions can apparently be remobilized directly in the germ line of animals that have been injected with a vector containing the *Mos* transposase under the control of the constitutive germ line promoter *Pglh-2* (Frokjaer-Jensen *et al.*, 2008).

In this protocol, a DNA mix containing the repair template (50 ng/ $\mu$ L), the *pJL43.1* plasmid (*Pglh-2::MosTase* at 50 ng/ $\mu$ L), and the *pPD118.33* (*Pmyo-2::GFP* at 5 ng/ $\mu$ L) is injected into the germ line syncitium of worms homozygous for the insertion to be remobilized. Each injected P0 animal is kept at 20°C on a NGM plate and screened for the presence of transgenic worms expressing GFP in the pharynx in their F1 progeny to verify that they were successfully injected. When using a phenotypic screening strategy, candidate *Mos*TIC alleles can be screened in the progeny of P0 animals that segregate transgenic animals. When using PCR screening, consider each P0 plate containing F1 progeny as a pool and use the sib-ling-selection strategy described previously.

Interestingly, we observed that *Mos*TIC alleles could arise at later generations using *Pglh-2::MosTase*. Therefore, it is worth isolating transgenic lines and screening for *Mos*TIC alleles at subsequent generations if no *Mos*TIC-engineered allele is recovered on the P0 plates. To minimize germ line silencing of the extrachromosomal array containing the *Mos* transposase source, we recommend to maintain the transgenic lines at 25°C.

Using this protocol, we constructed three independent KI alleles of *unc-29*, a gene coding for a subunit of the acetylcholine receptor expressed at neuromuscular junctions. Based on phenotypic screening, we were able to recover *Mos*TIC-engineered alleles with a frequency ranging from one *Mos*TIC-engineered allele out of 7 successfully injected animals (14%) to one *Mos*TIC-engineered allele out of 15 successfully injected animals (6.6%) (Fig. 5B).

When using a phenotypic screening strategy, engineered alleles can be obtained in less than 10 days after injection. When using a PCR-based strategy, engineered alleles are recovered in 3 weeks after injection.



**Fig. 5** Description of an alternative *MosTIC* procedure. (A) Overview of the procedure. In this strategy, the *Mos* transposase is under the control of a constitutive germ line promoter and *MosTIC*-engineered alleles can be screened directly in the F1 progeny of injected animals. (B) *MosTIC* efficiency using this strategy. The number of recovered *MosTIC*-engineered alleles is given compared to the number of successfully injected animals. (See color plate.)

## F. An Alternative Protocol for *MosTIC*-Engineered KO/del Alleles Selection

An alternative *MosTIC* protocol called *MosDEL* (which stands for *Mos1*-mediated deletion) was recently developed by the group of E. Jorgensen (University of Utah) to engineer and select KO/del alleles (Frokjaer-Jensen *et al.*, 2010). In the *MosDEL* protocol, a *C. briggsae unc-119(+)* fragment, which can rescue the mutant phenotype of *unc-119(ed3)* mutants, is inserted at the place of the deleted region. First, the *unc-119(ed3)* allele is crossed into a strain containing a *Mos1* insertion in the region to delete. Second, the repair template is injected in that strain together with a vector containing the *Mos* transposase under the control of the constitutive germ line promoter *Pglh-2* and 3 plasmids expressing mCherry in different tissues to mark extrachromosomal arrays. After two to three generations

(7 days), progeny of the injected animals are screened for *Mos*TIC-engineered KO/del alleles by looking at the worms that are rescued for *unc-119* but that do not express any of the injection markers. Most of them contain a *Mos*TIC-engineered allele in which the region to delete has been replaced by the *unc-119* marker.

Using this protocol, it has been possible to (i) generate several deletions at the *dpy-13* locus, some of them being 25-kb long, (ii) delete the tandem gene duplication containing the *cst-1* and *cst-2* genes, and (iii) delete the essential gene *dyn-1*. Interestingly, the presence of the *unc-119(+)* marker at the site of the deletion makes a perfectly balanced lethal chromosome. A detailed protocol is provided in Frokjaer-Jensen *et al.* (2010).

This method seems to be extremely simple and powerful to select *Mos*TIC-engineered KO/del alleles without using PCR screening. It remains, however, to evaluate the potential effect of the *unc-119(+)* fragment on the expression of nearby genes.

## G. Frequently Asked Questions

1. When is it Better to use the Alternative *Mos*TIC Protocol Based on Constitutive Germ Line Expression of the *Mos* Transposase Instead of the Standard Protocol?

The standard *Mos*TIC protocol, requires the microinjection of only few animals containing the *Mos1* insertion of interest in order to generate transgenic lines. The alternative protocol requires more time microinjecting but permits the identification of engineered alleles a few days after injection. However, the overall *Mos*TIC efficiency depends on the quality of the injection, which varies with the researcher injection ability and the genetic background. If the *Mos*TIC events are predicted to be rare (engineering site far from the *Mos1* insertion), we recommend to use a heat-shock-based strategy.

2. What is the Optimal Length for the Homologous Arms Present on the Repair Template?

Standard repair templates contain two 1.5 kb regions of homology (see Part III.A.1 for details). We tried to increase the length of one of this arm but did not observe a significant increase in *Mos*TIC efficiency (Robert and Bessereau, 2007 and Fig. 2B). On the contrary, shortening one of the arm length down to 700 bp dramatically decreased *Mos*TIC efficiency.

When designing a repair template, there are two points to keep in mind depending on the screening methodology. First, if a phenotypic screening strategy is used, it is essential that the repair template by itself does not rescue the mutant phenotype. Second, if screening by PCR, it is essential to keep at least one of the homologous arm short enough to be able to design primers outside of the repair template. In our experiments, we usually keep the short homologous arm at 1.5 kb and we design primers in such a way that PCR fragment will range between 1.8 and 2 kb.

3. Did you Try to Have the Repair Template and the Mos Transposase on Two Different Arrays? Why Don't you use the Existing hsp::Tpase Array *oxEx166* (Bessereau *et al.*, 2001; Boulin and Bessereau, 2007; Williams *et al.*, 2005)?

Initial *MosTIC* experiments were performed using the existing *oxEx166* array, which contains heat-shock inducible source of Mos transposase, combined with an independent array carrying the repair template. In this configuration, *MosTIC* was not more efficient than using a single array carrying both the transposase and the repair template. Therefore, we recommend to generate a single extrachromosomal array because it is sometimes problematic to select and amplify double transgenic animals. In addition, we think that using a freshly injected source of transposase for each experiment minimizes the risk of working with transgenes that might have undergone silencing.

4. Did you Try Other Coinjection Markers than *pPD118.33 (Pmyo-2::GFP)*?

Any coinjection marker might work as long as it is easy to detect. For example, we used the dominant injection marker *rol-6(su1006)* contained on plasmid *pRF4* (Mello *et al.*, 1991). However, extrachromosomal arrays are not always stable in mitosis or meiosis. As a consequence, transgenic animals are mosaic, meaning that, in a single animal, some cells carry the transgene whereas others have lost it (Stinchcomb *et al.*, 1985). Mosaicism can sometimes complicate the discrimination between transgenic animals from nontransgenic animals, which is an issue when screening by PCR. In our experiments, confusion between nontransgenic and transgenic animals happened when working with *pRF4*. On the contrary, we did not have significant problems discriminating between nontransgenic and transgenic animals when using *pPD118.33* as a transformation marker, which we therefore recommend.

5. Is *MosTIC* More Efficient in Genetic Backgrounds in the Absence of Germ Line Transgene Silencing?

No. In the *C. elegans* germ line, extrachromosomal arrays are submitted to silencing (see "Introduction"). It might result in decreased transposase expression and *MosI* excision frequency. To test this hypothesis, we performed *MosTIC* experiments in the *mut-7(pk204)* background where germ line transgene silencing is released (Ketting *et al.*, 1999; Robert and Bessereau, 2007). In this mutant background, *MosTIC* events were recovered at the same frequencies as in the wild-type background suggesting that in spite of transgene silencing, the amount of Mos transposase present in the germ line was not a restrictive factor for *MosTIC* efficiency.

6. Is *MosTIC* More Efficient in the Absence of End-Joining DSB Repair?

Several DSB repair mechanisms, including end-joining, are at work in the *C. elegans* germ line, raising the possibility that they might compete with transgene-instructed gene conversion and decrease its efficiency. To test this hypothesis,

we performed *Mos*TIC experiments in mutant backgrounds for *lig-4* and *cku-80*, two genes that are highly conserved and encode factors that are required for nonhomologous end-joining (Robert and Bessereau, 2007). We demonstrated that in the soma of the tested mutants, nonhomologous end-joining was defective to repair a *MosI*-excision induced DSB. However, in the germ line, we were not able to detect any quantitative or qualitative differences between the repair events generated in wild-type or *lig-4* and *cku-80* mutant backgrounds.

#### 7. Does Spontaneous Excision of *Mos1* Occur?

Yes, rarely. We indeed had noticed that some *MosI* insertions are not stable in the germ line and can be lost. For this reason, we recommend to check by PCR that the *MosI* insertion of interest is still present before starting the injection procedure. Similarly, this can also be done when thawing a *MosI* insertion containing strain, when a *MosI* insertion containing strain has been maintained for a long time or when it is difficult to identify a *Mos*TIC-engineered allele.

---

---

---

## IV. Materials

### A. Equipment

Standard *C. elegans* culture facility  
Standard germ line microinjection setup (Evans, 2006; Mello and Fire, 1995)  
Dissecting scope equipped for fluorescence detection  
Gradient thermal Cycler  
Standard setup for agarose gel electrophoresis  
Multichannel pipettes  
Repeating dispenser

### B. Solutions and Reagents

M9 buffer: 22 mM KH<sub>2</sub>PO<sub>4</sub>, 22 mM Na<sub>2</sub>HPO<sub>4</sub>, 85 mM NaCl, 1 mM MgSO<sub>4</sub>  
Worm lysis buffer: 50 mM KCl, 10 mM Tris pH 8.2, 25 mM MgCl<sub>2</sub>, 0.45% NP-40, 0.45% Tween-20, 0.01% Gelatin complemented with Proteinase K (final concentration: 1 mg/mL; Eurobio GEXPRK00-6R) before performing lysis.  
Phusion high-fidelity DNA Polymerase (Finnzymes, Cat#F-530)  
*Taq* DNA Polymerase (New England Biolabs, Cat#M0273)  
Qiaquick gel extraction kit (Qiagen, Cat#28704)  
Zero Blunt PCR Cloning Kit (Invitrogen, Cat#K2700)  
QuickChange II Site-directed Mutagenesis kit (Stratagene, Cat#200523)

### C. Plasmids

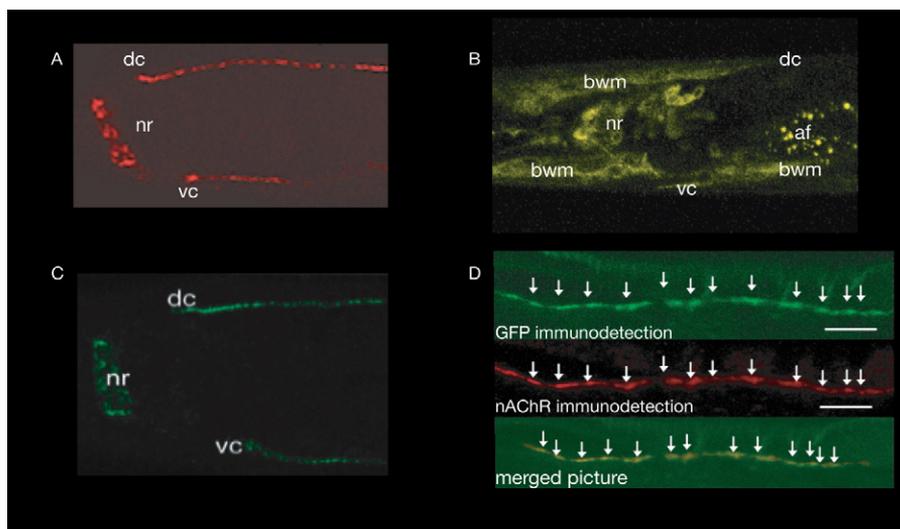
The plasmids *pJL44* and *pJL43* that contain the *Mos* transposase (Bessereau *et al.*, 2001) and *pPD118.33* (*Pmyo-2::GFP*) are available at addgene (<http://www.addgene.org>).

## V. Discussion: *Mos*TIC Applications and Derivatives

### A. Gene Function Analysis Using *Mos*TIC-Engineered Alleles

In our laboratory, we routinely use *Mos*TIC to tag genes of interest and our first study using *Mos*TIC-engineered alleles was published recently (Gendrel *et al.*, 2009). We characterized a protein complex involved in the clustering of acetylcholine receptors at the *C. elegans* neuromuscular junction. In this study, KI alleles were used for genetic, cell biology, and biochemistry experiments. It bypassed the expensive and time-consuming need of developing and testing antibodies specifically recognizing each protein of interest.

In the course of our experiments, we often observe that the use of *Mos*TIC-engineered alleles significantly minimizes experimental artifacts due to overexpression of engineered proteins expressed from repetitive transgenes. For instance, we expressed a functional YFP-tagged version of UNC-63, one of the subunits of an acetylcholine receptor present at the neuromuscular junction, from repetitive transgenes and by *Mos*TIC (Gendrel *et al.*, 2009). As illustrated in Fig. 6, endogenous



**Fig. 6** Expression of tagged proteins in *C. elegans*: repetitive transgenes versus engineered endogenous loci. (A) The levamisole-sensitive acetylcholine receptor is detected by immunostaining of its subunit UNC-29. It is localized in the muscle membrane at the nerve ring (nr), the dorsal nerve cord (dc) and the ventral nerve cord (vc). (B-E) Localization of an YFP-tagged version of UNC-63, another subunit of the levamisole-sensitive acetylcholine receptor. (B) The tagged UNC-63 subunit is expressed from a repetitive transgene. In the presence of this transgene, a loss-of-function allele of *unc-63* is rescued indicating that functional receptors are made and that at least some of them are located at the muscle membrane. However, most of the tagged subunit is diffusely localized in body wall muscles (bwm) (af: gut autofluorescence). (C) Confocal imaging of a strain homozygous for a *Mos*TIC-engineered allele expressing the same UNC-63-tagged version as in (B). Fluorescence is mostly detected in the nerve ring and ventral and dorsal nerve cords. (D) Immunostaining characterization of the strain observed in (C). An antibody against the YFP detects puncta in the ventral nerve cord that co-localize with the other subunits of the levamisole-sensitive acetylcholine receptor. (Images are courtesy of C. Gally and G. Rapti). (See color plate.)

UNC-63 is detected by immunostaining at synapses in the nerve ring, dorsal nerve cord, and ventral nerve cord. When expressed from an extrachromosomal array under the control of its endogenous promoter, UNC-63-YFP rescues the mutant phenotypes of an *unc-63* null but the fusion protein is diffusely distributed in intracellular muscle compartments and in the plasma membrane. On the contrary, when the *unc-63* locus is engineered by *MosTIC*, UNC-63-YFP is only present at synapses and colocalizes with the other subunits of the acetylcholine receptor.

For germ line-expressed genes, experimental evidence are still required but *MosTIC*-engineered alleles should be extremely valuable tools since it has been so far extremely difficult to construct healthy strains expressing tagged versions of genes expressed in the germ line.

### B. Single-Copy Transgene Genomic Integration by MosSCI

Quickly after its establishment, *MosTIC* protocol was adapted to derive an efficient technique of single-copy transgene insertion known as MosSCI (for *MosI*-mediated single-copy insertion) ((Frokjaer-Jensen *et al.*, 2008) and <http://sites.google.com/site/jorgensenmossci/Home>). Using MosSCI, a transgene can be inserted as a single-copy at an intergenic position defined by a pre-existing *MosI* insertion. In this case, the repair template is made of the sequence of interest flanked by two DNA arms, which are homologous to the genomic regions flanking the *MosI* insertion point. After induction of *MosI* excision in the presence of the Mos transposase, positive and negative-counter selections are used to select DSB repair events resulting in the single-copy integration of the extrachromosomal repair template. Using MosSCI, single-copy transgene integrations could be recovered at least at two “neutral” intergenic positions defined by *MosI* insertions that were chosen because they apparently do not induce phenotypic changes. MosSCI alleles exhibit a stable and robust expression both in the germ line and the soma (Frokjaer-Jensen *et al.*, 2008) and first examples of studies based on the analysis of MosSCI alleles have been recently published (Lehrbach *et al.*, 2009; Pagano *et al.*, 2009).

---

---

---

## VI. Summary

The *MosTIC* technique described in this chapter provides a way to engineer customized alleles in *C. elegans*. It requires an insertion of the *MosI* transposon in the target locus. Excision of the transposon triggered by germ line expression of the Mos transposase creates a DNA DSB that locally stimulates homologous recombination. Sequence variations contained in a repair template provided in a transgene can be copied into the broken locus by gene conversion. In the standard *MosTIC* protocol, *MosI* transposition is induced by heat-shocking transgenic lines carrying a Mos transposase expression vector under the control of a heat-shock inducible promoter and an engineered repair template. *MosTIC* alleles can be identified in the progeny of the heat-shocked animals by PCR or based on the appearance of a new

phenotype caused by the *MosTIC* allele. This protocol was successfully used to engineer point mutations, KO/del and KI alleles.

The use of *MosTIC*-engineered alleles is increasing and it has been demonstrated that (i) they are compatible with genetics, cell biology, and biochemistry approaches, (ii) their use minimizes experimental artifacts due to misexpression of transgenic sequences in *C. elegans*, and (iii) gene tagging by homologous recombination can bypass the time consuming and expensive task of developing specific antibodies against each protein of interest.

## Acknowledgment

The authors thank Luis Briseno-Roa and Bérangère Pinan-Lucarré for critical reading of this manuscript.

## References

- Barrett, P. L., Fleming, J. T., and Gobel, V. (2004). Targeted gene alteration in *Caenorhabditis elegans* by gene conversion. *Nat. Genet.* **36**, 1231–1237.
- Bazopoulou, D., and Tavernarakis, N. (2009). The NemaGENETAG initiative: Large scale transposon insertion gene-tagging in *Caenorhabditis elegans*. *Genetica* **137**, 39–46.
- Berezikov, E., Bargmann, C. I., and Plasterk, R. H. (2004). Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40.
- Berset, T. A., Hoier, E. F., and Hajnal, A. (2005). The *C. elegans* homolog of the mammalian tumor suppressor Dep-1/Sccl inhibits EGFR signaling to regulate binary cell fate decisions. *Genes Dev.* **19**, 1328–1340.
- Bessereau, J. L. (2006). Insertional mutagenesis in *C. elegans* using the *Drosophila* transposon Mos1: A method for the rapid identification of mutated genes. *Methods Mol. Biol.* **351**, 59–73.
- Bessereau, J. L., Wright, A., Williams, D. C., Schuske, K., Davis, M. W., and Jorgensen, E. M. (2001). Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74.
- Boulin, T., and Bessereau, J. L. (2007). Mos1-mediated insertional mutagenesis in *Caenorhabditis elegans*. *Nat. Protoc.* **2**, 1276–1287.
- Broverman, S., MacMorris, M., and Blumenthal, T. (1993). Alteration of *Caenorhabditis elegans* gene expression by targeted transformation. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4359–4363.
- Cheeseman, I. M., Niessen, S., Anderson, S., Hyndman, F., Yates 3rd, J. R., Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18**, 2255–2268.
- Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the *C. elegans* germ line. *Genes Dev.* **14**, 1578–1583.
- Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S., and Smithies, O. (1987). Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**, 576–578.
- Duverger, Y., Belougne, J., Scaglione, S., Brandli, D., Beclin, C., and Ewbank, J. J. (2007). A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res.* **35**, e11.
- Evans, T. C., ed. Transformation and microinjection (April 6, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.108.1, <http://www.wormbook.org>.
- Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383.

- Frokjaer-Jensen, C., Davis, M. W., Holloper, G., Taylor, J., Harris, T. W., Nix, P., Lofgren, R., Prestgard-Duke, M., Bastiani, M., Moerman, D. G., and Jorgensen, E. M. (2010). Targeted gene deletions in *C. elegans* using transposon excision. *Nature Methods* **7**, 451–453.
- Gendrel, M., Rapti, G., Richmond, J. E., and Bessereau, J. L. (2009). A secreted complement-control-related protein ensures acetylcholine receptor clustering. *Nature* **461**, 992–996.
- Gloor, G. B., Nassif, N. A., Johnson-Schlitz, D. M., Preston, C. R., and Engels, W. R. (1991). Targeted gene replacement in *Drosophila* via P element-induced gap repair. *Science* **253**, 1110–1117.
- Green, R. A., Audhya, A., Pozniakovsky, A., Dammermann, A., Pemble, H., Monen, J., Portier, N., Hyman, A., Desai, A., and Oegema, K. (2008). Expression and imaging of fluorescent proteins in the *C. elegans* gonad and early embryo. *Methods Cell Biol.* **85**, 179–218.
- Hartl, D. (2001). Discovery of the transposable element mariner. *Genetics* **157**, 471–476.
- Hobert, O. (2002). PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques* **32**, 728–730.
- Jacobson, J. W., and Hartl, D. L. (1985). Coupled instability of two X-linked genes in *Drosophila mauritiana*: Germinal and somatic mutability. *Genetics* **111**, 57–65.
- Jacobson, J. W., Medhora, M. M., and Hartl, D. L. (1986). Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8684–8688.
- Jantsch, V., Pasierbek, P., Mueller, M. M., Schweizer, D., Jantsch, M., and Loidl, J. (2004). Targeted gene knockout reveals a role in meiotic recombination for ZHP-3, a Zip3-related protein in *Caenorhabditis elegans*. *Mol. Cell Biol.* **24**, 7998–8006.
- Kelly, W. G., Xu, S., Montgomery, M. K., and Fire, A. (1997). Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238.
- Ketting, R. F., Haverkamp, T. H., van Luenen, H. G., and Plasterk, R. H. (1999). Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**, 133–141.
- Ketting, R. F., and Plasterk, R. H. (2000). A genetic link between co-suppression and RNA interference in *C. elegans*. *Nature* **404**, 296–298.
- Lehrbach, N. J., Armisen, J., Lightfoot, H. L., Murfitt, K. J., Bugaut, A., Balasubramanian, S., and Miska, E. A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **16**, 1016–1020.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the *C. elegans* germline. *Curr. Biol.* **18**, 1476–1482.
- Paabo, S., Irwin, D. M., and Wilson, A. C. (1990). DNA damage promotes jumping between templates during enzymatic amplification. *J. Biol. Chem.* **265**, 4718–4721.
- Pagano, J. M., Farley, B. M., Essien, K. I., and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20252–20257.
- Plasterk, R. H. (1991). The origin of footprints of the Tc1 transposon of *Caenorhabditis elegans*. *EMBO J.* **10**, 1919–1925.
- Plasterk, R. H., and Groenen, J. T. (1992). Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. *EMBO J.* **11**, 287–290.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Praitis, V., Ciccone, E., and Austin, J. (2005). SMA-1 spectrin has essential roles in epithelial cell sheet morphogenesis in *C. elegans*. *Dev. Biol.* **283**, 157–170.
- Robert, V., and Bessereau, J. L. (2007). Targeted engineering of the *Caenorhabditis elegans* genome following Mos1-triggered chromosomal breaks. *EMBO J.* **26**, 170–183.
- Robert, V. J., Davis, M. W., Jorgensen, E. M., and Bessereau, J. L. (2008). Gene conversion and end-joining-repair double-strand breaks in the *Caenorhabditis elegans* germline. *Genetics* **180**, 673–679.

- Robert, V. J., Katic, I., and Bessereau, J. L. (2009). Mos1 transposition as a tool to engineer the *Caenorhabditis elegans* genome by homologous recombination. *Methods* **49**, 263–269.
- Robert, V. J., Sijen, T., van Wolfswinkel, J., and Plasterk, R. H. (2005). Chromatin and RNAi factors protect the *C. elegans* germline against repetitive sequences. *Genes Dev.* **19**, 782–787.
- Scherer, S., and Davis, R. W. (1979). Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4951–4955.
- Sijen, T., and Plasterk, R. H. (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **426**, 310–314.
- Stiernagle, T. Maintenance of *C. elegans* (February 11, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.101.1, <http://www.wormbook.org>.
- Stinchcomb, D. T., Shaw, J. E., Carr, S. H., and Hirsh, D. (1985). Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol. Cell Biol.* **5**, 3484–3496.
- Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512.
- Vazquez-Manrique, R. P., Legg, J. C., Olofsson, B., Ly, S., and Baylis, H. A. (2010). Improved gene targeting in *C. elegans* using counter-selection and FLP-mediated marker excision. *Genomics* **95**, 37–46.
- Williams, D. C., Boulin, T., Ruaud, A. F., Jorgensen, E. M., and Bessereau, J. L. (2005). Characterization of Mos1-mediated mutagenesis in *Caenorhabditis elegans*: A method for the rapid identification of mutated genes. *Genetics* **169**, 1779–1785.
- Wilm, T., Demel, P., Koop, H. U., Schnabel, H., and Schnabel, R. (1999). Ballistic transformation of *Caenorhabditis elegans*. *Gene* **229**, 31–35.

---

---

## CHAPTER 4

# RNAi Methods and Screening: RNAi Based High-Throughput Genetic Interaction Screening

**Patricia G. Cipriani**<sup>\*,†</sup> and **Fabio Piano**<sup>\*,†,‡</sup>

<sup>\*</sup>Center for Genomics and Systems Biology

<sup>†</sup>Department of Biology, New York University, New York, NY 10003

<sup>‡</sup>New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

---

### Abstract

- I. Introduction: Large-Scale RNAi Screening in *C. elegans*
  - II. Methods
    - A. Large-Scale RNAi Screening
    - B. Scoring of Interactions/Phenotypes
    - C. Quantification of Embryonic Lethality or Survival on Agar Plates
  - III. Materials
    - A. Reagents
    - B. Buffers and Media
    - C. Supplies and Equipment
- Acknowledgments  
References

---

---

### Abstract

Expanding on decades of mutational analyses, numerous genome-scale RNAi screens have now been performed in *C. elegans*, leading to estimates that the majority of genes with essential functions that can be revealed by single-gene perturbations have already been identified in this organism. To build on this basic foundation and uncover condition-dependent or combinatorial effects of non-essential genes will require even higher-scale screening. Here we describe a method for performing high-throughput RNAi-based screens in *C. elegans* in liquid in 96-well plates, and we explain how to systematically test for enhancement and

suppression of temperature-sensitive mutations. This chapter covers our entire set of protocols, from setting up the experiment and screening schedule, to scoring the results. The rapid acquisition of high-quality images of each experiment allows the management of a large number of samples per screening cycle and opens up new possibilities for quantitative scoring, computerized image analysis, and the ability to review results independent of the time constraints that are associated with large-scale screening.

---

---

---

## I. Introduction: Large-Scale RNAi Screening in *C. elegans*

A powerful way to help decipher a gene's function is to disturb it and analyze the effect that is produced. The result provides a clue about the processes in which the gene product is required, or more generally, the response of the system to the perturbation. Extending this idea to genome-wide analyses has given new insights into the molecular mechanisms underlying basic cell biological processes (Mohr *et al.*, 2010; Perrimon and Mathey-Prevot, 2007). *C. elegans* in particular has been used successfully as a model animal for large-scale genetic, RNAi, and other types of screening – including chemical genetic screens to link genes, proteins, or small molecules with biological roles (e.g., Burns *et al.*, 2006; Kwok *et al.*, 2006; Min *et al.*, 2007). The reasons that *C. elegans* is such a good *in vivo* model system both are biological and technical. *C. elegans* displays many developmental programs and behaviors that are conserved across metazoans, yet is unique in that its development is completely mapped out to single-cell resolution, so that its entire lineage of cell fates is defined (Sulston and Horvitz, 1977). On a practical, experimental level, a number of functional genomic tools are available (i.e., described in Wormbook: <http://www.Wormbook.org>). *C. elegans* can grow well within wells of a 96-well plate (van Haften *et al.*, 2004), and it is small enough to be handled through microfluidic devices like liquid-handling robots or Fluorescence Activated Cell Sorting machines without damage (e.g., Ben-Yakar and Bourgeois, 2009; Ben-Yakar *et al.*, 2009; Doitsidou *et al.*, 2008; Fernandez *et al.*, 2010; Stoeckius *et al.*, 2009). Combining these features makes *C. elegans* a premiere animal model for high-throughput *in vivo* biology.

In large-scale screening, we are now moving toward ever higher throughput. A systematic approach that explores condition-specific genetic analysis is potentially infinite. The simple case of reducing the function of two genes simultaneously under laboratory conditions explores an experimental space of over 200,000,000 possible interactions among the ~20,000 genes in the *C. elegans* genome. Adding complexities derived from specific alleles, from diverse genetic background, and from the possibility of cell-specific interactions can expand the systematic search space dramatically. A solution to this problem is to perform even-more complex screens that take advantage of strategies that look for multiple genetic effects on phenotypes (Rockman and Kruglyak, 2009). Other approaches include exploring systematically genetic interactions, as has been successfully done in yeast (Boone *et al.*, 2007; Costanzo *et al.*, 2010;

Davierwala *et al.*, 2005; Tong *et al.*, 2004) and has begun in *C. elegans* (Byrne *et al.*, 2007; Holway *et al.*, 2005; Lehner *et al.*, 2006; Tischler *et al.*, 2006).

Multiple approaches will be needed to map all genetic interactions in a complex multi-cellular organism, including ways that rely on dramatically expanded abilities to perform, archive, and analyze ultra-large systematic screens. Here we explore double genetic combinations using a method to test RNAi effects of single genes in different genetic backgrounds.

In *C. elegans*, mutants affecting most cellular processes are available from a public repository, the *Caenorhabditis* Genetics Center (CGC). In addition, there are several ongoing projects to produce mutants targeting almost every gene in the genome, including the *C. elegans* Knockout Consortium (<http://celeganskoconsortium.omrf.org/>), the *C. elegans* National Bioresource Project of Japan (<http://shigen.lab.nig.ac.jp/c.elegans/index.jsp>), and the NemaGENETAG project (<http://elegans.imbb.forth.gr/nemagenetag/>).

Early large-scale RNAi studies concentrated on individual gene analyses (Fernandez *et al.*, 2005; Fraser *et al.*, 2000; Gonczy *et al.*, 2000; Gunsalus and Piano, 2005; Kamath *et al.*, 2003; Maeda and Sugimoto, 2001; Piano *et al.*, 2000; Rual *et al.*, 2004; Sonnichsen *et al.*, 2005). A result from these studies was that less than 20% of genes produced a clear phenotype under laboratory conditions. Although this observation could be due to a number of reasons, including incomplete functional depletion using RNAi or not scoring for all phenotypes, the effect of losing one gene can be masked by genetic redundancy or other compensatory mechanisms. Since genes interact with one another to modulate cellular systems and generate specific phenotypes, it is useful to develop methods that reveal genetic interaction networks in a multi-cellular organism in large scale.

A major technical bottleneck of increasing the throughput of large-scale screening has been the limited window of time in which the results need to be scored. In the following, we describe the pipeline that we have implemented in our lab to overcome this issue. We modified a liquid RNAi protocol in 96-well plates (Ahringer, 2006) to use with temperature-sensitive (ts) mutants, and we use it to perform ~40,000 RNAi experiments per 9-day cycle. In addition, we developed an image-capturing platform that quickly records and archives high-quality images of all the experiments. This has given us the ability to separate the experimental testing from the analysis of the data. We can score data long after it is collected and review it multiple times. Moreover, this approach has opened the door to using computer vision to help develop automation and quantitative data analysis for the phenotypic scoring (White *et al.*, 2010).

We implemented these methods in screens using conditional mutations (e.g., temperature-sensitive), enabling us to control the timing of when we assay for genetic interactions (our unpublished data). Since our studies focus on embryonic development, we maintain the larvae at permissive temperature and do not perturb the gene function until the last larval stage (L4) to bypass genetic interactions that could lead to adult phenotypes before embryos are produced.

---

---

## II. Methods

### A. Large-Scale RNAi Screening

#### 1. Selection of Worm Strains and Gene Targets

Depending on the focus of the research and the phenotype under scrutiny, the screening can be done in the standard laboratory *C. elegans* strain N2, in RNAi hypersensitive strains such as *rrf-3(pk1426)* or *eri-1(mg369)*, *lin-15(n744)*, or in other mutant strains. The protocol that we present in this chapter is tailored for use with ts strains, but could be adapted to use with transgenic reporter strains to evaluate the presence or absence of fluorescence, or to screen for genetic modifiers of the effects of small molecules dissolved in the liquid media.

For the evaluation of genetic interactions that affect embryonic development, we can easily differentiate wild-type (WT), embryonic lethality, and sterility in the images that we produce. Many other phenotypes can be readily observed such as ruptured vulva, protruding vulva, dumpy, patchy coloration, larva lethal, and larval arrest (Fig. 1).

#### Mutants and Temperature-Sensitive Strains

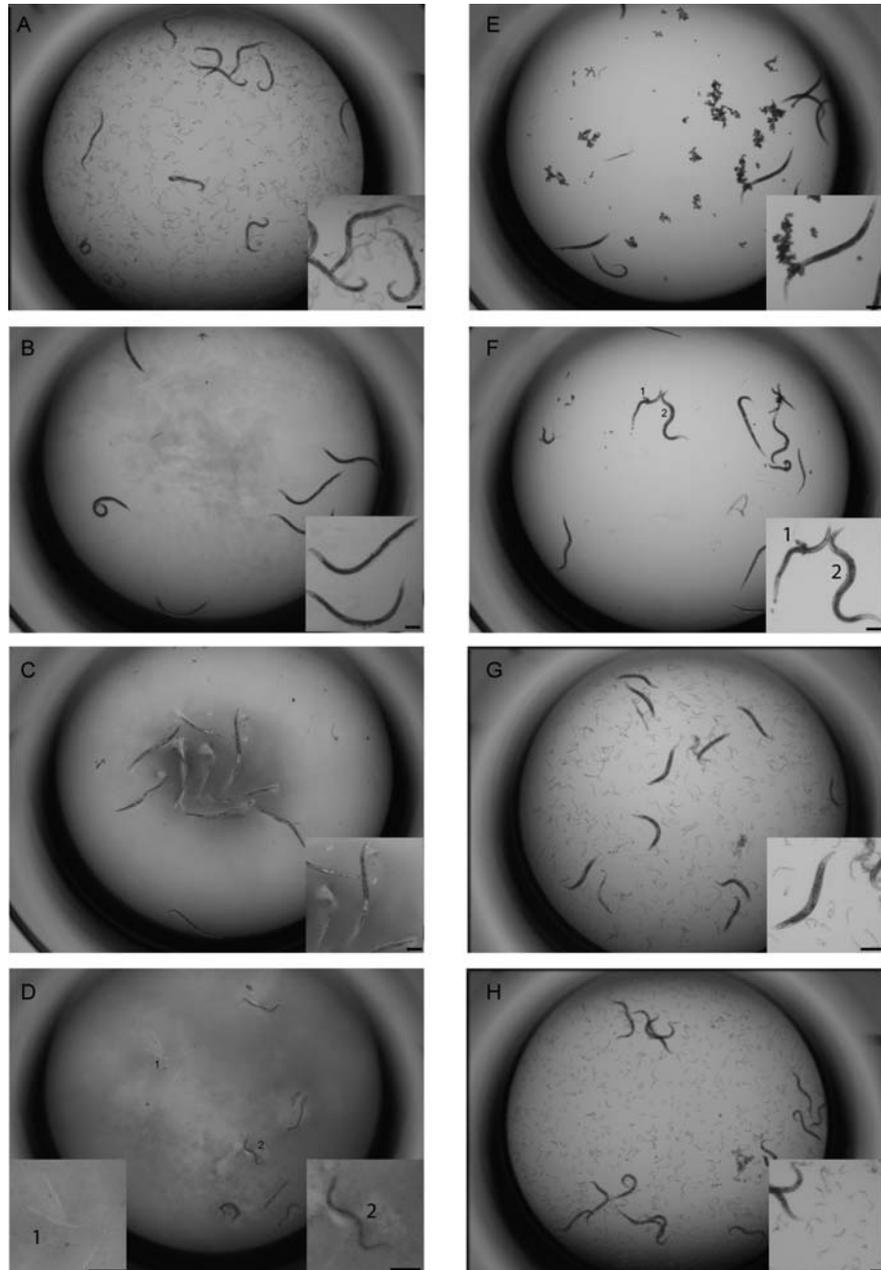
Searching for modifiers of a gene of interest is possible by comparing the results of an RNAi screen in a mutant strain with the results in the control strain N2. Using hypomorphic or conditional alleles, it is possible to evaluate a gene's effect on lethality through the RNAi knock-down of a second gene. Our lab focuses on screening modifiers of genes that play a role in early embryonic development. We use ts strains that allow us to do two types of screens at the same time: one looking for an increase in lethality (enhancement) when the screen is performed at a semi-permissive temperature, and one looking for a reduction in lethality (suppression) when the screen is performed at a semi-restrictive temperature.

#### Temperature Sensitivity Curves

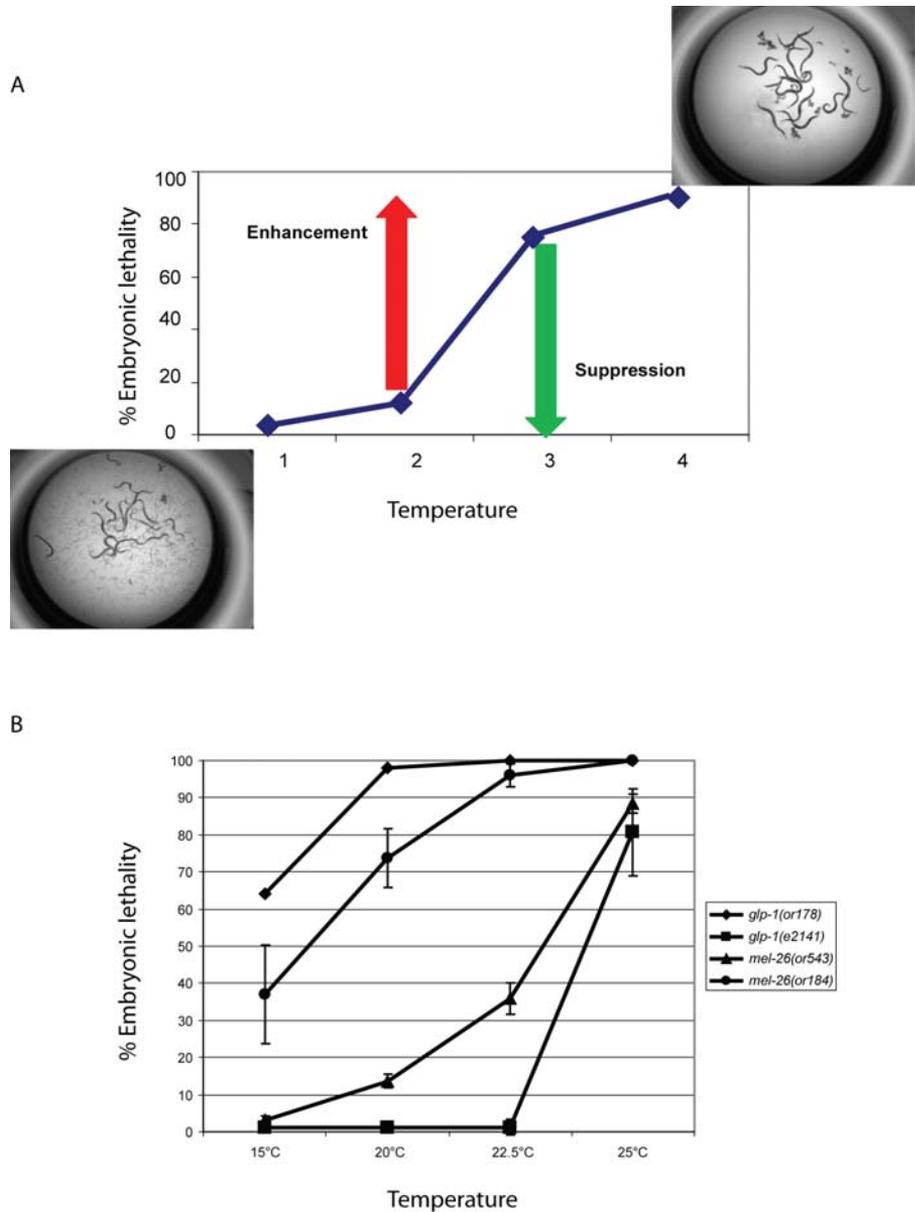
Each ts mutant allele shows specific rates of lethality at varying temperatures. The strain(s) of interest must be tested at a minimum of four different temperatures to determine the best allele to use for the screen (if more than one is available) and to calculate the semi-permissive (less than 20% lethality) and semi-restrictive (more than 80% lethality) temperatures (Fig. 2A). In our lab we begin by testing 15, 20, 22.5, and 25 °C, because even different alleles of the same gene could demonstrate different temperature sensitivities (Fig. 2B).

The procedure for determining temperature sensitivity is as follows:

1. *Pick and plate L4s*. For each mutant strain, collect at least 20 L4s on a medium (60 mm × 15 mm) NGM plate seeded with OP50. From this plate, transfer five worms, 1 worm per well (labeled A1–A5), to each of four 12-well NGM plates seeded with 20 µL OP50 bacteria per well.



**Fig. 1** Phenotypes observed in 96-well liquid RNAi. (A) Wild type. (B) Sterile (C) Patchy coloration and paralyzed. (D) 1. Larva lethal, 2. Larval arrest of P0 (original worms do not develop). (E) Embryonic lethal. (F) 1. Ruptured 2. Protruding vulva. (G) Dumpy (shorter and thicker). (H) Larval arrest of F1 (hatched progeny do not develop). Insets are enlarged sections of the pictures to show details. Scale bars 10  $\mu\text{m}$ .



**Fig. 2** Strain-dependent changes in the percentage of lethality produced in response to incubation temperature. (A) Ideal temperature curve depicting the concept of temperature selection for a modifier screen. (B) Examples of observed differences in the response to temperature of strains harboring mutations in the same gene. (For color version of this figure, the reader is referred to the web version of this book.)

2. *Incubate*. Separate the four plates to the four different temperatures being tested. Allow the worms to develop into adults and lay progeny over a 24-h period.
3. *Transfer worms*. Transfer each of the five worms to a new well in the same 12-well plate, 1 worm per well (labeled B1–B5). We leave the two center wells empty to separate between A and B wells, and because these wells tend to show different conditions from the outer wells.
4. *Incubate*. Place the 12-well plates at their respective temperatures for another 24-h period. Then, remove the worms from the B wells.
5. *Score percent lethality*. After well A or well B sits without adults for a 24-h period, allowing time for any eggs to hatch, count the total brood as well as the number of larvae and eggs for each well. Calculate percent lethality by dividing the number of eggs by the total number of larvae and eggs and multiplying by 100.

### Gene Target Selection

Depending on the project resources and the scale of the screen, the investigator must determine which and how many RNAi clones to use, defining a “target set.” This set could encompass a library of genes selected by different criteria (i.e., GO term category), or could be as large as all clones available to target the genome.

The Ahringer’s RNAi library (Kamath *et al.*, 2003) (<http://www.geneservice.co.uk/products/rnai/Celegans.jsp>) is the largest source of bacterial-feeding clones. The collection comprises 16,757 clones divided into 52 384-well bacterial clone plates. The ORFeome-RNAi v1.1 library developed in the Vidal Lab (Rual *et al.*, 2004) (<http://www.openbiosystems.com>) contains 11,511 RNAi clones divided into 1308 96-well plates. The latter collection includes clones for ~1700 genes not currently targeted by the Ahringer’s library. Combining both libraries, the overall predicted coverage of the *C. elegans* genome is ~88% of the gene models in WormBase release WS200. Specialized feeding RNAi libraries are also available from Geneservice (<http://www.geneservice.co.uk/>), which collects clones representing genes for specific “processes” such as chromatin (257 clones), phosphatases (166 clones), and transcription factors (387 clones).

### 2. Preparing the Bacterial Library

To assemble a library of the target set from the bacterial-feeding collection, the selected clones must be cherry-picked from their original plates and frozen onto new plates. The plate and position of the clones can be obtained from WormBase (<http://www.wormbase.org>), or from Geneservice ([http://www.geneservice.co.uk/products/tools/Celegans\\_Finder.jsp](http://www.geneservice.co.uk/products/tools/Celegans_Finder.jsp)) using the “*C. elegans* finder” tool for the Ahringer library, or by searching a downloadable Excel file available from Thermo Fisher Scientific Inc. for the ORFeome-RNAi v1.1 library (<http://www.openbiosystems.com/ProductDataFiles.aspx?AliasPath=/GeneExpression/Non-Mammalian/Worm/CelegansORF-RNAi&CatalogNumber=RCE1181>).

The bacterial clones can be re-arrayed manually into new plates by first identifying the positions of the clones of interest, and then retrieving their source plates from the  $-80^{\circ}\text{C}$  freezer (place them on a bucket with dry ice to prevent thawing). Each clone is picked with a single-channel  $200\ \mu\text{L}$  pipette by poking the foil that covers the plate and dabbing the frozen bacterial culture. This culture is streaked onto a LB agar plate with ampicillin at a concentration of  $100\ \mu\text{g}/\text{mL}$  and tetracycline at a concentration of  $12.5\ \mu\text{g}/\text{mL}$  for selection. After 24 h of incubation at  $37^{\circ}\text{C}$ , an individual colony per clone is inoculated in one well of a 96-well deep well plate containing  $800\ \mu\text{L}$  per well of LB broth with ampicillin at a concentration of  $50\ \mu\text{g}/\text{mL}$ . After 24 h of shaking the culture at  $37^{\circ}\text{C}$  and 250 rpm,  $100\ \mu\text{L}$  from each well is transferred to a 96-well round bottom plate, mixed with 90% glycerol in a 1:1 ratio, and frozen at  $-80^{\circ}\text{C}$  for future use.

When complete RNAi libraries are being used as the target set, the ORFeome-RNAi v1.1 library is provided in 96-well microtiter plates and can be used directly for the screening in this format. The Ahringer bacterial collection is provided in 384-well plates and must be re-arrayed in 96-well format. Each 384-well plate will result in four 96-well plates by positioning the first pin of a short 96-pin replicator tool in the wells A1, A2, B1, and B2 of the 384-well plate, and successively “spotting” (not poking) four different LB agar plates with ampicillin at a concentration of  $100\ \mu\text{g}/\text{mL}$  and tetracycline at a concentration of  $12.5\ \mu\text{g}/\text{mL}$  (Note 1). After incubating these plates at  $37^{\circ}\text{C}$  for 24 h, the cultures can be inoculated in liquid using a long 96-pin replicator tool, and the procedure to freeze the new plates is the same as described above (Note 2).

### 3. Large-Scale Worm Amplification

One way to produce a large quantity of worms from a strain maintained on medium ( $60\ \text{mm} \times 15\ \text{mm}$ ) plates is to let the worms grow until they are freshly starved, when there are mostly L1 larvae on the plate. Then, cut out a piece of agar (“chunk”) containing about 500 L1 larvae, and place it on a large plate. Prepare at least 10 plates this way.

When the L1 larvae become gravid adults (about three to four days after seeding, depending on the incubation temperature), follow the pseudo-synchronization protocol (subheading II.A.4), and dispense the resulting L1 larvae on seeded large ( $100\ \text{mm} \times 15\ \text{mm}$ ) or extra-large ( $150\ \text{mm} \times 15\ \text{mm}$ ) NGM plates. In our experience, one extra-large plate corresponds to approximately 10 large plates. One extra-large plate can be seeded with 10,000 L1 worms (Note 3).

### 4. Larvae Pseudo-Synchronization

We use a “bleaching” protocol to obtain thousands of freshly hatched and pseudo-synchronized first-instar larvae (L1 larvae). The procedure is as follows:

1. Wash adult worms from an extra-large plate into a 50 mL conical tube, using about 20 mL of M9 buffer (Note 4).

2. Let the tubes with the worms settle for 10 min at the appropriate temperature (in the 15 °C incubator for most ts strains), and then aspirate the supernatant, being careful not to aspirate the adult worms.
3. Add M9 to 10 mL, and then divide the liquid into two 15 mL conical tubes (Note 5). Centrifuge for 2 mins at 2000 rpm and then aspirate the supernatant. The pellet in each tube must be of similar size.
4. Add 10 mL of bleaching solution. Periodically shake the tubes and observe under the dissecting microscope. After approximately 4 min, the adult worms start to break apart (Note 6).
5. As soon as adult breakage begins, centrifuge for 2 min at 2000 rpm. Aspirate the supernatant as completely and quickly as possible without disturbing the pellet. Total contact with the bleaching solution should not exceed 10 min.
6. Add 10 mL of M9 to each tube. Invert or shake the tube to break up the pellet. Centrifuge for 2 min at 2000 rpm, and then aspirate. Repeat this washing procedure four times.
7. Resuspend in 10 mL M9 and incubate with rocking movement for 24 h (Note 7).

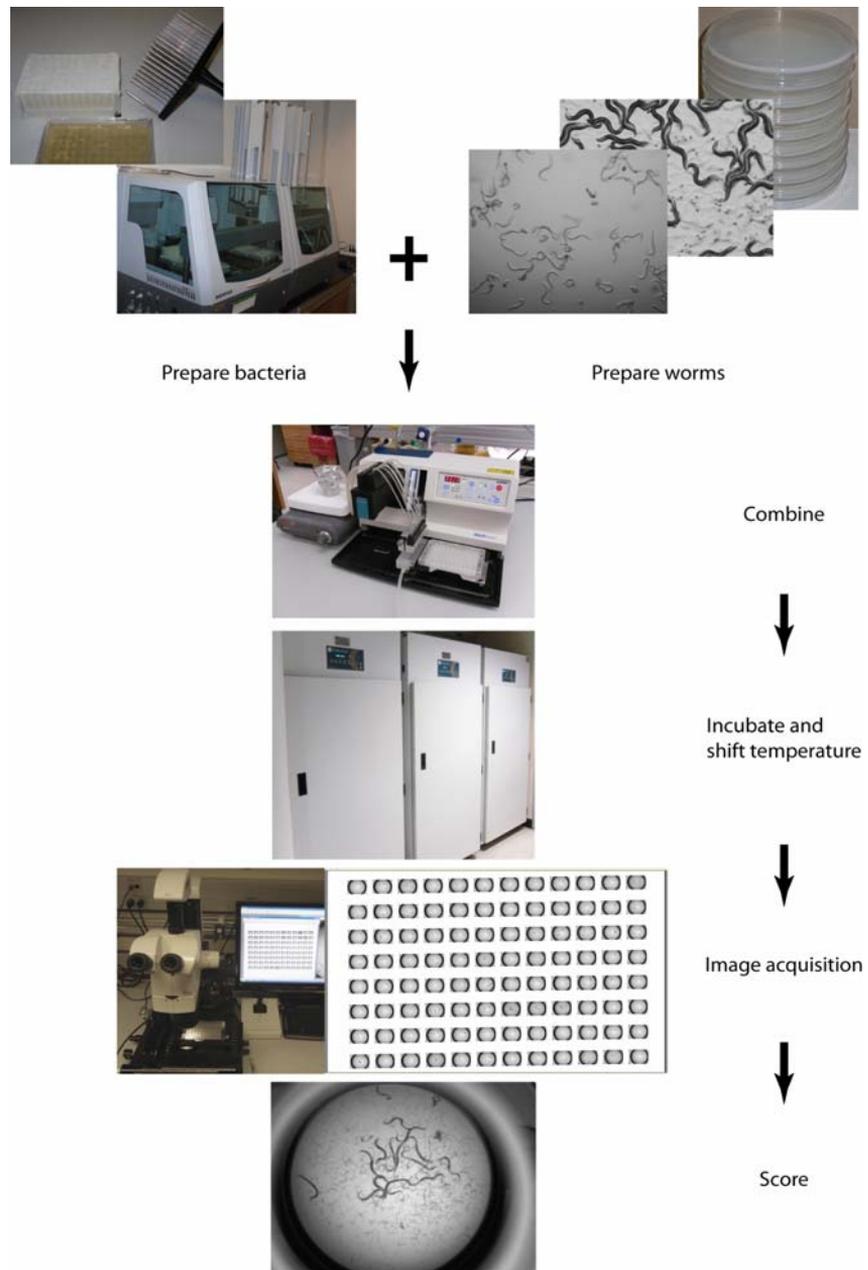
One extra-large plate with 10,000 gravid adults produces about 10 mL of a suspension with 20–60 worms/ $\mu$ L. When this suspension is diluted to 10 worms/20  $\mu$ L, it is enough to dispense L1 larvae in more than 200 96-well screening plates.

## 5. Screening Protocol

Briefly, we first prepare the RNAi bacterial plates, as well as the *C. elegans* strains pseudo-synchronized to the L1 stage. Using a liquid-handling robot (we use a Tecan Aquarius<sup>TM</sup>, equipped with a carousel), the bacteria are dispensed onto plates with flat-bottom wells. Worms are dispensed into these wells using an automated liquid dispenser (we use a Matrix WellMate<sup>®</sup>). Plates are incubated at permissive temperature for three days, allowing the worms to reach the L3–L4 stage. They are then transferred to the appropriate temperature for enhancement or suppression screening. Incubate for 3–5 days, after which the worms will have developed into adults and produced progeny. When the worms clear the wells of bacteria, an image of each well is captured and stored in a database for future analysis. Our protocol is shown schematically in Fig. 3.

### Defrost Bacteria

Bacteria from source plates at –80 °C are replicated to LB agar with ampicillin at a concentration of 100  $\mu$ g/mL and tetracycline at a concentration of 12.5  $\mu$ g/mL on one-well rectangular plates. Replicate the bacteria with a 96-pin tool with short pins, using a bucket with dry ice to hold the frozen stocks. Be careful not to puncture the agar. Incubate the plates at 37 °C for 24 h, and then store at 4 °C until inoculation. Do not store plates at 4 °C for more than 1 week.



**Fig. 3** Procedure for high-throughput genetic interaction screening by RNAi in liquid. (For color version of this figure, the reader is referred to the web version of this book.)

## Amplify Worms

- See subheading II. A. 3

The screening pipeline follows a 7–9 days cycle. Days 8 and 9 are image acquisition days and overlap with days 2 and 3 of the following cycle (Fig. 4).

## Day 1

*Worm pseudo-synchronization*

1. Bleach adult worms (subheading II.A.4).
2. Leave washed eggs to hatch in M9 buffer, rocking overnight at permissive temperature.

*Bacteria inoculation*

1. Prepare inoculation plates by using the liquid dispensing machine (Wellmate®) to dispense 800  $\mu\text{L}$  per well of LB broth with 100  $\mu\text{g}/\text{mL}$  ampicillin on 96-well deep well plates (Note 8). This medium is previously filtered using with 0.22  $\mu\text{m}$  membranes to maintain sterility and to avoid any potential debris that can affect image analysis.
2. Use a sterile “long” 96-pin replicator tool to inoculate bacteria from the LB agar plates into the deep wells. Cover plates with AirPore tape sheets. Shake plates overnight at 37 °C.

## Day 2

*Bacteria induction and re-suspension*

1. Using the Wellmate®, add 80  $\mu\text{L}$  of a 10 mM IPTG solution to each well of bacteria culture. Re-seal the plates with AirPore tape sheets. Continue shaking at 37 °C for 4 h to induce the bacteria to produce double-stranded RNA. (Note 9).
2. Centrifuge deep well plates at 3500 rpm for 5 min, and discard supernatant (Note 10). Using the Wellmate®, add 300  $\mu\text{L}$  of sterilized and filtrated “complete” S-Basal medium supplemented with 100  $\mu\text{g}/\text{mL}$  ampicillin, 0.1  $\mu\text{g}/\text{mL}$  Fungizone, 0.01% Tween, and 1 mM IPTG to the bacterial pellets.
3. Securely cover deep well plates with transparent plastic adhesive plate seals to avoid cross-contamination of bacterial cultures from adjacent wells, and then carefully vortex to re-suspend the bacteria in the S-Basal complete medium. Continue re-suspension with multichannel pipette if necessary.
4. Using the liquid-handling robot Aquarius™ from Tecan, dispense 30  $\mu\text{L}$  per well of bacteria from the 96-well deep well plates to 96-well flat-bottom plates (Note 11).

*Worm re-suspension and seeding*

5. Filter L1 worms from the bleached worm suspension with 40  $\mu\text{m}$  cell strainers to remove any debris, recovering the filtrate in 50 mL conical tubes (Note 12).

	Day 1- cycle 1	Day 2-cycle 1	Day 3-cycle 1	Day 4-cycle 1	Day 5- cycle 1	Day 6- cycle 1
	1. Worm pseudo-synchronization 2. Bacteria inoculation	1. Bacteria induction and re-suspension 2. Worm re-suspension and seeding			Temperature shift of screening plates	
Defrost Bacteria-cycle 1		Worm amplification at 15°C for cycle 2				
Day 7- cycle 1	Day 8- cycle 1 Day 1- cycle 2	Day 9- cycle 1 Day 2-cycle 2	Day 3-cycle 2	Day 4-cycle 2	Day 5- cycle 2	Day 6- cycle 2
Image acquisition	Image acquisition	Image acquisition				
	1. Worm pseudo-synchronization 2. Bacteria inoculation	1. Bacteria induction and re-suspension 2. Worm re-suspension and seeding			Temperature shift of screening plates	
Defrost Bacteria-cycle 2		Worm amplification at 15°C for cycle 3				

**Fig. 4** A typical schedule of two-week activities for high-throughput RNAi screening showing two overlapping cycles.

6. Wash L1 larvae twice with 10 mL of complete S-Basal. After each wash, centrifuge for 2 min at 2000 rpm and remove the supernatant using a vacuum aspirator. Leave about 0.5 mL of liquid, because the worms will not form a solid pellet.
7. Dilute the suspension of L1 larvae in complete S-Basal to 10 worms/20  $\mu$ L. To determine the dilution factor, check several samples in a stereo-microscope, counting the number of L1 larvae in a 10  $\mu$ L drop (Note 13).
8. Using the Wellmate<sup>®</sup>, add 20  $\mu$ L of diluted worm suspension ( $\sim$ 10 worms) to each well of the 96-well flat-bottom plates (Note 14).
9. Incubate plates in humid chambers at permissive temperature (Note 15).

#### Day 5

##### *Temperature shift*

When the worms are at the L3–L4 stage, transfer the plates from permissive temperature to either semi-permissive temperature for enhancement or to semi-restrictive temperature for suppression. Continue incubation until the worms reach adulthood and produce progeny, and the bacteria clear from the wells. The amount of time is variable depending on the particular incubation temperature (see following image acquisition times).

#### Day 7

##### *Image acquisition*

Photograph N2 worms that were shifted to 22.5 °C or higher temperature.

#### Day 8

##### *Image acquisition*

Photograph N2 or mutant worms that were shifted to the semi-permissive temperature of 20 °C or higher, and mutants shifted to restrictive temperature.

#### Day 9

##### *Image acquisition*

Photograph control or mutant worms shifted to the semi-permissive temperature of 17.5 °C or lower.

## 6. Image Acquisition

To acquire an image of each well of the experiment, we built a system around a Z16 dissecting scope with a DFC340 FX camera (both from Leica Microsystems Inc.), a Bio-precision motorized stage from Ludl, Inc. (with adaptors for the 96-well plates and stage fittings), and the Surveyor software from Media Cybernetics, Inc. to control the camera and the stage.

After testing different exposure times, we found that 1.2 ms was ideal to avoid shading from the rapid movement of the worms in the pictures. With this exposure time, we acquire 96 high-quality individual images from a 96-well plate in less than 1 min. We use a magnification of 17.3 $\times$ , which not only provides sufficient detail for image analysis ( $>30$  pixels/embryo), but also permits most of the well to fit within the image frame.

## 7. Notes

1. Pay close attention not to thaw the original plate. Always keep it over dry ice. Before grabbing the bacterial culture with the short 96-pin tool, remove the foil cover, and cover with a new one when finished.
2. The 96-pin tool must be sterilized after each inoculation. We use three empty tip boxes: the first filled with commercial bleach, the second with water (to rinse the bleach), and the third with ethanol. We dip the tool successively in these liquids and then quick-flame it. Special caution must be taken when flaming the tool. Never leave the 96-pin tool soaking in any liquid as it will damage the pins.

The deep wells can also be inoculated using a multichannel pipette and a box of tips for each plate. This method increases the time and cost necessary for the process.
3. Using extra-large plates reduces the space needed for the plates in the incubator, and reduces the time needed to wash the adult worms from the plates during the pseudo-synchronization protocol.
4. Washed worms from two plates can be collected in a 50 mL conical tube. If using large plates, about 5 mL of M9 buffer is enough to wash one plate. Letting the adult worms settle by gravity, and then removing the supernatant by aspiration will eliminate most of the smaller larvae.
5. During centrifugation, worms pellet more quickly and compactly in 15 mL conical tubes than in 50 mL conical tubes.
6. The time in contact with the bleaching solution varies based on worm age and strain, as well as the brand or batch of bleach being used. If eggs are observed popping from the worms before 4 min, proceed to centrifuge the samples immediately.
7. To maximize larval hatching after the bleaching, remove as much of the supernatant as possible after each wash to remove any trace of bleaching solution.
8. The liquid dispensing can be done manually using a multichannel pipette. The liquid dispensing machine helps to reduce the dispensing time when preparing many 96-well deep well plates.
9. IPTG can be stored frozen in 0.5 M or 1 M aliquots and diluted with sterile water.
10. Remove the supernatant by turning the deep well plates upside down over a container. Then tap the plate upside down on paper towels to remove as much LB as possible. This will help prevent further growth of the bacterial culture.
11. A small number of plates can be dispensed with manual multichannel pipettes; the number can be increased with electronic repetitive multichannel pipettes (i.e., Matrix 12-channel electronic multichannel 15–850  $\mu$ L). For large number of plates, this step is very labor intensive. We have achieved a maximum throughput of dispensing about 500 96-well flat-bottom plates in 3 h using the Aquarius<sup>TM</sup> from Tecan, Inc., a liquid-handling robot with a 96-well head.

12. Extra L1 larvae from day 3 can be used to seed extra-large plates for the following week's experiment. In our lab, we plate 10,000 L1 larvae on each extra-large plate, and we prepare four plates for each strain. This produces enough L1 larvae to dispense more than 200 flat-bottom plates and to seed NGM plates for the following week's experiment.
13. To count the number of hatched L1 larvae in the suspension, dilute 100  $\mu\text{L}$  of larvae suspension with 900  $\mu\text{L}$  of M9, and mix well. From this dilution, place three drops of 10  $\mu\text{L}$  on the lid of a medium petri plate. Count the number of larvae in each drop, and calculate the dilutions necessary to achieve a final concentration of 10 larvae/20  $\mu\text{L}$ . Dilute with complete S-Basal. One flat-bottom plate requires about 2 mL of larvae suspension to fill all wells.

When calculating the volume of larvae suspension to dilute, consider that the Wellmate® has a 10 mL dead volume in the cartridge, and depending on the diameter of the container used to dispense the worms, extra volume is needed to avoid aspirating air into the tubing.

14. To dispense the larvae suspension, we use a small sterile beaker or a disposable sterile plastic container with a stirring bar to continuously mix the suspension. This keeps the suspension homogeneous throughout the process. We have found that with the dilution we use, 92% of our assay wells show a range of 6–16 worms, and all look healthy.
15. Assemble the humid chambers using plastic Tupperware containers lined with wet paper towels. It is important to make sure that the containers are hermetically sealed with Parafilm, especially if incubating at higher temperatures, to avoid any potential dehydration of the samples.

## B. Scoring of Interactions/Phenotypes

### 1. Visual

Before scoring the images, define the variables to be evaluated and the method of evaluation. As a general guide, we describe below the criteria we have developed in our laboratory to score genetic enhancers and suppressors, though each laboratory will most likely want to develop customized assay protocols for their own purposes.

For each experiment, we analyze two technical replicates (test plates A and B) with the ts mutant fed on RNAi bacteria targeting a gene of interest (“double-knock-downs”). These are compared to three controls: N2 strain fed on the empty vector bacteria L4440 (WT worms), N2 fed on the same RNAi bacteria as the test plates (effect of the RNAi alone), and the mutant strain fed on L4440 (effect of the mutation). To facilitate scoring, we usually open a set of images from the same experiment together (this can be done using a variety of common image visualization programs), which allows us to scroll quickly through the results of all 96 wells from the same plate. We record all phenotype scores in an Excel worksheet designed to mimic the layout of a 96-well plate.

## Scoring Enhancement by Visual Inspection

We first examine the images of the *ts* mutant fed on the L4440 empty vector control. We typically inspect at least 48 images to gain a sense of the variation in number of adults and progeny (larvae/eggs) per well in the absence of any RNAi effect. Usually, a *ts* mutant strain fed on the L4440 control at semi-permissive temperature will produce numerous larvae and few eggs.

Next, we score the 96 images of the first double-knock-down plate (plate A), comparing always with what is expected from the L4440 control. Any wells that resemble the *ts* mutant fed on the L4440 control are labeled “not enhanced” (NE). We note any differences in viability of the progeny and brood size, recording Embryonic lethality (Emb), Sterility (Ste), and deviations in the number of larvae and eggs. We also score any detectable post-embryonic phenotypes (including larval arrest (Lva), larval lethality (Lvl), Growth arrest (Gro), ruptured worms (Rup), protruding vulva in adults (Pvl), and egg laying defective (Egl). We then score the 96 images of the second test plate for the same double-knock-downs (plate B) in the same way.

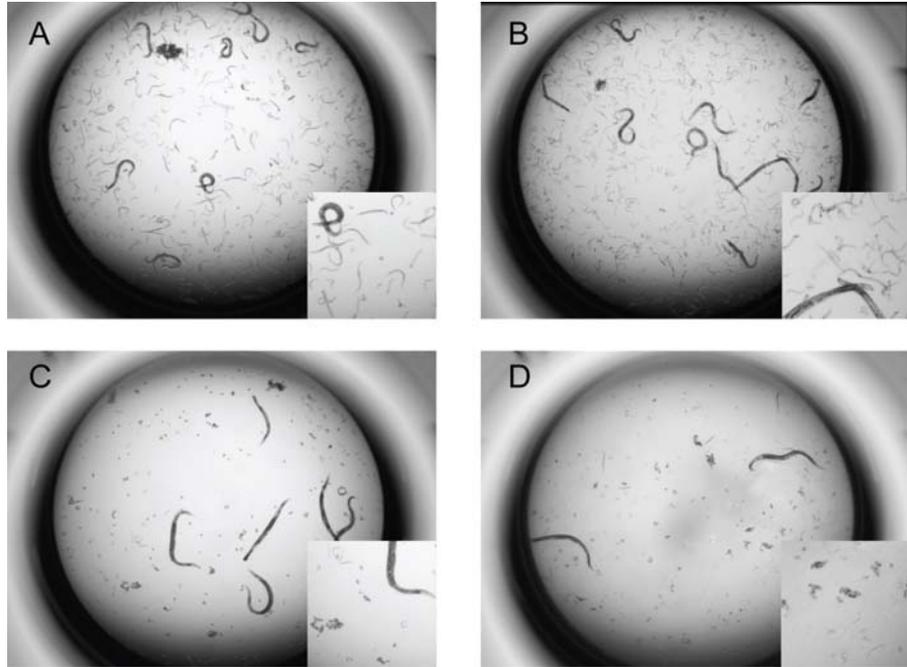
From the scores of plates A and B, any image with a phenotype deviating from the L4440 control is compared to the corresponding image of N2, which was fed on the same RNAi bacteria. If the N2 RNAi looks WT, then we score the target gene in the double-knock-down as a putative enhancer, which we place into one of the four categories: we score a result as “strong” if it shows Emb, Ste, or a highly reduced number of larvae (VLL); as “medium” if it shows a higher proportion of unhatched embryos (>E, typically signifying incomplete penetrance lethality) or a smaller brood size (identified as fewer larvae (LL) or less embryos (<E) depending on the screening temperature, signifying either incomplete sterility or reduced fertilization); as “weak” if the N2 control shows some RNAi effect or if no conclusive score can be assigned (?); or as “post-embryonic enhancement” (PEE) if we detect any post-embryonic phenotypes not present in the N2 RNAi. Fig. 5 shows an example of an enhancement interaction.

Once we obtain a score for both replicates from each experiment (copies A and B), we perform secondary screening of all putative strong, medium, and post-embryonic enhancers identified in at least one replicate and of putative weak enhancers identified in both replicates.

## Scoring Suppression by Visual Inspection

As for scoring enhancement, we first examine ~48 images of the *ts* mutant fed on control vector L4440 to gain a sense of the variation among the wells. Usually, a *ts* mutant strain fed on the L4440 control at semi-restrictive temperature will produce numerous eggs and few larvae, if any.

Next, we score the images from both experimental replicates in comparison with the expected phenotype from the control L4440 RNAi plate. We record results with

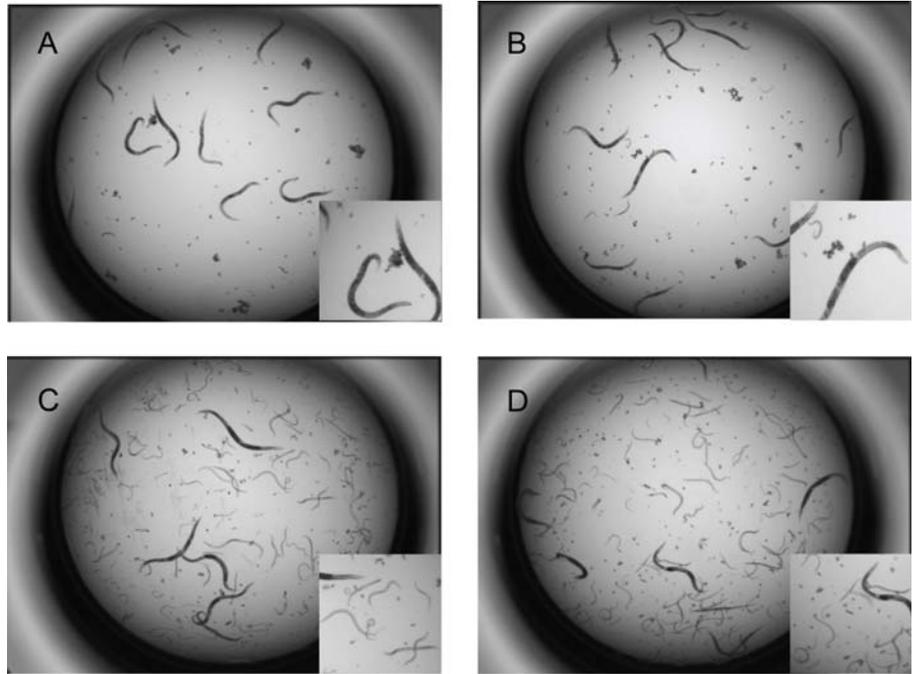


**Fig. 5** Example of a positive enhancement interaction. (A) Control for the mutation: *zyg-8(b235) L4440 (RNAi)* at 22.5 °C is not lethal. (B) Control for RNAi of *egl-27*: *egl-27 (RNAi)* on N2 at 22.5 °C is not lethal. (C–D) Lethality observed for the interaction *zyg-8(b235) egl-27 (RNAi)* at 22.5 °C repeats A and B, respectively. Insets are enlarged sections of the pictures to show details.

no detectable increase in survival as “not suppressed” (NS) and place putative suppressors into one of the four categories based on the number of larvae produced: “strong” (SUP, numerous larvae), “medium” (sup, more larvae than the control), “weak” (sup?, few larvae but more than the control), or “post-embryonic” (PE). As for scoring enhancers, we perform secondary screening of all putative strong, medium, and post-embryonic suppressors identified in at least one replicate and of putative weak suppressors identified in both replicates. Fig. 6 shows an example of a suppression interaction.

## 2. Automated Computerized Image Analysis

Visual inspection can be very laborious when the scale is too large. Some algorithms are being developed for similar but not identical images (Fontaine *et al.*, 2006; Geng *et al.*, 2004; Huang *et al.*, 2006; O’Rourke *et al.*, 2009). An algorithm for the automated analysis of images is being developed in our lab to streamline this process and provide quantitative results (White *et al.*, 2010).



**Fig. 6** Example of a positive suppression interaction. (A–B) Repeats of the control for the mutation: *emb-30(g53)* at 25 °C is lethal. (C–D) Suppression of lethality of *emb-30(g53)* at 25 °C by the RNAi of the gene *smg-1*, repeats A and B, respectively. Insets are enlarged sections of the pictures to show details.

### C. Quantification of Embryonic Lethality or Survival on Agar Plates

The positive results in the large-scale screen in liquid for enhancement or suppression can be repeated on agar plates to confirm and quantify the results.

#### Day 1

Replicate bacteria from –80°C stock to LB agar with 50 mg/L ampicillin and 10 mg/L tetracycline rectangular plates and place overnight at 37 °C in a dry incubator.

#### Day 2

Grow bacteria in liquid LB with 50 mg/L ampicillin shaking at 250 rpm for 12–16 h, but no more than 18 h.

#### Day 3

1. For each worm strain-RNAi clone pair to be tested (including the empty vector, L4440), seed three RNAi plates with 30 µL of bacteria and one RNAi plate with 300 µL of bacteria, and let dry for 24 h.

2. Wash adult worms from agar plates with M9 buffer solution and bleach them (see larvae pseudo-synchronization protocol, subheading II.A.4). Let eggs hatch overnight at 15 °C.

#### Day 4

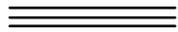
1. Count the number of L1 worms hatched per 10  $\mu\text{L}$  of solution and dispense about 50 worms into the RNAi plate previously seeded with 300  $\mu\text{L}$  of bacteria.
2. Incubate the L1 larvae at 15 °C until they reach the L4 stage (about 3 days).

#### Day 7

Move 10 L4 larvae to each of the three RNAi plates previously seeded with 30  $\mu\text{L}$  of each bacterial clone. Incubate plates at the appropriate temperature, until the L4 larvae become adults and lay eggs.

#### Day 9 or after

After this incubation period, remove the adult worms, and replace plates at the appropriate temperature. After 24 h, count the number of larvae and eggs to calculate the percentage of hatched larvae.



### III. Materials

#### A. Reagents

1. Ahringer RNAi feeding library – Geneservice Ltd, UK (<http://www.geneservice.co.uk/products/rnai/Celegans.jsp>)
2. Orfeome V1.1 RNAi feeding library – Thermo Fisher, USA (<http://www.openbiosystems.com>)
3. *C. elegans* strains and OP50 bacteria – the *C. elegans* Genetics Center (<http://www.cbs.umn.edu/CGC/>)

#### B. Buffers and Media

1. NGM agar: 3 g NaCl, 2.5 g peptone, 17 g agar, to 973 mL with water. Sterilize by autoclaving. Let it cool down. While it is still warm, add the following sterile solutions: 1 mL of 1 M  $\text{CaCl}_2$ , 1 mL of 1 M  $\text{MgSO}_4$ , and 25 mL of 1 M buffer phosphate, pH 6.0. Dispense 28 mL for large plates (100 mm  $\times$  15 mm) and 55 mL for extra-large plates (150 mm  $\times$  15 mm).
2. Seeded extra-large NGM plates: inoculate LB broth with freshly streaked OP50. Incubate overnight at 37 °C, 250 rpm. Pellet bacteria (3500 rpm for 20 min) and concentrate it about 25 times. Under the laminar flow chamber and using a glass bacterial spreader, distribute 2 mL of bacterial suspension over the surface of the NGM plate. Let dry overnight at room temperature and store at 4 °C until use.
3. M9 buffer: 5 g NaCl, 3 g  $\text{KH}_2\text{PO}_4$ , 6 g  $\text{Na}_2\text{HPO}_4$ , water to 999 mL. Sterilize by autoclaving. Add 1 mL of sterile 1 M  $\text{MgSO}_4$ .

4. S-Basal buffer: 5.85 g NaCl, 50 mL 1 M buffer phosphate pH 6.0, water to 1 L. Sterilize by autoclaving.
5. 1 M buffer phosphate pH 6.0: For 1 L, mix 868 mL 1 M  $\text{KH}_2\text{PO}_4$  with 132 mL 1 M  $\text{K}_2\text{HPO}_4$ .
6. 1 M  $\text{KH}_2\text{PO}_4$ : 136.09 g  $\text{KH}_2\text{PO}_4$ , water to 1 L. Sterilize by autoclaving.
7. 1 M  $\text{K}_2\text{HPO}_4$ : 125.4 g  $\text{K}_2\text{HPO}_4$ , water to 1 L. Sterilize by autoclaving.
8. S-Basal complete: To 970 mL of S-Basal buffer add: 10 mL 1 M potassium citrate, 3 mL 1 M  $\text{MgSO}_4$ , 3 mL 1 M  $\text{CaCl}_2$ , and 10 mL Trace metals solution. Sterilize by filtration. Add 1 mL of 5 mg/mL cholesterol dissolved in ethanol, 0.5 mL 100  $\mu\text{g}/\mu\text{L}$  ampicillin, 0.4 mL 250  $\mu\text{g}/\text{mL}$  Fungizone, 1 mL 1 M IPTG, and 1 mL 10% Tween-20. This can be prepared the day before and stored at 15 °C.
9. Trace metals solution: Disodium EDTA 1.86 g,  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  0.69 g,  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$  0.20 g,  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$  0.29 g,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  0.025 g. Dissolve in 1 L water; aliquot into 50 mL tubes and store in dark.
10. LB broth: 5 g yeast extract, 10 g tryptone, 10 g NaCl, water to 1 L. Sterilize by autoclaving. To grow the bacteria for the RNAi experiments, LB broth is supplemented with ampicillin at a concentration of 50  $\mu\text{g}/\text{mL}$ .
11. LB agar with ampicillin and tetracycline plates: Add 16 g/L agar to LB broth. Sterilize by autoclaving. After it has cooled (just before pouring), add ampicillin at a concentration of 100  $\mu\text{g}/\mu\text{L}$  and tetracycline at a concentration of 12.5  $\mu\text{g}/\mu\text{L}$ . Pour 45 mL into each 86 × 128 mm rectangular plate. Store at 4 °C wrapped in aluminum foil until use.
12. Bleaching solution (100 mL): 4 mL 10 M NaOH, 20 mL Clorox Germicidal Bleach (6.15% sodium hypochlorite), 76 mL autoclaved water. Prepare fresh every time.

### C. Supplies and Equipment

1. Petri dishes by Fisherbrand from Fisher
  - a. 60 × 15 mm Stackable Mono Plate (#08-757-13A)
  - b. 100 × 15 mm Stackable Mono Plate (#08-757-12)
  - c. 150 × 15 mm Slippable Mono Plate (#08-757-14)
2. OmniTray 86 × 128 mm rectangular plates by Nunc (#264728)
3. 96-well flat-bottom plates by Costar, Corning (#3598)
4. Matrix 96-well deep-well plates Deep Well/2 mL Blocks by Thermo Scientific (#4222)
5. BD Falcon cell strainers, nylon mesh size: 40  $\mu\text{m}$ , from Fisher
6. Stericups w/500 mL funnel, 45 mm neck size, PES membrane, Millipore filters (#SCGP-T05-RE)
7. Fisherbrand 4oz/118 mL containers (#14-375-147)
8. AirPore tape sheets, from Quiagen (#1957)
9. Short 96-pin replicator (V&P Scientific, Inc. model VP 381)
10. Long 96-pin replicator (V&P Scientific, Inc. model VP408A)

11. Benchtop Sorvall Legend RT Centrifuge with adapters for plate centrifugation
12. 15–25 °C Percival Worm incubators
13. 37 °C bacteria incubator
14. 37 °C shaking incubator
15. Small rocking platform
16. Matrix WellMate (ThermoFisher) + Tubing cassettes
17. Aquarius (Tecan)
18. Corning aluminum sealing tapes for 96-well plates (#6570)
19. Abgene transparent adhesive plate seals (#AB-0580)
20. Clorox Germicidal Bleach (6.15% sodium hypochlorite), from Fisher.
21. Surveyor software (Media Cybernetics Inc.)
22. Leica stereoscopic microscopes with Luld Bioprecision stage and adaptors

## Acknowledgments

The authors thank Eliana Munarriz for suggestions to the protocol, Kris Gunsalus, Katherine Erickson, and Annalise Paaby for critical reading and editing of the manuscript, Carrie Nygard and Gina Destefano for proofreading, and Amelia White and Huey-Ling Kao for suggestions to the figures. This work has been supported by NIH grant #5R01HD046236-08.

## References

- Ahringer (ed.), J. (2006). Reverse Genetics. the *C. elegans* Research Community. *Wormbook. Wormbook*. 1551–8507.
- Ben-Yakar, A., and Bourgeois, F. (2009). Ultrafast laser nanosurgery in microfluidics for genome-wide screenings. *Curr. Opin. Biotechnol.* **20**, 100–105.
- Ben-Yakar, A., Chronis, N., and Lu, H. (2009). Microfluidics for the analysis of behavior, nerve regeneration, and neural cell biology in *C. elegans*. *Curr. Opin. Neurobiol.* **19**, 561–567.
- Boone, C., Bussey, H., and Andrews, B. J. (2007). Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449.
- Burns, A. R., Kwok, T. C., Howard, A., Houston, E., Johanson, K., Chan, A., Cutler, S. R., McCourt, P., and Roy, P. J. (2006). High-throughput screening of small molecules for bioactivity and target identification in *Caenorhabditis elegans*. *Nat. Protoc.* **1**, 1906–1914.
- Byrne, A. B., Weirauch, M. T., Wong, V., Koeva, M., Dixon, S. J., Stuart, J. M., and Roy, P. J. (2007). A global analysis of genetic interactions in *Caenorhabditis elegans*. *J. Biol.* **6**, 8.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z. Y., Liang, W., Marback, M., Paw, J., San Luis, B. J., Shuteriqi, E., Tong, A. H., van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibzadeh, S., Papp, B., Pal, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A. C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., and Boone, C. (2010). The genetic landscape of a cell. *Science* **327**, 425–431.
- Davierwala, A. P., Haynes, J., Li, Z., Brost, R. L., Robinson, M. D., Yu, L., Mnaimneh, S., Ding, H., Zhu, H., Chen, Y., Cheng, X., Brown, G. W., Boone, C., Andrews, B. J., and Hughes, T. R. (2005). The synthetic genetic interaction spectrum of essential genes. *Nat. Genet.* **37**, 1147–1152.
- Doitsidou, M., Flames, N., Lee, A. C., Boyanov, A., and Hobert, O. (2008). Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nat. Methods* **5**, 869–872.
- Fernandez, A. G., Gunsalus, K. C., Huang, J., Chuang, L. S., Ying, N., Liang, H. L., Tang, C., Schetter, A. J., Zegar, C., Rual, J. F., Hill, D. E., Reinke, V., Vidal, M., and Piano, F. (2005). New genes with roles in

- the *C. elegans* embryo revealed using RNAi of ovary-enriched ORFeome clones. *Genome Res.* **15**, 250–259.
- Fernandez, A. G., Mis, E. K., Bargmann, B. O., Birnbaum, K. D., and Piano, F. (2010). Automated sorting of live *C. elegans* using laFACS. *Nat. Methods* **7**, 417–418.
- Fontaine, E., Burdick, J., and Barr, A. (2006). Automated tracking of multiple *C. elegans*. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1**, 3716–3719.
- Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408**, 325–330.
- Geng, W., Cosman, P., Berry, C. C., Feng, Z., and Schafer, W. R. (2004). Automatic tracking, feature extraction and classification of *C. elegans* phenotypes. *IEEE Trans. Biomed. Eng.* **51**, 1811–1820.
- Gonczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S., Flohrs, K., Goessen, A., Leidel, S., Alleaume, A. M., Martin, C., Ozlu, N., Bork, P., and Hyman, A. A. (2000). Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* **408**, 331–336.
- Gunsalus, K. C., and Piano, F. (2005). RNAi as a tool to study cell biology: building the genome-phenome bridge. *Curr. Opin. Cell. Biol.* **17**, 3–8.
- Holway, A. H., Hung, C., and Michael, W. M. (2005). Systematic RNA-interference-mediated identification of mus-101 modifier genes in *Caenorhabditis elegans*. *Genetics* **169**, 1451–1460.
- Huang, K. M., Cosman, P., and Schafer, W. R. (2006). Machine vision based detection of omega bends and reversals in *C. elegans*. *J. Neurosci. Methods* **158**, 323–336.
- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237.
- Kwok, T. C., Ricker, N., Fraser, R., Chan, A. W., Burns, A., Stanley, E. F., McCourt, P., Cutler, S. R., and Roy, P. J. (2006). A small-molecule screen in *C. elegans* yields a new calcium channel antagonist. *Nature* **441**, 91–95.
- Lehner, B., Crombie, C., Tischler, J., Fortunato, A., and Fraser, A. G. (2006). Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. *Nat. Genet.* **38**, 896–903.
- Maeda, I., and Sugimoto, A. (2001). Systematic functional analysis of the *C. elegans* genome. *Tanpakushitsu Kakusan Koso* **46**, 2432–2435.
- Min, J., Kyung Kim, Y., Cipriani, P. G., Kang, M., Khersonsky, S. M., Walsh, D. P., Lee, J. Y., Niessen, S., Yates 3rd, J. R., Gunsalus, K., Piano, F., and Chang, Y. T. (2007). Forward chemical genetic approach identifies new role for GAPDH in insulin signaling. *Nat. Chem. Biol.* **3**, 55–59.
- Mohr, S., Bakal, C., and Perrimon, N. (2010). Genomic screening with RNAi: results and challenges. *Annu. Rev. Biochem.* **79**, 37–64.
- O'Rourke, E. J., Conery, A. L., and Moy, T. I. (2009). Whole-animal high-throughput screens: the *C. elegans* model. *Methods Mol. Biol.* **486**, 57–75.
- Perrimon, N., and Mathey-Prevot, B. (2007). Applications of high-throughput RNA interference screens to problems in cell and developmental biology. *Genetics* **175**, 7–16.
- Piano, F., Schetter, A. J., Mangone, M., Stein, L., and Kempthues, K. J. (2000). RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. *Curr. Biol.* **10**, 1619–1622.
- Rockman, M. V., and Kruglyak, L. (2009). Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* **5**, e1000419.
- Rual, J. F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S., and Vidal, M. (2004). Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162–2168.
- Sonnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A. M., Artelt, J., Bettencourt, P., Cassin, E., Hewitson, M., Holz, C., Khan, M., Lazik, S., Martin, C., Nitzsche, B., Ruer, M., Stamford, J., Winzi, M., Heinkel, R., Roder, M., Finell, J., Hantsch, H., Jones, S. J., Jones, M., Piano, F., Gunsalus, K. C., Oegema, K., Gonczy, P., Coulson, A., Hyman, A. A., and Echeverri, C. J.

- (2005). Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469.
- Stoeckius, M., Maaskola, J., Colombo, T., Rahn, H. P., Friedlander, M. R., Li, N., Chen, W., Piano, F., and Rajewsky, N. (2009). Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. *Nat. Methods* **6**, 745–751.
- Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156.
- Tischler, J., Lehner, B., Chen, N., and Fraser, A. G. (2006). Combinatorial RNA interference in *Caenorhabditis elegans* reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biol.* **7**, R69.
- Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G. F., Brost, R. L., Chang, M., Chen, Y., Cheng, X., Chua, G., Friesen, H., Goldberg, D. S., Haynes, J., Humphries, C., He, G., Hussein, S., Ke, L., Krogan, N., Li, Z., Levinson, J. N., Lu, H., Menard, P., Munyana, C., Parsons, A. B., Ryan, O., Tonikian, R., Roberts, T., Sdicu, A. M., Shapiro, J., Sheikh, B., Suter, B., Wong, S. L., Zhang, L. V., Zhu, H., Burd, C. G., Munro, S., Sander, C., Rine, J., Greenblatt, J., Peter, M., Bretscher, A., Bell, G., Roth, F. P., Brown, G. W., Andrews, B., Bussey, H., and Boone, C. (2004). Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813.
- van Haften, G., Vastenhouw, N. L., Nollen, E. A., Plasterk, R. H., and Tijsterman, M. (2004). Gene interactions in the DNA damage-response pathway identified by genome-wide RNA-interference analysis of synthetic lethality. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12992–12996.
- White A., et al., (13–18 June 2010). Rapid and accurate developmental stage recognition of *C. elegans* from high-throughput image data. *2010 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)* 3089–3096.

---

---

## CHAPTER 5

# Dissection of Genetic Pathways in *C. elegans*

**Zheng Wang and David R. Sherwood**

Department of Biology, Duke University, Durham, North Carolina, USA

---

Abstract

- I. Introduction
  - II. Strategies for Identifying Components in Pathways of Interest
    - A. Forward Genetic Screens
    - B. Functional Genomic Approaches to Identify Components of Pathways
    - C. Using Systems Biology Approaches to Identify Components of Pathways
  - III. Ordering Genes into Pathways
    - A. Determining Whether Two Genes Function in the Same Pathway
    - B. Genetic Ordering of Pathways
    - C. Properties of Gene Products for Gene Ordering
    - D. Localization Dependency of Gene Products for Gene Ordering
    - E. Gene Expression Dependency for Gene Ordering
    - F. Functional Genomic Approaches for Gene Ordering
  - IV. Future Outlook
- Acknowledgments  
References

---

---

### Abstract

With unique genetic and cell biological strengths, *C. elegans* has emerged as a powerful model system for studying many biological processes. These processes are typically regulated by complex genetic networks consisting of genes. Identifying those genes and organizing them into genetic pathways are two major steps toward understanding the mechanisms that regulate biological events. Forward genetic screens with various designs are a traditional approach for identifying candidate genes. The completion of the genome sequencing in *C. elegans* and the advent of high-throughput experimental techniques have led to the development of two additional powerful approaches: functional genomics and systems biology. Genes that are discovered by these approaches can be ordered into interacting pathways through a variety of

strategies, involving genetics, cell biology, biochemistry, and functional genomics, to gain a more complete understanding of how gene regulatory networks control a particular biological process. The aim of this review is to provide an overview of the approaches available to identify and construct the genetic pathways using *C. elegans*.

---

---

---

## I. Introduction

*C. elegans* has emerged as a powerful model system for identifying the genes and genetic pathways that regulate a diverse array of fundamental biological processes. The strengths of *C. elegans* include its invariant cell lineage, simplified cellular landscape, transparency, short life cycle, and hermaphroditic reproduction, which favor rapid genetic analysis. Complementing these traditional attributes, more recent functional genomic technologies, including genome-scale transcriptional and phenotypic profiling as well as physical interaction mapping, have led to more instruments in the *C. elegans* researcher's tool box to uncover the genetic networks that control developmental, behavioral, cell biological, and physiological processes. Highlighting the utility of *C. elegans* is the leading role that this model system has played in elucidating the genetic pathways regulating key biological processes such as cell-fate specification (Greenwald *et al.*, 1983; Sternberg, 2004; Sternberg and Horvitz, 1986), apoptosis (Kimble and Hirsh, 1979; Metzstein *et al.*, 1998; Sulston *et al.*, 1983; Sulston and Horvitz, 1977), RNA interference (RNAi) (Fire *et al.*, 1998), microRNA biology (Reinhart *et al.*, 2000; Simon *et al.*, 2008), axon guidance (Chan *et al.*, 1996; Hao *et al.*, 2001; Walthall and Chalfie, 1988), cell polarity (Goldstein and Hird, 1996; Kemphues *et al.*, 1988), and aging (Friedman and Johnson, 1988; Kimura *et al.*, 1997).

The goal of this review is to provide a comprehensive overview of the approaches available to *C. elegans* researchers to identify and construct the genetic pathways that control biological events. Often these start with a simple genetic screen to identify the key nonredundant genes that regulate a process of interest. Alternatively, researchers sometimes stumble into a biological process when conducting reverse genetics – knocking out or reducing the function of a gene of interest (often disease-related) and studying the resulting phenotype. Once key genes are identified, reverse genetics and sensitized screening approaches can be used to better focus and identify redundant or modulatory genes. These approaches can often be complemented with functional genomic and systems-level studies to gain a more complete understanding of genetic pathways and the networks that guide these processes. In this review, we first discuss representative designs underlying diverse genetic screens in *C. elegans* with a uniform aim of identifying genes involved in a biological process of interest. We then discuss new techniques being utilized to discover candidate genes, including functional genomic techniques and their integration into systems-level analysis. This is followed by an outline of strategies for constructing pathways with genes identified. Throughout we provide specific examples of screening and genetic pathway construction to illustrate how these techniques are implemented.

---

---

---

## II. Strategies for Identifying Components in Pathways of Interest

Three general strategies are currently used by *C. elegans* biologists to discover the specific components of pathways: (1) forward genetic screens that encompass a diverse array of screening methods and can be complemented with reverse genetic approaches; (2) functional genomic approaches that utilize genome-wide analysis of transcription, protein–protein interaction, DNA-binding site analysis, and loss-of-function techniques; and (3) systems biology approaches that integrate the use of functional genomic techniques.

### A. Forward Genetic Screens

#### 1. Forward Genetics and Reverse Genetics

Biological processes are precisely regulated by highly coordinated gene regulatory networks that are comprised of numerous interacting signaling pathways. The basic constituent of signaling pathways are proteins encoded by corresponding genes. Perturbation of these genes could cause deregulation of associated pathways or even the entire network. Deregulation sufficient to disrupt a biological function may result in an observable outcome referred to as a phenotype. The collective status of perturbed genes is called genotype. The biological relationship in which a genotype determines a phenotype is the foundation for dissecting genetic pathways.

Forward genetic approaches (investigation directed from phenotype to genotype) and reverse genetic approaches (investigation directed from genotype to phenotype) are two powerful ways of elucidating the function of genes that regulate a biological process of interest. Forward genetic screens identify genes in an unbiased manner based on phenotypes of mutants. The screens start by searching for a desired phenotype caused by a mutation that is introduced into a gene by mutagens, such as EMS (1-methylsulfonyloxyethane, also known as ethyl methanesulfonate) or ENU (1-ethyl-1-nitrosourea, also known as N-ethyl-N-nitrosourea). Identity of the mutation-harboring genes can be determined by positional cloning or candidate-gene testing. Forward genetic studies in *C. elegans* have made significant contributions to our understanding of a wide range of developmental processes. For example, the forward genetic screens pioneered by Nobel Prize Laureate Robert Horvitz for mutants defective in programmed cell death (PCD) identified the underlying genetic pathways that direct apoptosis, a process conserved among metazoans, including humans (Metzstein *et al.*, 1998).

Similarly, reverse genetics has provided considerable insights into many biological processes. Reverse genetic approaches begin with a set of genes with known sequences that are of particular interest such as disease-related genes (Ahringer, 1997; Barr *et al.*, 2001; Derry *et al.*, 2001). Genes are inactivated by target-selected approaches such as creation of deletion mutants using chemical mutagens or UV light (Gengyo-Ando and Mitani, 2000; Jansen *et al.*, 1997; Liu *et al.*, 1999),

transposon (Tc1) insertion (Rushforth *et al.*, 1993), and RNAi (Fire *et al.*, 1998). The availability of the complete and well-annotated genome sequence and RNAi libraries currently covering over 94% of the predicted genes in *C. elegans* (Ahringer, 2006) allow investigation of nearly any gene in *C. elegans*. For example, Derry *et al.* (2001) used sequence analysis to identify *cep-1*, the *C. elegans* ortholog of the mammalian p53 tumor suppressor gene (Rubin *et al.*, 2000). By generating a *cep-1* deletion mutant and using RNAi for functional assays, they identified and characterized the roles of CEP-1 in regulating apoptosis, stress response in somatic cells, and chromosome segregation in the germ line. This work laid the foundation for subsequent genetic screens that have added novel insights into how p53 mediates these conserved processes (Fuhrman *et al.*, 2009; Gao *et al.*, 2008; Schumacher *et al.*, 2005; Sandoel *et al.*, 2010).

As two independent approaches for deciphering gene function, forward and reverse genetics can often complement each other. A notable example comes from studies on the *C. elegans* orthologs of two mammalian polycystin proteins, PKD1 and PKD2, which are defective in human autosomal dominant polycystic kidney disease (ADPKD), one of the most common monogenic human disorders, affecting 1 in 400–1000 individuals (Igarashi and Somlo, 2002). In a forward genetic screen for male mutants defective in the ability to locate the hermaphrodite vulva, Barr and Sternberg (1999) isolated a mutation in *lov-1*. Cloning the *lov-1* gene revealed it to be the ortholog of the human disease gene PKD1, which encodes a large transmembrane receptor-like protein (Harris and Torres, 2009; Xiao and Quarles, 2010). Barr and Sternberg found this gene to be exclusively expressed in male-specific sensory neurons. Loss of *lov-1* function displayed no phenotype in hermaphrodites, which is likely the reason this gene was not previously identified, as male-specific phenotypes, especially behavioral, are not often examined. With this knowledge of *lov-1* in hand, they employed a reverse genetic strategy and generated a deletion mutant of *pkd-2*, the worm ortholog of a second PKD disease gene, which encodes a transient receptor potential channel (Clapham, 2003). Strikingly, they found a similar mating defect in males (Barr *et al.*, 2001). Both *lov-1* and *pkd-2* localize to the ciliated endings of male-specific sensory neurons, the site of sensory mechanotransduction, but they are not required for ciliogenesis. This was important, as this study was the first to indicate that these genes may function in sensory transduction in cilia, a location where mammalian PKD1 and PKD2 genes were later found to reside and mediate mechanosensation (Nauli *et al.*, 2003; Pazour *et al.*, 2002; Yoder *et al.*, 2002). Dysfunction of these genes may cause ADPKD due to the inability of renal epithelial cells to sense fluid flow, which might alter various cell functions, including gene expression, growth, differentiation, and apoptosis (Nauli *et al.*, 2003).

## 2. Direct Simple Screen

Dissecting genetic pathways involved in a biological process usually starts from identifying functionally nonredundant components of those pathways. To search for these key players, direct simple screens are often used. In this type of screen, mutants

with desired phenotypes are isolated by direct inspection of descendants of mutagenized or RNAi-treated worms. For example, the very first direct simple screen using EMS as a mutagen in *C. elegans* was performed by Sydney Brenner with particular interests in mutants defective in coordinated movement (Brenner, 1974). He identified mutations in 77 genes affecting movement. Notably, one of these genes, *unc-6*, was later shown to be the ortholog of the vertebrate netrin gene, encoding an important extracellular cue directing axon outgrowth and broadly conserved across the animal kingdom (Harris *et al.*, 1996; Hedgecock *et al.*, 1990; Ishii *et al.*, 1992; Kennedy *et al.*, 1994; Lauderdale *et al.*, 1997; Mitchell *et al.*, 1996; Serafini *et al.*, 1994).

### 3. Forward Screens with the Aid of the Green Fluorescence Protein (GFP)

In contrast to phenotypes observed in direct simple screens, many phenotypical changes, particularly cellular biological ones, are invisible at the behavioral and light microscope levels. For some phenotypes, this limitation can be overcome with fluorescent proteins. Fluorescent markers, in particular, green fluorescence protein (GFP) from the jellyfish *Aequoria victoria* (Chalfie *et al.*, 1994), have facilitated screens in many biological processes, including axon guidance (Zallen *et al.*, 1998), vesicle transportation (Grant and Hirsh, 1999; Grant *et al.*, 2001; Sato *et al.*, 2008), and synapse formation (Liao *et al.*, 2004; Shen and Bargmann, 2003; Zhen *et al.*, 2000). In these screens, GFP is utilized as a visual indicator of phenotypic alterations for identification of mutants, as it allows selective visualization of normally invisible proteins, subcellular structures, specific cells or tissues, and even gene transcription status by placing the GFP open reading frame downstream of genes' *cis*-regulatory regions. In these types of screens, worms that are engineered to transgenically express GFP are mutagenized and examined for changes in GFP expression levels or patterns. For instance, to investigate the mechanisms underlying left–right functional asymmetry of chemoreceptor gene expression between two morphologically symmetrical neurons, ASE left (ASEL) and ASE right (ASER), Chang *et al.* (2003) performed a screen on transgenic worms with GFP expression in either ASEL or ASER under the control of the cell-specific *cis*-regulatory regions. They identified mutations that lost left–right functional asymmetry by isolating mutants that symmetrically expressed GFP in both cells or neither. From this screen, they uncovered several microRNAs and transcription factors that formed a complex regulatory cascade directing left–right asymmetrical chemoreceptor gene expression, thus shedding light on chemosensory neuron differentiation (Chang *et al.*, 2004; Johnston and Hobert, 2003).

The manual isolation of mutants using fluorescence markers is often laborious, as it requires visual inspection of a large number of mutagenized worms at the microscopic level. Recently, an automated worm sorter (Complex Object Parametric Analysis and Sorter, COPAS), which is a flow cytometry machine used to sort worms based on their optical sizes, density, changes in color, and fluorescence intensity (Doitsidou *et al.*, 2008), has been developed to facilitate isolation of mutants. For

example, to study genes involved in executing the dopaminergic cell fate, Doitsidou *et al.* (2008) devised a worm-sorter-based screen using a transgenic strain with all dopaminergic neurons exclusively labeled by a cell type-specific GFP reporter. As a failure to execute dopaminergic cell differentiation can result in fewer GFP positive neurons, mutants with reduced GFP fluorescence were sought. To control for the variability in fluorescence intensities among individual worms, they introduced a broadly expressed red fluorescence protein (RFP) reporter into the transgenic strain as an internal reference for GFP/RFP ratiometric measurements. The worm sorter was accordingly set to detect a reduced GFP/RFP ratio. This screening strategy was highly sensitive as it allowed the identification of mutants lacking GFP expression in only one or two of the eight dopaminergic neurons. In comparison with a manual screen performed in parallel, the worm-sorter-based screen displayed a higher efficiency in isolating mutants, as  $\sim 50,000$  individual worms per hour were screened – in contrast to  $\sim 1000$  manually screened with a microscope. The automated screen identified 22 mutants over a few days, whereas the manual screen isolated 10 mutants over a few months.

#### 4. Enhancer Screens

It is estimated that only  $\sim 30\%$  of the approximately 20,000 genes encoded in the *C. elegans* genome show a visible, lethal, or sterile phenotype after loss or reduction in function (Hodgkin and Herman, 1998; Johnsen and Baillie, 1991). The majority of genes are phenotypically silent upon loss of function under laboratory conditions. Sometimes, it is because the function of these genes can be compensated either by homologous genes with high structural and functional similarity or through buffering of regulatory networks via nonhomologous genes acting in related pathways (Hartman *et al.*, 2001; Wagner, 2000). These genes, which are thought to constitute a large part of most genetic pathways, are unlikely to be recovered in direct simple screens. Identifying these genes necessitates the loss of two or more genes simultaneously. One screening strategy to accomplish this task is the use of an enhancer screen, which is usually conducted on a starting strain with a defined phenotypic defect caused by a mutation in a single gene. Any gene whose functional disruption can enhance the defects of the starting perturbation is referred to as an enhancer of the starting mutant.

Two types of genetic outcomes are possible between enhancers and a starting gene: (1) synergistic enhancement in which the combined severity is more than the sum of both single mutant phenotypes; (2) additive enhancement in which the severity of the combined defects equals to the sum of the individual defects. The type of genetic enhancement can be informative for constructing genetic pathways (see also discussion in Part III).

To devise an enhancer screen, genetic nature, particularly dosage effects (loss of function or reduction of function) of starting alleles used for enhancer screens should be taken into careful consideration because the dosage nature of the starting alleles affects the types of enhancer genes that can be identified.

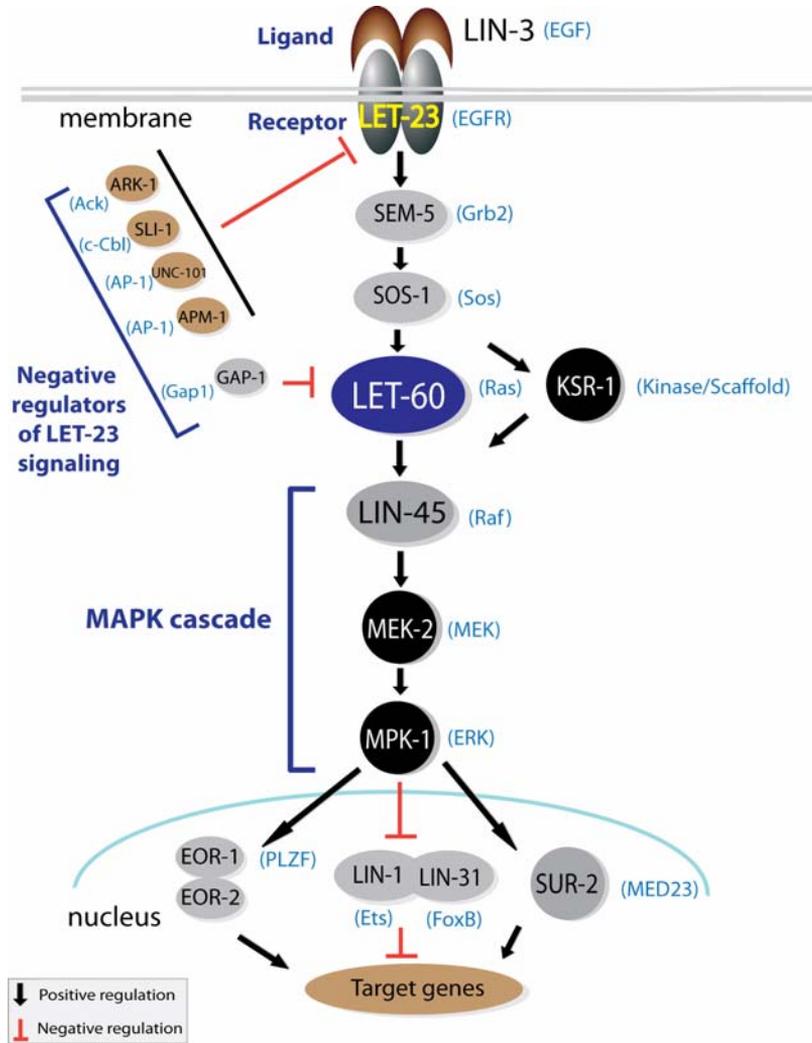
### Enhancer Screens Using Null Alleles

Enhancer screens using null alleles (mutations causing complete loss of function in the corresponding genes) as starting mutations are a widely used and effective way to isolate functionally redundant genes involved in the same biological process. This approach was used to study the negative regulation of a Receptor Tyrosine Kinase – Ras GTPase – Mitogen Activated Protein Kinase signaling cascade (RTK/Ras/MAPK) mediated by LET-23, an epidermal growth factor receptor (EGFR) like-tyrosine kinase, in vulval induction in *C. elegans* (Fig. 1). Activity of LET-23 was known to be dampened by several functionally redundant negative regulators. A mutation in any individual regulator is phenotypically silent with regard to vulval induction but a combination of any two of them displays a hyperinduced-vulval phenotype due to increased activity of LET-23 (Sternberg *et al.*, 1994). To identify new LET-23 negative regulators that might be masked by this redundancy, Hopper *et al.* (2000) conducted an enhancer screen for the hyperinduced-vulval phenotype in the background of a null allele of *sl-1* (an ortholog of the Cbl family of ubiquitin ligases), a known negative regulator of LET-23, showing no phenotype on its own (Hopper *et al.*, 2000). This screen identified a novel negative regulator, *ark-1*, which encodes an ortholog of the Ack-related nonreceptor tyrosine kinase. This gene was later found to be a target of LIN-12/Notch lateral signaling and to mediate the interaction between LET-23 signaling and LIN-12 pathway during their cooperative regulation of vulval cell-fate specification (Yoo *et al.*, 2004).

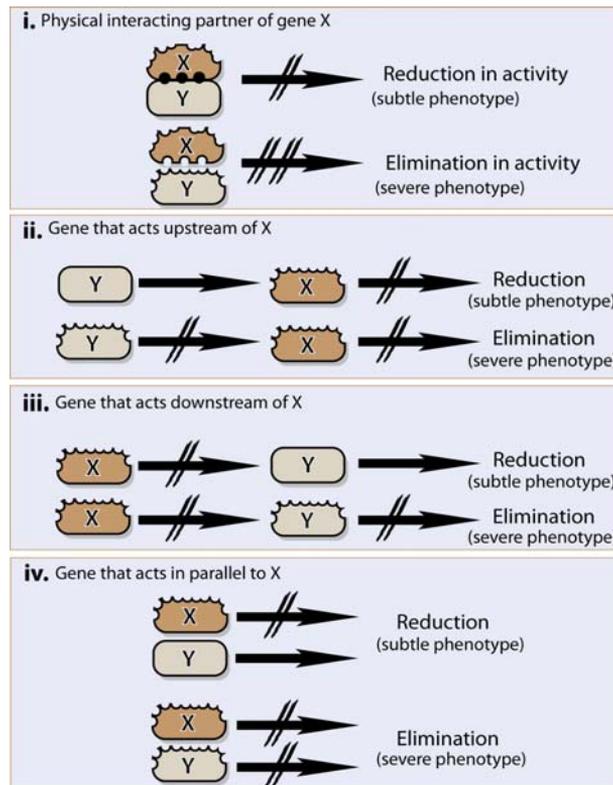
Although an enhancer screen using a null allele is powerful in identifying functionally redundant genes, it has limitations. As a null allele does not produce a protein to regulate downstream components or be influenced by gene products acting upstream, the linearity of the signaling relay is interrupted and the activity of the corresponding pathway is lost. Additional mutations in genes that act either upstream or downstream would not further enhance the initial phenotype of the null allele. Thus, an enhancer screen using a null allele is ineffective in identifying components that act upstream or downstream of the pathway where the starting null allele resides.

### Enhancer Screens Using Hypomorphic Alleles

A hypomorphic mutation causes partial loss of gene function, which leads to a reduction in the activity of the encoded gene product in the signaling pathway in which this gene is involved. This reduction can be enhanced in various situations (Fig. 2): (1) loss or reduction in function of a gene that encodes a physical interacting partner of this hypomorphic allele; (2) loss or reduction in function of another gene acting in the same or the parallel pathways that functionally compensate for each other. Thus, an enhancer screen using a hypomorphic mutation can identify a broad range of interactors that function in either the same physical complex or the same pathway (upstream or downstream), as well as genes that act in redundant pathways. For example, in an enhancer screen using a hypomorphic allele of *lin-45*, a critical



**Fig. 1** The RTK/Ras/MAPK signaling pathway in vulval induction in *C. elegans*. The ligand, its receptor, and core signal transducers are indicated. LIN-3 encodes the worm epidermal growth factor (EGF) ligand and is secreted from the gonadal anchor cell. It binds to its receptor LET-23 on the vulval precursor cells which then dimerizes and undergoes autophosphorylation (Aroian *et al.*, 1990; Hill and Sternberg, 1992). SEM-5, an adaptor protein, binds phosphorylated LET-23 (Clark *et al.*, 1992) and recruits SOS-1 (Chang *et al.*, 2000), a guanine nucleotide exchange factor (GEF) that activates LET-60, a Ras GTPase. KSR-1, a putative scaffold protein, is required for robust activation of LET-60 downstream signaling (Kornfeld *et al.*, 1995b; Sundaram and Han, 1995). LET-60 activates a mitogen-activated protein kinase (MAPK) cascade, including LIN-45 (Raf), MEK-2 (MAPK kinase), and MPK-1 (MAPK) (Chong *et al.*, 2003; Han *et al.*, 1993; Kornfeld *et al.*, 1995a; Lackner *et al.*, 1994; Sternberg *et al.*, 1993; Wu *et al.*, 1995; Wu and Han, 1994). Activated MPK-1 inhibits LIN-1, a transcription factor forming a complex with LIN-31. Upon LIN-1 inhibition, LIN-31 is released to promote the acquisition of vulval cell fates (Tan *et al.*, 1998). MPK-1 also indirectly activates SUR-2, a mediator protein that positively regulates vulval cell fate (Singh and Han,



**Fig. 2** Possible genetic scenarios where the phenotype of a hypomorphic allele can be enhanced. A hypomorphic mutation (X) causes partial loss of gene function, which only slightly reduces the activity of a corresponding signaling pathway. This reduction can be dramatically enhanced by loss or reduction in the activity of a gene (Y) that encodes a physical interacting partner (i), acts upstream (ii), downstream (iii) of the same pathway, or functions in a parallel pathway that functionally compensate for each other (iv). (For color version of this figure, the reader is referred to the web version of this book.)

1995). EOR-1, a putative transcription factor related to the human oncogene PLZF (Hoepfner *et al.*, 2004), and EOR-2, a novel protein (Hoepfner *et al.*, 2004), act downstream or in parallel to MPK-1 and function redundantly with LIN-1 to regulate transcription of target genes (Howell *et al.*, 2010; Rocheleau *et al.*, 2002). LET-23 signaling is also regulated by several negative regulators, include ARK-1 (Ack) (Hopper *et al.*, 2000), SLI-1 (c-Cbl), which targets activated LET-23 for internalization and degradation (Rubin *et al.*, 2005; Swaminathan and Tsygankov, 2006), UNC-101 and APM-1, which encode medium chains of the AP-1 adaptin and promote LET-23 endocytotic recycling (Lee *et al.*, 1994; Shim *et al.*, 2000), and GAP-1, which stimulates LET-60 GTP hydrolysis (Hajnal *et al.*, 1997). More details about this pathway can be found elsewhere (Sundaram, 2006). The mammalian homologs of these genes are indicated in parentheses. (For color version of this figure, the reader is referred to the web version of this book.)

RTK-Ras-MAPK component encoding a Raf protein, Rocheleau *et al.* (2002) identified novel alleles of known components of this pathway that function either upstream or downstream of *lin-45*, including *sem-5*, *sos-1*, *lin-1*, and *ksr-1*, and alleles in two new components, *eor-1* and *eor-2*, likely acting downstream of or in parallel to *mpk-1* (Fig. 1). The spectrum of genes identified highlights the efficiency of using hypomorphic alleles in enhancer screens.

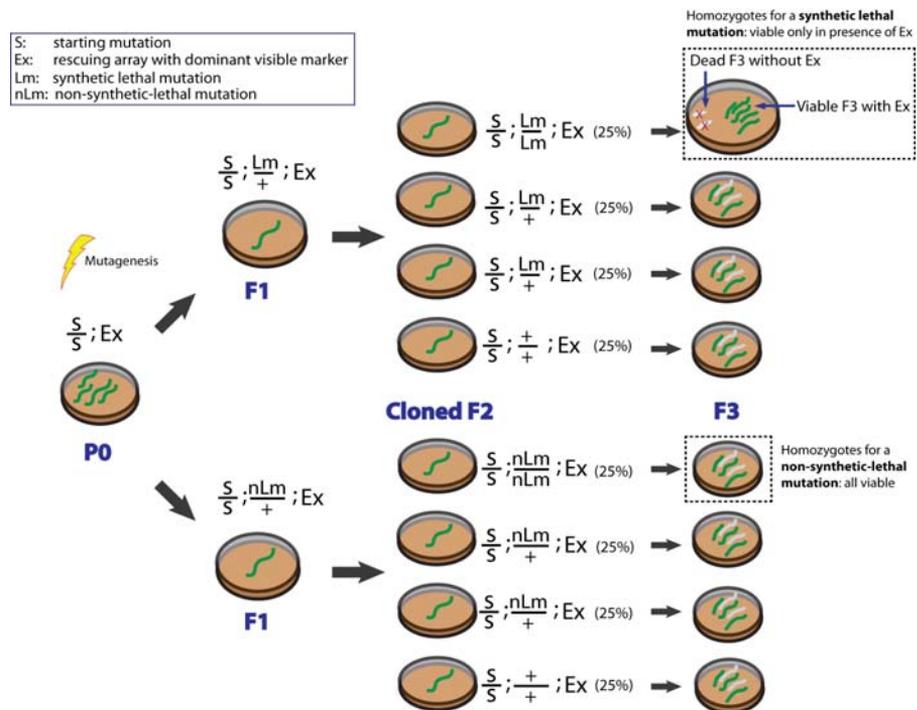
Another example showing the usage of hypomorphic alleles in identifying genes with redundant functions comes from a screen carried out by Schwabiuk *et al.* (2009). They discovered a novel function for a gene, *sdn-1*, in regulating migration of the distal tip cells (DTC) that lead gonad extension. *sdn-1*, which encodes an ortholog of a type I transmembrane proteoglycan syndecan-2, was previously implicated in axon guidance (Rhiner *et al.*, 2005) and in epidermal enclosure (Hudson *et al.*, 2006). Its new functional role in DTC migration was found in a screen for mutations that enhanced the DTC migration defects caused by a hypomorphic allele of *unc-5*, which encodes a receptor for the axon guidance cue, netrin. *sdn-1* would not otherwise have been identified in simple screens using wild-type worms or enhancer screens using a null allele of *unc-5*, because the ensuing functional characterization revealed that all *sdn-1* alleles (null and hypomorphic) were phenotypically silent in DTC migration on their own and importantly no enhancement was observed in the double mutants between an *unc-5* null allele and any of *sdn-1* alleles, indicating that *sdn-1* function is linked to *unc-5* activity.

For some enhancer screens, temperature-sensitive alleles (“hypomorphic” equivalents) can also be used. These alleles often display incomplete phenotypic penetrance at intermediate temperature between restrictive and permissive temperatures. Additional mutations that increase original penetrance of phenotypes at intermediate temperature are isolated and cloned. This temperature strategy has been used in a number of enhancer screens, such as a screen to identify genes involved in the neddylation process (Dorfman *et al.*, 2009), as well as LIN-12/Notch signaling (Qiao *et al.*, 1995).

### Synthetic Lethality Screen Using Extrachromosomal Arrays

In some enhancer screens, the combination of two mutations can cause synthetic lethality: the disruption of a single gene displays no discernable or a very subtle phenotype, whereas simultaneous disruption of two or more genes causes lethality. Synthetic lethality poses a challenge for mutation recovery as it is usually unknown whether a starting mutation has synthetic lethal partners. One way to recover synthetic lethal mutations is to devise an enhancer screen using a starting mutant carrying an extrachromosomal array that rescues potential lethality (Fig. 3). These arrays are transgenes introduced into worms by gonad microinjection (Mello and Fire, 1995). Due to their extrachromosomal nature, these transgenes are stochastically lost during meiosis and mitosis, and are only expressed in the progeny or daughter cells containing them. For enhancer screens, the transgenes carried by the starting mutant are often designed to express two proteins simultaneously: (1) a wild-type copy of a starting-mutation-harboring gene for avoiding synthetic

lethality; (2) a dominant visible marker (e.g., GFP) driven by a ubiquitous promoter to indicate the presence of the transgenes. The progeny of this transgenic strain will be comprised of two populations: the marker-positive worms (*transgene*<sup>+</sup>) and the marker-negative worms (*transgene*<sup>-</sup>), the frequency of each being dependent on the frequency of stochastic loss of the extrachromosomal array. If a synthetic lethal mutation were to be introduced into one of these transgene-containing worms, only progeny containing the transgene would be viable, as the marker-negative population would be all dead due to the absence of an extrachromosomal array expressing the wild-type protein. Conversely, both *transgene*<sup>+</sup> and *transgene*<sup>-</sup> progeny would be present and viable with nonlethal enhancers.



**Fig. 3** Synthetic lethality screen using extrachromosomal arrays. The starting strain (P0) with an initial mutation (S) and a rescuing array containing a dominant visible marker (Ex) is mutagenized. The progeny (F1) carrying the visible marker (green) are cloned onto individual plates. As an F1 worm is heterozygous for a recessive mutation, only 25% of the progeny (F2) will be homozygous. Thus, several F2 progeny carrying the marker from each F1 plate are then cloned onto individual plates (here, we show four F2 worms per F1). The progeny (F3) of each cloned F2 are inspected for the presence of the marker. For an F2 that is homozygous for a synthetic lethal mutation (Lm), all viable F3 progeny will be marker-positive (green) because the progeny lacking the marker (white) will be dead (dotted box). In the case of F2 worms derived from an F1 worm carrying a nonlethal mutation (nLm), the progeny F3 of each F2 worm will be both marker-positive (green) and marker-negative (white). This screening process is also known as an F2 clonal screen. (See color plate.)

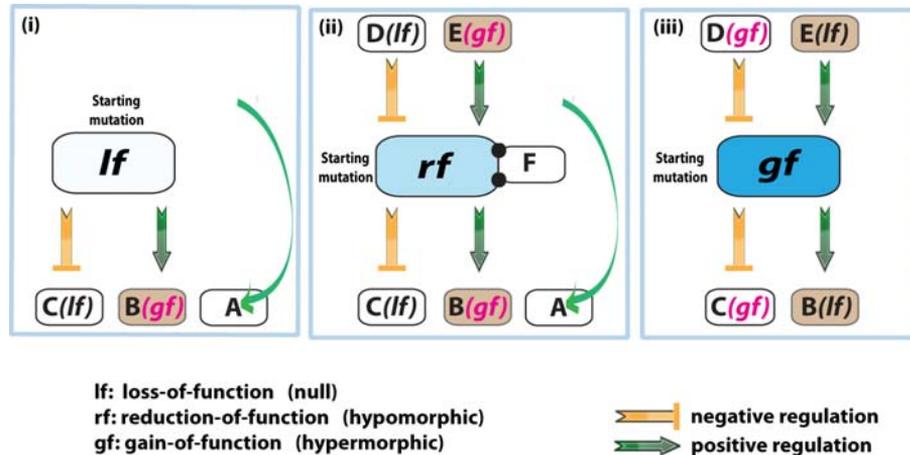
Using this screening strategy, Fay *et al.* (2002) discovered that the gene *fzr-1* functions redundantly with *lin-35*, the *C. elegans* ortholog of the Rb (Retinoblastoma protein) tumor suppressor gene, in controlling cell proliferation. Unlike Rb knockout mutants in flies and mice (Clarke *et al.*, 1992; Du and Dyson, 1999; Jacks *et al.*, 1992), worms with strong loss of function of *lin-35/Rb* are viable, and display relatively subtle defects in development (Fay *et al.*, 2002; Lu and Horvitz, 1998), suggesting the existence of genes acting redundantly with *lin-35*. To identify these genes, they performed a synthetic lethality screen on *lin-35* mutants carrying an extrachromosomal array expressing the wild-type LIN-35 protein and a ubiquitously expressed GFP marker for tracking worms with the array. After mutagenesis, F3 worms derived from each of 10,000 single cloned F2 (four F2 per F1) were examined for the presence or absence of the GFP marker (see Fig. 3 for details). They recovered seven mutations defining seven distinct loci that displayed synthetic lethality. The gene characterized in this study was *fzr-1* encoding a protein orthologous to a regulatory subunit of the anaphase-promoting complex required for anaphase initiation and exit from mitosis (Dawson *et al.*, 1995; Schwab *et al.*, 1997; Sigrist *et al.*, 1995). In addition to synthetic lethality, the double mutants of *fzr-1* and *lin-35* also showed extensive tissue hyperproliferation affecting a wide range of cell types, indicating new functional roles of *lin-35*. As hyperproliferation caused by mutations in mammalian Rb genes is an important genetic event during multi-step carcinogenesis, this study provided support that uncontrolled proliferation in *C. elegans* follows the same genetic pattern of oncogenesis in mammals and revealed a possible connection of this process with the anaphase-promoting complex.

## 5. Suppressor Screens

Suppressor screens are a powerful way to identify interacting genes that regulate biological processes. The “suppression” design enables this type of screen to bypass synthetic lethality, which can be a limitation of an enhancer screen. A suppressor refers to a gene whose dysfunction (loss, reduction, or gain of function) can suppress the well-defined phenotypic defect of a mutation in another gene (starting allele). Prior to screening, it is of great importance to decide the appropriate starting alleles for suppression. Alleles identified in direct simple screens often exhibit highly penetrant phenotypic defects, which may limit their utility in enhancer screens but makes them potentially ideal starting alleles for suppressor screens. Second, the genetic nature (null, hypomorphic, or hypermorphic, i.e., gain of function) of a starting allele should be taken into careful consideration as it determines the types of mutants that can be recovered from the screen.

### Suppression of a Null Allele

Suppression of a null allele in a gene of interest allows isolation of mutations in genes that bypass the need for the original gene. In addition, the screen can also identify gain-of-function (*gf*) mutations in downstream genes that are positively



**Fig. 4** The choice of starting mutations affects the types of mutations recovered from suppressor screens. (i) A suppressor screen using a loss-of-function (*lf*) mutation as a starting mutation allows isolation of mutations (A) in genes that bypass the gene of interest, gain-of-function (*gf*) mutations (B) in downstream genes that are positively regulated by the gene of interest, and loss-of-function (*lf*) mutations (C) in downstream genes that are negatively regulated by the pathway. (ii) In addition to the similar spectrum of mutations recovered from suppression of a null mutation, suppression of a reduction-of-function (*rf*) mutation can identify *lf* mutations (D) in upstream negative regulators and *gf* mutations (E) in upstream positive regulators, and mutations (F) in direct physical interactors. (iii) Suppression of a *gf* mutation can obtain a wide range of mutations with attributes opposite to those identified in suppression of a *lf* mutation. (For color version of this figure, the reader is referred to the web version of this book.)

regulated by the gene of interest as well as loss-of-function (*lf*) mutations in downstream genes that are negatively regulating the pathway (Fig. 4i).

Suppressor screens using null mutations are in general not effective in recovering mutations in genes acting upstream of a pathway or directly interacting with the starting genes, because null alleles produce no protein products for upstream genes to regulate and for direct interactors to modulate. A way to identify genes that act upstream and direct interacting genes is to use hypomorphic alleles as the basis of a suppressor screen.

### Suppression of a Hypomorphic Allele

Suppression of hypomorphic alleles, that is, reduction-of-function (*rf*) mutations, not only yields the range of mutations similar to suppression of null mutations, but also allows recovery of mutations in upstream genes, such as *lf* mutations in upstream negative regulators and *gf* mutations in upstream positive regulators, and mutations in genes directly interacting with the gene of interest (Fig. 4ii). For example, UNC-4, a homeodomain protein, regulates synaptic connectivity of VA motor neurons, which mediates backward movement in *C. elegans*. *unc-4* mutants

are unable to crawl backwards. To identify gene products that directly interact with UNC-4, Miller *et al.* (1993) screened for mutants that suppressed a hypomorphic temperature sensitive allele of *unc-4* and successfully identified a *gf* mutation in a gene, *unc-37*, which was found to encode a Groucho-like protein forming a complex with UNC-4 to regulate transcription (Winnier *et al.*, 1999). Importantly, this mutation in *unc-37* was not able to suppress *unc-4* null alleles, suggesting that this suppression was dependent on UNC-4 activity.

### Suppression of a Hypermorphic Allele

Suppression of hypermorphic alleles, that is, *gf* mutations, yields a spectrum of alleles in genes with attributes opposite to those identified in suppression of *lf* mutations (Fig. 4iii). This type of suppression is a powerful way to identify components of a biological pathway. A good example comes from the highly conserved RTK/Ras/MAPK pathway, which is involved in vulval induction in *C. elegans* (see Fig. 1). The worm Ras gene, *let-60*, is a crucial component of this pathway. Gain-of-function mutations in *let-60* cause a multiple-vulva (Muv) phenotype. A group of key components acting downstream of *let-60*, including *lin-45* (the worm Raf gene) (Hsu *et al.*, 2002), *mpk-1* (the worm ERK gene) (Wu and Han, 1994), *mek-2* (the worm MEK gene) (Wu *et al.*, 1995), *ksr-1* (a *C. elegans* kinase suppressor of Ras) (Sundaram and Han, 1995), and *sur-2* (Singh and Han, 1995), were all identified in screens for suppressors of the Muv phenotype caused by a *gf* allele of *let-60*.

### Suppression of Engineered Gain-of-Function (*gf*) Alleles

*gf* mutations used for suppressor screens are not limited to genetically defined alleles. By engineering into worms transgenes (“artificial *gf* alleles”) expressing mutant proteins, one can produce a phenotype that facilitates selection and identification of suppressors. For instance, Zheng *et al.* (2004) devised a suppressor screen hunting for genes interacting with an important class of neurotransmitter receptor, ionotropic glutamate receptors (iGluRs), that mediate most excitatory synaptic signaling between neurons (Zheng *et al.*, 2004). The screen was performed in a transgenic strain engineered to express a non-N-methyl-d-aspartate (NMDA) type ionotropic glutamate receptor (GLR-1) subunit containing a dominant mutation. Expression of this transgene under the control of the *glr-1* promoter resulted in a hyper-reversal phenotype, that is, transgenic animals show a higher frequency of reversing direction during movement than wild-type worms. By searching for mutations suppressing this phenotype, Zheng and colleagues identified a gene encoding a type I transmembrane protein, SOL-1, that can bind to GLR-1 and participate in the gating of non-NMDA iGluRs.

### Suppressors of Engineered Pathological Processes

Greater than 40% of human-disease-related genes have clear *C. elegans* orthologs (Culetto and Sattelle, 2000). *C. elegans* has been used to model human diseases,

including human ADPKD (Barr *et al.*, 2001; Barr and Sternberg, 1999), muscular dystrophy (Kim *et al.*, 2004), cancer (Bergamaschi *et al.*, 2003; Polanowska *et al.*, 2004), diabetes and obesity (Pierce *et al.*, 2001), and neurodegenerative diseases (Lakso *et al.*, 2003). The applications of *C. elegans* for various disease models have been extensively reviewed (Dimitriadi and Hart, 2010; Kaletta and Hengartner, 2006; Kirienko *et al.*, 2010). With the unique genetic advantages of *C. elegans*, large-scale screens that are impractical in vertebrates can be readily performed in *C. elegans* to identify highly conserved genes that may modulate human diseases. To identify those genes, suppressor screens are often performed on a transgenic strain generated to resemble a pathological process of interest. For example, one hallmark of some notable neurodegenerative diseases is the abnormal aggregation of proteins, such as wild-type or mutated tau protein that normally functions to stabilize microtubules and promote their polymerization. The aggregation of tau is seen in a group of neurodegenerative diseases, including Alzheimer's disease and frontotemporal dementia with parkinsonism chromosome 17 type (FTDP-17T) (Lee *et al.*, 2001). To identify genes participating in tau neurotoxicity, Guthrie *et al.* (2009) carried out a forward genetic screen for suppressors of the Unc (uncoordinated movement) phenotype caused by accumulation of exogenous mutated human tau in a transgenic strain that was engineered to express this tau protein in all neurons. Using this transgenic worm as a model of human tauopathy disorders, they revealed that loss of function in a gene, *sut-2*, which encodes a highly conserved subtype of CCCH zinc finger protein, was able to suppress tau neurotoxicity. The identification of this gene suggested a novel neuroprotective strategy to interrupt tau pathogenesis (Guthrie *et al.*, 2009).

### Some Considerations Regarding Nature of Suppressors

There are two specialized types of suppression that can arise when performing a suppressor screen, which may not provide insight into the biological processes of interest. These are important to be aware of when interpreting results of suppressor screens. One is informational suppression caused by mutations in genes involved in the general machinery of transcription, RNA processing, and protein translation. This suppression is allele-specific, gene-nonspecific. A large number of EMS-induced null alleles are nonsense point mutations causing early stop codons. Suppressor screens using these nonsense alleles can produce tRNA mutations that recognize stop codons as sense codons so that the starting allele can be translated to the protein with biological activity. As an informational suppressor is allele-specific, it may not suppress the phenotype of other alleles of the gene of interest, which provides a good way to test if a suppressor mutant is an informational mutation. Similarly, mutations of components in the nonsense decay system that is responsible for degrading premature mRNA can also suppress the phenotype of the starting mutation. Informational mutations are valuable for studies on regulation of transcription, RNA processing, and translation; however, they do not provide insights into the genetic networks controlling the biological processes likely to be of interest in the screens.

Another type of suppression is referred to as intragenic suppression, where a suppressor mutation in the same gene that harbors the starting mutation reverses its defects. For example, the suppressor may introduce a functional favorable mutation into the protein, which may compensate for the reduction of activity caused by the original deleterious mutation. Sometimes, intragenic suppressor mutations can affect splicing, which causes a skip of the original mutation and produces a protein with biological activity. More details on this topic can be found in a review by Hodgkin (2005).

## 6. Selection Screens

Depending on the design of the genetic screen, the isolation of mutants can be laborious. Many screens require careful examination of nearly every single worm for the presence or absence of a desired phenotype. For example, in the previously mentioned screen for suppressors of an *unc-4* allele, 500,000 progeny of mutagenized worms were individually tapped on the head with a platinum pick to test if the worms were able to move backward (Miller *et al.*, 1993). Selection screens are designed to rapidly facilitate identification of mutants with a specific phenotype by eliminating animals that do not carry a desired mutation. A drug screen is one type of selection screen in which mutants resistant to a particular drug will be readily selected because all other worms are either killed or display a specific phenotype when lacking resistance. For example, acetylcholine, a neurotransmitter, is released from synaptic vesicles at neuromuscular junctions, where it induces muscle contraction. Acetylcholine is normally degraded by the enzyme acetylcholinesterase. Pesticides, such as Aldicarb, block the activity of this enzyme and cause accumulation of acetylcholine, which ultimately kills animals, including worms, because of excessive excitation of muscles. A screen searching for mutants that are resistant to Aldicarb (e.g., viable in its presence) identified *unc-17*, which was later cloned and found to encode a broadly conserved acetylcholine transporter involved in uptake of acetylcholine into synaptic vesicles (Alfonso *et al.*, 1993; Brenner, 1974).

Inducible transgenes can also be used for selection screens, where a transgene engineered into a strain induces a particular phenotype that when suppressed enables easy selection of corresponding mutations. For example, GOA-1, a *C. elegans*  $\alpha$ -subunit of the major heterotrimeric G protein in the nervous system, regulates many behaviors, including locomotion and egg laying (Hajdu-Cronin *et al.*, 1999). To identify genes that interact with G<sub>o</sub> signaling, Hajdu-Cronin *et al.* (1999) generated a transgenic strain overexpressing a constitutively active GOA-1 mutant protein under the control of a heat-shock promoter. Upon heat shock, this strain displays a severe phenotype, paralysis. By screening for mutants that restored locomotion, they identified two regulators of G<sub>o</sub> signaling, *dgk-1* (first identified and named as *sag-1* in the study), which encodes a diacylglycerol kinase (Miller *et al.*, 1999; Nurrish *et al.*, 1999), and *eat-16*, which encodes a protein orthologous to the mammalian regulators of G protein signaling (RGS) 7 and RGS9 (Hajdu-Cronin *et al.*, 1999).

In addition, selection screens can be conducted using temperature sensitive (*ts*) lethal mutants. Growing such mutants at restrictive temperature usually induces embryonic or larval lethality. In screens for suppressors of lethality of *ts* alleles, mutants that restore the viability of embryos or larvae can be easily identified. For instance, PAR (partitioning defective) proteins, first identified in forward genetic screens in *C. elegans*, are highly conserved regulators of cell polarity and asymmetric cell division (Kemphues *et al.*, 1988). To gain insight into the precise mechanisms by which PAR-1, a Ser/Thr kinase, regulates embryonic asymmetric cell division, Spilker *et al.* (2009) performed a genome-wide RNAi screen on a temperature sensitive *par-1* allele and identified several genes that when their activity is reduced specifically suppress the embryonic lethality of *par-1*. One of the identified suppressors was *mpk-1*, which encodes a mitogen-activated protein (MAP) kinase. Reduced activity of *mpk-1* restored the asymmetric distribution of cell-fate specification markers in *par-1* mutants. In addition, disrupting the function of other components of the MAPK signaling pathway also suppressed *par-1* embryonic lethality. These results revealed that MAP kinase signaling is involved in antagonizing PAR-1 activity during early *C. elegans* embryonic polarization.

## 7. Sequential Screens

The specific phenotypes of a biological process of interest may not always be suitable for large-scale forward genetic screens. One common reason is the difficulty in observing the phenotype. To make genetic screens applicable for such a biological process, it is often possible to score a more readily detected phenotype, such as lethality or uncoordinated movement (primary screen), that allows isolation of a broader scope of mutations including those specific for the process of interest. The specific mutations are then identified through a secondary screen. A good example using a sequential strategy was a screen aimed at identifying genes involved in regulation of presynaptic terminal formation. This was accomplished by seeking suppressors of a mutation in the RING finger/E3 ubiquitin ligase gene *rpm-1*, a key regulator of synapse formation (Nakata *et al.*, 2005). Mutations in *rpm-1* result in a disorganized presynaptic structure but they cause no defects in locomotion. Although this phenotype can be observed with fluorescent synaptic markers, its subcellular microscopic-level nature limits the screening scale. To facilitate the identification of genes interacting with *rpm-1*, they devised a sequential screen for suppressors of an *rpm-1* mutation. In the primary screen, instead of screening solely in the *rpm-1* mutant background, they screened for suppressors of a set of easily scored phenotypes, severe defects in locomotion, and reduction in body size, caused by introducing (with *rpm-1*) a mutation in a synaptogenesis gene, *syd-1*. The SYD-1 protein regulates the distribution of presynaptic components and when mutated leads to mild defects in locomotion and egg-laying (Egl) behavior. Specific suppressors of *rpm-1* could be identified by restoration of locomotion and body size but not the Egl phenotype of the *syd-1* mutation. The alternative easily scored phenotypes used for suppression in the primary screen allowed a large

screening scale. As suppressors isolated in the primary screen may not all be specific to loss of *rpm-1* activity, a secondary screen at the microscope level was then conducted to identify *rpm-1*-specific suppressors that restored synaptic morphology using a synapse fluorescent marker. The highlight from this screen was the identification of three MAP kinases, *dlk-1* (MAPKKK), *mkk-4* (MAPKK), and *pmk-3* (p38-like MAPK), which were found to form a previously uncharacterized p38 MAP kinase cascade that is negatively regulated by RPM-1 during synapse formation. The success of this screen speaks to the specificity and utility of a sequential screening strategy when designed appropriately.

## 8. Evaluation and Limitations of Forward Genetic Screens

Regardless of various screening designs, a high-quality screen should be able to identify many or most of the nonredundant components of a biological pathway of interest. A key question that then arises is how to determine the degree of saturation (all genes that can be mutated to display a specific phenotype) that a screen reaches. Empirically, if a screen is saturated, (1) mutations in the same gene, particularly in small-size genes that usually are less frequently hit, will be repeatedly isolated. This will be reflected by the fact that multiple alleles fail to complement each other in a complementation test, which is a genetic experiment to determine if two alleles reside in the same gene by comparing the phenotype of transheterozygotes of these two alleles with that of homozygotes for each allele. If the phenotypes of transheterozygotes and homozygotes are the same, it indicates that the two alleles fail to complement each other, suggesting that the two alleles likely correspond to the same gene. Conversely, if a screen is not saturated, it is common to see that each mutation defines a distinct locus. For example, in the *lin-35/Rb* synthetic lethality screen described in Section II.A.4, Fay and his colleagues recovered seven mutations defining seven distinct loci, indicating the screen was not saturated. (2) Unusual hypomorphic alleles of lethal genes will be identified that normally are less likely to be recovered than null alleles because these alleles require changes in specific amino acids. Besides qualitative judgment, the degree of saturation can also be analyzed in a quantitative manner. A statistic method using Bayesian and maximum-likelihood calculation may be used to estimate the number of alleles that remain to be found (Pollock and Larkin, 2004).

Many components of genetic pathways have been identified through mutagenesis-based forward genetic screens. However, as an experimental approach, forward genetic screens have intrinsic weaknesses in identifying some pathway components. Even though a forward genetic screen can be performed at a large scale, many genes may still be missed or rarely hit for several reasons: (1) small-size genes may be missed because they are too small to be effective targets for mutagenesis; (2) genes with pleiotropic functions might not be identified because they preferentially give a phenotype that masks their other functions (Jorgensen and Mango, 2002); (3) genes that when mutated confer early lethality often prevent identification of their later functions; (4) functional redundant genes that ensure robustness and plasticity to

biological processes often have no observable phenotypes when individually mutated (Wagner, 2000). Although an enhancer screen can overcome some redundancy, this approach does not always exhaust its multiple layers. RNAi-mediated forward genetic screens can also bypass substantial redundancy and lethality by causing partial loss of gene function; however, RNAi has its own limitations. RNAi phenotypes are often variable in penetrance and RNAi is ineffective in neurons (Tavernarakis *et al.*, 2000; Tewari *et al.*, 2004). To find genes that are not easily identified through forward genetic screens, functional genomic and systems-level approaches that complement conventional genetic screens can be used.

## B. Functional Genomic Approaches to Identify Components of Pathways

The completion of sequencing *C. elegans*, *D. Melanogaster*, and *H. sapiens* genomes along with rapidly evolving high-throughput techniques have changed the methodological ways biologists study gene function and dissect genetic pathways. The genome-sequencing project in *C. elegans* revealed a significant number of novel genes with unknown function. Undoubtedly, these genes are involved in a wide variety of biological functions. To decipher their function, traditional forward genetic screening still remains useful but now functional genomic studies can also be utilized. Functional genomics uses high-throughput approaches, such as genome-wide RNAi, DNA microarray, Serial Analysis of Gene Expression (SAGE), *cis*-regulatory analysis, yeast-two-hybrid/yeast-one-hybrid techniques, and mass spectrometry, to acquire information about genome-wide patterns of gene expression, protein–DNA interactions, and protein–protein interactions. By analyzing this information, biologists can begin to elucidate the organization and regulation of genetic pathways at a global level. Nevertheless, during analysis, experimental validation using conventional single-gene genetic approaches, including genetic perturbation analysis and reverse genetic approaches, is indispensable for confirming results obtained from functional genomic studies. The combination of functional genomic approaches with conventional methods has emerged as an effective way to gain a more complete understanding of the gene networks that guide biological processes (see also reviews by Grant and Wilkinson, 2003; Kim, 2001; Piano *et al.*, 2006).

### 1. Genome-Wide RNAi Screens

RNAi is an endogenous cellular process during which double-stranded RNA (dsRNA) complementary to sequences of target messenger RNAs (mRNA) mediates degradation of these mRNAs, resulting in reduction of expression of corresponding genes (Boutros and Ahringer, 2008). Since RNAi was discovered (Fire *et al.*, 1998), it has rapidly been adopted as an experimental means to silence expression of genes in a range of organisms (Boutros and Ahringer, 2008; Gildorf *et al.*, 2010). In *C. elegans*, RNAi assays can be conveniently carried out by feeding worms with bacteria expressing dsRNA constructs (Ahringer, 2006), soaking them in nematode growth media containing these bacteria (Lehner *et al.*, 2006), or injecting dsRNA

into gonads of young adult hermaphrodites to obtain progeny with mutant phenotypes. Essentially, RNAi can be used, instead of a mutagen, for variously designed forward genetic screens already discussed. RNAi-mediated forward genetic screens have some unique advantages. For instance, the identity of genes whose inactivation causes phenotypes is immediately known, in contrast to the time-consuming cloning of mutation-harboring genes from forward mutagenesis screens. RNAi also has temporal flexibility. It can be applied to animals at different developmental stages to avoid embryonic or larval lethality caused by inactivation of corresponding genes at the early developmental stages. Moreover, RNAi usually results in reduction of function of gene activity rather than complete loss, which allows effective investigation of the roles of essential genes (Kemphues, 2005).

There are currently two RNAi feeding libraries available for *C. elegans* research. One library constructed by the Vidal lab has 11,511 clones containing full-length gene cDNAs that were cloned into a double T7 vector by the Gateway cloning method (Rual *et al.*, 2004). This library is commercially named the *C. elegans* ORF-RNAi Collection V1.1, available through Open Biosystems. The other library was constructed by the Ahringer lab and has 16,757 clones containing the genomic sequences of genes (Fraser *et al.*, 2000; Kamath *et al.*, 2003). This library is commercially available through Geneservice. With the availability of two RNAi libraries, which together target 94% of *C. elegans* genes (Ahringer, 2006), RNAi screens are often carried out at a genome-wide scale and more frequently in an automated high-throughput fashion. For example, to identify genetic interactors with the RTK/Ras/MAPK pathway, Lehner *et al.* (2006) performed an RNAi screen in the background of loss-of-function mutations in the 12 known components of the RTK/Ras/MAPK pathway (Kamath *et al.*, 2003). To perform this screen in a high-throughput manner, the RNAi was delivered in 96-well plates in which mutants were soaked in a liquid containing RNAi feeding bacteria. They screened for RNAi clones that produced synthetic lethality with any of these 12 known components. Notably, 16 genes that had no previously reported roles in RTK/Ras/MAPK signaling were found to genetically interact with two or more components of the pathway. Nine out of these 16 genes were shown to regulate RTK/Ras/MAPK signaling during vulval induction (see Fig. 1), perhaps the best-characterized function of this pathway. This study highlights how high-throughput functional genomic approaches can rapidly identify new components of a specific pathway.

The growing number of large-scale RNAi studies carried out in *C. elegans* have produced a wealth of RNAi-induced phenotypic information. This information is being deposited into online databases, such as Wormbase (Harris *et al.*, 2010; Rogers *et al.*, 2008), RNAiDB (Gunsalus *et al.*, 2004), and PhenoBank (Sonnichsen *et al.*, 2005), to facilitate gene function studies. For example, using these databases, our group narrowed our search for genetic regulators of anchor cell invasion in a sequential RNAi screen (Matus *et al.*, 2010). Anchor cell invasion through basement membranes, which mediates formation of uterine–vulval attachment in *C. elegans*, has been used as a simple *in vivo* model for investigating cell invasion (Sherwood, 2006; Sherwood *et al.*, 2005; Ziel *et al.*, 2009). As a failure of anchor cell invasion

causes a Protruding-vulva (Pvl) phenotype, we first compiled a list of 539 genes whose reduction in activity was reported to result in the Pvl phenotype from a number of whole-genome RNAi screens (Kamath *et al.*, 2003; Rual *et al.*, 2004; Simmer *et al.*, 2003). We then performed a focused RNAi screen on these genes by examining anchor cell invasion using differential interference contrast (DIC) optics. Through these efforts, we identified 99 genes that are required for anchor cell invasion. Most of these genes have not been previously implicated in cell invasion, potentially expanding new targets for cancer therapeutics.

## 2. Gene Expression Profiling Approach

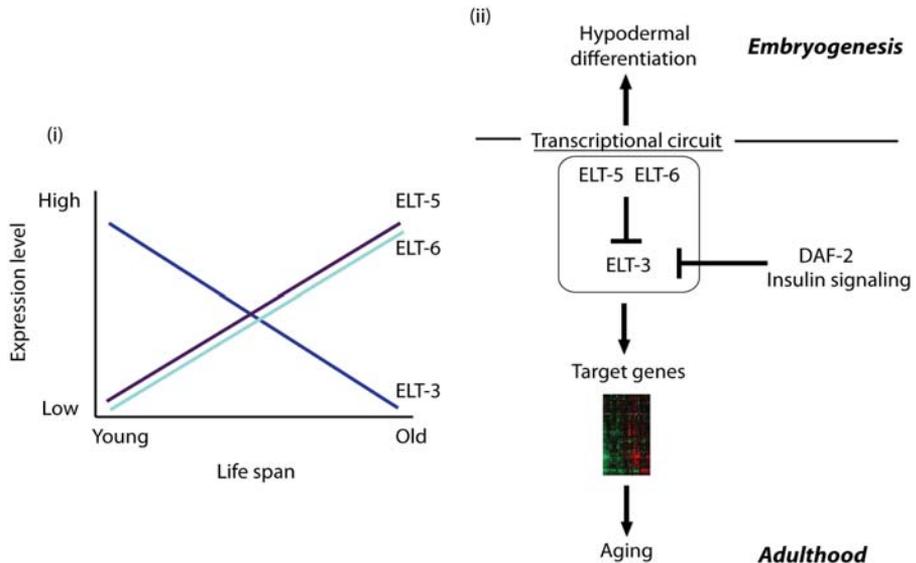
Examination of gene expression at a genome-wide scale within the whole organism, specific tissues, and even single cells has proven to be a valuable approach in identifying pathway components and characterizing gene function (Grant and Wilkinson, 2003). Genes responsible for the same biological process tend to be regulated in a similar manner. By profiling gene expression changes associated with a biological process, it is possible to identify a common set of genes whose expression dynamics and spatiotemporal localization share the same pattern under different conditions and in various mutant backgrounds.

A way to measure expression levels of genes is to quantify transcripts of corresponding genes, which can now be readily achieved through a variety of high-throughput technologies at a genome-wide level, including hybridization-based approaches and sequence-based approaches (Wang *et al.*, 2009). A typical example of hybridization-based approaches is a DNA microarray assay that profiles expression of individual genes at a genomic scale through hybridization of oligonucleotide DNA probes with fluorescently labeled cDNAs of nearly every gene. Unlike hybridization techniques, sequence-based approaches obtain quantitative gene expression data by sequencing gene transcripts. For example, SAGE, a sequencing-based method, quantifies gene expression by counting the number of times a particular transcript is found in a pool of short diagnostic sequence tags isolated from an mRNA sample (Velculescu *et al.*, 1995). Recently, with advances in deep sequencing technologies, RNA-Seq, a new high-throughput and more precise sequencing-based method, allows quantification of all transcripts by directly sequencing fragmented cDNA (30–400 bp) converted from a population of RNA (Wang *et al.*, 2009). In addition to acquiring information about levels of gene expression, a project aimed to profile spatiotemporal patterns of gene expression at a large scale (localizome) has been initiated (Dupuy *et al.*, 2007). In this project, worms have been engineered to express transgenes in which the open reading frame of GFP was placed downstream of the proximal promoters of 1610 predicted genes. The expression of GFP from these promoters has been characterized using a worm sorter that profiles tissue expression at various developmental stages in a high-throughput fashion. The relevant expression data can be found at the web site <http://localizome.dfci.harvard.edu/>. The ultimate goal of this project is the characterization of all of the ~20,000 genes in the *C. elegans* genome.

Excellent examples of how expression analyses have facilitated identification of the genetic networks controlling diverse biological processes include aging (Budovskaya *et al.*, 2008; Murphy *et al.*, 2003), development (Baugh *et al.*, 2009), and innate immunity (Styer *et al.*, 2008). Illustrating its utility in the aging field, Budovskaya *et al.* (2008) recently discovered a development-related transcriptional circuit that guides the aging process. By comparing DNA microarray profiles, they identified a common set of 1254 genes that showed age-dependent expression changes (both upregulated and downregulated) during normal aging. This pattern of gene expression was found to be shared with dauer larvae (developmentally arrested worms whose life spans are ten times longer than normal worms), and longevity mutants displaying either extended or shortened life spans. For example, genes that show increased expression with age tend to have increased expression in dauer larvae and long-lifespan mutants, but show decreased expression in short-lifespan mutants. To search for transcription factors that regulate these age-dependent expression changes, they analyzed the upstream regulatory regions of these 1254 genes and identified a common consensus motif recognized by a GATA transcription factor, *elt-3*, in 602 of them. The importance of *elt-3* was validated by showing that RNAi depletion of *elt-3* activity resulted in decreased expression of 12 representative GATA-site-containing age-regulated genes. The expression of *elt-3* itself over the normal life span was negatively regulated by two other GATA transcription factors, *elt-5* and *elt-6* (Fig. 5), which were previously known to function with *elt-3* to regulate hypodermis differentiation in embryos (Gilleard *et al.*, 1999; Gilleard and McGhee, 2001). Consistent with a role for these transcription factors in regulating longevity, loss of *elt-3* suppressed the long-lifespan phenotype of a longevity mutant of *daf-2* (encoding a *C. elegans* insulin/IGF receptor), whereas *elt-5* or *elt-6* promoted aging as reduction in activity of either *elt-5* or *elt-6* caused lifespan extension. Thus, using a combination of transcriptional profiling, *cis*-regulatory analysis, and reverse genetic approaches for validation, this study identified a development-related transcriptional circuit consisting of three GATA transcription factors and revealed its novel role in regulating aging late in life.

### 3. Protein Interaction Screens

Many genetic interactions are realized in the form of direct protein–protein interactions. Analyzing interactions among proteins not only provides insight into the function of their corresponding genes, but also helps unravel the genetic topology of pathways and networks regulating biological processes. To acquire interaction information between proteins in *C. elegans*, two strategies are often employed: yeast two-hybrid (Y2H) screens (Fields and Song, 1989) and affinity-based protein isolation coupled with mass spectrometry (MS) analysis. In a typical Y2H assay, one protein is used as a bait and fused with the DNA-binding domain of a transcription factor, whereas the other protein functions as a prey and is fused with the activating domain of the transcription factor. These two fusion proteins are introduced into the

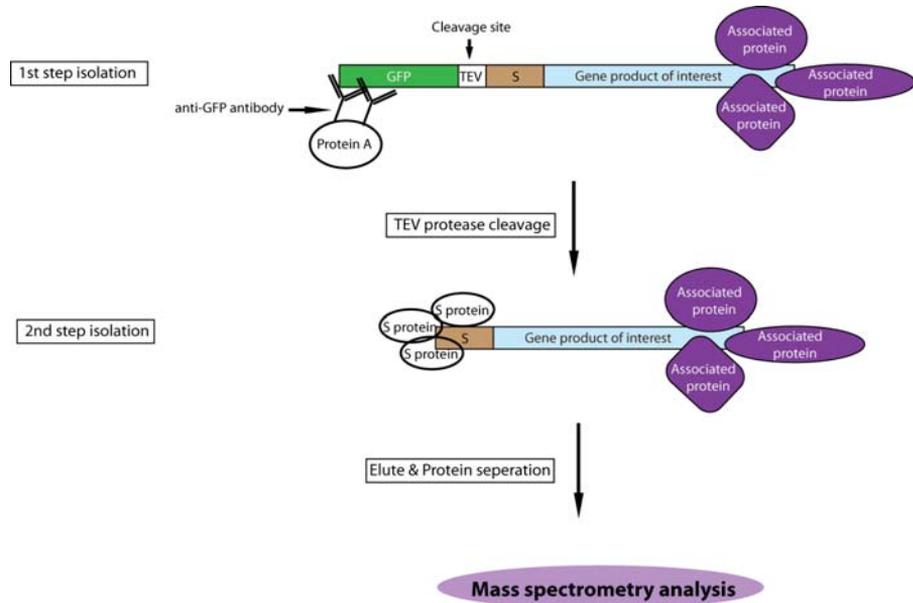


**Fig. 5** Expression changes of *elt-3/elt-5/elt-6* during normal aging and the model for this transcriptional circuit. (i) During the normal aging process, the expression of the GATA transcription factors *elt-5* and *elt-6* increases. This represses the expression of the GATA transcription factor *elt-3*. (ii) The transcriptional circuit consisting of *elt-3*, *elt-5*, and *elt-6* regulates hypodermal differentiation during embryogenesis and aging during adulthood. *elt-5* and *elt-6* promote normal aging by negatively regulating *elt-3* expression. ELT-3 regulates a group of age-dependent genes identified from DNA microarray analysis. Consistent with a functional role in aging, loss of function of *elt-3* suppresses the long-lifespan phenotype of *daf-2* mutants. The expression of *elt-3* is repressed by the DAF-2 mediated insulin signaling, indicating that this transcriptional circuit may also modulate the effects of the insulin signaling. (For color version of this figure, the reader is referred to the web version of this book.)

yeast system. If the two proteins physically interact, the two domains are brought into close proximity, which triggers transcription of a reporter gene indicating that an interaction has taken place. As an effort toward understanding protein–protein interactions, a large-scale protein interaction (interactome) mapping project based on Y2H screens is in progress ([http://interactome.dfci.harvard.edu/C\\_elegans/](http://interactome.dfci.harvard.edu/C_elegans/)) (Simonis *et al.*, 2009). This interactome map (Worm Interactome version 8) currently contains 3864 known binary protein–protein interactions. Through interactome mapping, many interactive connections involving novel proteins have been found between disparate biological processes in *C. elegans* (Boulton *et al.*, 2002; Li *et al.*, 2004; Reboul *et al.*, 2003). For example, using the Y2H screening strategy, Tewari *et al.* (2004) uncovered eight *daf-7*/TGF- $\beta$  pathway modifiers that were not identified by conventional means. They first employed six known components of the *daf-7*/TGF- $\beta$  pathway as bait for the first round Y2H screen using a *C. elegans* cDNA library. In order to identify more novel interaction links, they conducted the second round of Y2H screening using genes identified in the initial screen

as bait. Such a sequential multi-round Y2H screening approach is termed “interactome walking”, which can identify novel interactors linking distinct functions (Cusick *et al.*, 2005). In this study, they identified 71 interactions among 59 proteins, comprising a complex interactome map. Because Y2H screens might not reflect the *in vivo* functional relationships between two genes, other functional genomic techniques and conventional genetic approaches are often used for independent experimental validation. In this study, the identified interactions were confirmed by coaffinity purification assays and functionally validated by double genetic perturbation analysis in which RNAi was used to inactivate genes in loss-of-function mutant backgrounds of known *daf-7/TGF- $\beta$*  pathway genes. Through these approaches, nine genes were confirmed to interact with the *daf-7/TGF- $\beta$*  pathway, eight of which had not previously been reported to have roles in the *daf-7/TGF- $\beta$*  pathway. Given the high false-negative rate (approximately 60–70%) of Y2H screens (Walhout *et al.*, 2000) and intrinsic limitations of RNAi (discussed above), they speculated that many more interactors were likely missed in this study. Nevertheless, this study highlights the power of coupling large-scale protein interaction mapping with conventional genetic perturbation to identify components of signaling pathways in *C. elegans*.

Another strategy for identifying interacting proteins is affinity-based protein isolation coupled with MS analysis (Aebersold and Mann, 2003). This strategy is used to identify proteins that are associated with a protein of interest. These associated proteins can be isolated through two affinity purification techniques: (1) immunoprecipitation, an antibody-based purification, in which an antibody against a protein of interest is used to isolate it and its associated proteins; (2) tandem affinity purification (TAP), in which two tags separated by an enzyme-cleavable site are fused to a protein of interest and the associated proteins are isolated in two steps sequentially using antibodies or binding proteins against the two tags (Fig. 6). The tags can be fluorescent proteins, such as GFP, which allow both dynamic imaging studies and proteomic analysis. This type of tag is also referred to as the “localization and affinity purification” (LAP) tag (Cheeseman and Desai, 2005; Rigaut *et al.*, 1999). Compared with the Y2H technique in which the proteins cannot undergo some of the post-translation modifications required for particular interactions in metazoans, the tag-based strategy has several advantages: (1) the fully processed and modified protein can be used as a bait; (2) bound proteins are isolated from the native cellular environment where interactions take place; (3) multiple associated components can be isolated and analyzed at a single time (Ashman *et al.*, 2001). An example of the application of TAP/LAP coupled with MS for identifying genes was a proteomic study on *C. elegans* kinetochores (Cheeseman *et al.*, 2004), a specialized organelle that regulates chromosome segregation in mitosis and meiosis (Maiato *et al.*, 2004). To isolate the proteins involved in the assembly and function of kinetochores, they first generated a transgenic strain expressing two newly identified kinetochore proteins that were used as bait and fused with two tags, GFP and the S peptide domain. These two tags were separated by a sequence recognized by the tobacco etch virus (TEV) protease. Two sequential rounds of affinity isolation were performed. The bound proteins were first isolated using an antibody against

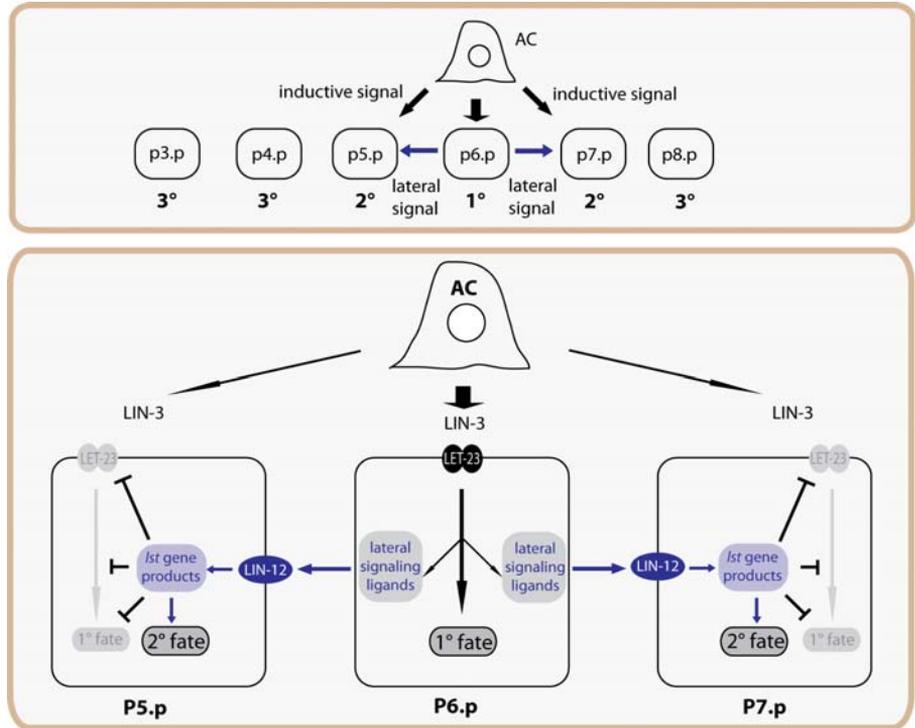


**Fig. 6** Tandem affinity purification (TAP) coupled with mass spectrometry (MS) analysis. The two tags, GFP and S peptide domain, are fused to a gene product of interest. These two tags are separated by a tobacco etch virus (TEV) protease cleavage site. The proteins associated with the gene product of interest are isolated using the antibody against the first tag, GFP, which is then released by TEV protease cleavage. S protein that binds to the S peptide domain is used for the secondary affinity purification. The associated proteins are then separated from the gene product of interest and subject to MS analysis. (For color version of this figure, the reader is referred to the web version of this book.)

GFP that was removed by TEV protease digestion. A second round of affinity purification using the second tag enriched the kinetochore complex components, which were then subject to the MS analysis. MS analysis indicated that this two-step purification process removed most of the nonspecific proteins that were present in the single-step antibody-based immunoprecipitation that was also performed in parallel. The study identified 10 kinetochore proteins, of which seven were previously uncharacterized.

#### 4. Using Bioinformatics Tools

Bioinformatics is the analysis of biological systems, especially systems involving genetic materials, using computer science, statistics, engineering, and information theory. Bioinformatic tools have been widely applied in biological research, ranging from sequence-based analysis, transcriptome analysis to computational proteomics (Rhee *et al.*, 2006). In *C. elegans*, bioinformatic tools in conjunction with functional perturbation are effective in dissecting biological processes in some circumstances, such as microRNA prediction (Grad *et al.*, 2003), gene identification by homology search (Berset *et al.*, 2001; Chen and Greenwald, 2004), and *cis*-regulatory sequence



**Fig. 7** Model for the crosstalk between LIN-3/LET-23 (ligand/receptor tyrosine kinase) signaling and LIN-12/Notch signaling in specification of vulval cell fate. Through LET-23/EGFR receptor, the graded inductive signal, LIN-3/EGF, secreted from the anchor cell (AC) promotes the 1° fate and activates expression of genes encoding ligands for the LIN-12 receptor in P6.p. Lateral signaling mediated by these ligands via LIN-12 promotes gene expression specific for the 2° fate in P5.p and P7.p, including *lst* (lateral signal target) genes. The *lst* genes that encode inhibitors of the LIN-3 signaling pathway antagonize the 1°-fate inductive effects of LIN-3 on P5.p and P7.p. (For color version of this figure, the reader is referred to the web version of this book.)

prediction (Budovskaya *et al.*, 2008; Hwang *et al.*, 2007; Yoo *et al.*, 2004). An excellent example showing the power of this approach was a study on LIN-12/Notch signaling in vulval development in *C. elegans* (Yoo *et al.*, 2004) (Fig. 7). Vulval patterning is precisely regulated through crosstalk between two pathways: the RTK/Ras/MAPK pathway mediated by LIN-3/LET-23 (ligand/receptor tyrosine kinase) and the Notch signaling pathway mediated by LIN-12 (a Notch-like receptor). *C. elegans* has six vulval precursor cells (VPCs) named consecutively P3.p to P8.p that adopt one of the three cell fates: primary fate (1°), secondary fate (2°), or tertiary fate (3°). Only the descendants of the 1° and 2° cells form the vulva. In wild-type animals, P6.p adopts the 1° fate, whereas P5.p and P7.p adopt the 2° fate. The LIN-12/Notch signaling pathway signals through a proteolytically freed intracellular fragment of LIN-12 receptor complexed with the transcription

factor LAG-1. To identify transcriptional target genes of LIN-12/Notch that antagonizes LIN-3/LET-23 signaling, Yoo *et al.* (2004) utilized computational programs to determine genes whose promoter regions contain clusters of the binding sites for LAG-1. One hundred sixty three genes were identified, two of which were previously reported to, respectively, antagonize or interact with the LIN-3/LET-23 signaling pathway during vulval development. By comparing the 5' regulatory regions of these two genes, they deduced two additional motifs that they postulated conferred tissue specificity of the 2° fated vulval cells. Through searching for genomic regions containing these motifs in the vicinity of the LAG-1 binding site clusters, they identified 10 candidate LIN-12 target genes that might act to antagonize LIN-3/LET-23 signaling in the 2° vulval cells. Of these 10 genes, five were experimentally verified as novel negative regulators of LIN-3/LET-23 signaling as the depletion of their activity using RNAi resulted in the increased activity of the LIN-3/LET-23 pathway in the 2° vulval cells. Importantly, in no case did elimination of activity of any of these genes individually disrupt 2° fate specification, suggesting that they function redundantly to inhibit LIN-3/LET-23 activity in the vulval cells. This work powerfully underscores the usefulness of functional genomic approaches, such as bioinformatic analysis, in identifying gene regulatory networks and in circumventing genetic functional redundancy.

### C. Using Systems Biology Approaches to Identify Components of Pathways

Systems biology approaches biological questions from the holistic and system-wide perspective rather than by a reductionist's gene-by-gene method. Though sometimes confusing in its definition, systems biology is emerging as an important approach for identifying and understanding gene networks in biology. All of the previously discussed genomic approaches are tools utilized in systems biology. These functional genomic approaches produce a large number of heterogeneous datasets on a genome-wide scale, such as gene expression data (transcriptome), protein–protein interaction data (interactome), and RNAi phenotypic data (phenome). Computational integration and systematic analysis of these datasets can reveal meaningful correlations that point to functionally associated components and lead to the generation of testable models that may assemble a more comprehensive picture of how gene regulatory networks control a particular biological process.

Evidence of correlations between any two types of datasets – transcriptome, interactome, and phenome – obtained from studies in yeast and worm has suggested that interacting genes appear to share similar expression, protein–protein interaction, and phenotypic profiles (Boulton *et al.*, 2002; Ge *et al.*, 2003; Jansen *et al.*, 2002; Jeong *et al.*, 2001; Kamath *et al.*, 2003; Li *et al.*, 2004; Oltvai and Barabasi, 2002; Piano *et al.*, 2002; Walhout *et al.*, 2002). Correlations across all these three types of data have been utilized to study and model the mechanisms underlying early embryogenesis in *C. elegans* (Gunsalus *et al.*, 2005). By integrating coexpression, protein–protein interaction, and phenotypic similarity datasets, Gunsalus *et al.* (2005) generated a network of 661 genes involved in early embryogenesis, including

the ribosome, proteasome, anaphase-promoting complex, and COPI coatmer, as well as complexes involved in translation initiation, nucleocytoplasmic transport, and cell polarity. They validated their predictions by testing localization of ten function-unknown genes that were predicted to associate with these “molecular machines”. Their results suggest that early embryogenesis in *C. elegans* is regulated by the coordination of a limited set of molecular machines, and provided hundreds of new putative molecular components of these machines.

Integration of functional genomic datasets can also be used across species as well as within. Because of functional conservation of orthologous genes and their genetic interactions, integration of disparate datasets from multiple organisms can provide stronger predictability for genetic interactions. An example illustrating the strength of this strategy comes from a study aimed at acquiring a global view of genetic interactions in *C. elegans* (Zhong and Sternberg, 2006). Zhong and Sternberg used a probability-based scoring system to integrate different datasets (interactome data, gene expression data, phenotype data, and functional annotation data curated from the literature) across three organisms (*S. cerevisiae*, *C. elegans*, *D. melanogaster*). They then generated a genetic interaction network consisting of 2254 genes and 18,183 predicted interactions with probability values for interactions. As a part of their experimental validation, they chose to verify the predicted interactions for two genes that have been the subjects of a number of genetic screens, *let-60*, a worm Ras gene, which plays a critical role in vulval induction, and *itr-1*, a worm 1,4,5-trisphosphate (IP3) receptor gene, which regulates pharyngeal pumping. As individual disruption of most of putative interacting genes caused no phenotype, double genetic perturbation was used for interaction validation. These genes were depleted by RNAi one at a time in the *let-60* or *itr-1* mutant backgrounds. Resulting enhancement or suppression of the phenotypic defects caused by the *let-60* or *itr-1* mutation indicated that interactions took place. Twelve of 49 predicted genes were confirmed to interact with *let-60*, and two of six genes for *itr-1*. Importantly, these 14 verified genes were novel modifiers that appeared to be missed in conventional screens, once again highlighting the ability of functional genomic/systems approaches in effectively identifying new genes and in particular functionally redundant genes or weak modifiers.

---

---

---

### III. Ordering Genes into Pathways

Historically, the ordering of genes into pathways in *C. elegans* was accomplished through genetic analysis [see review by Huang and Sternberg (1995)] with an emphasis on arranging linear genetic pathways controlling developmental processes. One realization over the past decade is that most pathways controlling biological processes are not simply linear, but rather are highly regulated, buffered with redundancy, and often branched with multiple feedback or feedforward mechanisms. Modern pathway analysis involves the combination of genetic, biochemical, cell biological, and functional genomic approaches. We outline here how these diverse strategies are used to order genes into pathways that control biological processes.

### A. Determining Whether Two Genes Function in the Same Pathway

Following the recovery of mutations based upon the strategies discussed above, great effort should be put to carefully characterize the phenotypes and genetic nature of mutations (i.e., null, hypomorphic, or hypermorphic). This information is essential for genetic interaction analysis in which the defects of double mutants are compared to those of single mutants, allowing one to determine whether mutated genes act in the same or distinct pathways that regulate a particular biological process. It is particularly important for this analysis to use null alleles or a null allele and a strong-loss-of-function allele with no activity in the assayed biological process for double mutant construction. Double mutants with a phenotypic severity similar to that of the more severe single mutant suggest that the two genes work together or in series within the same pathway (Fig. 8i). Lacking a genetic interaction, the phenotype of the double mutant would be expected to be equal to the additive defects of the single mutants (Fig. 8ii). Double mutants displaying a more severe phenotype than the expected combined loss of each indicate a synergistic (or synthetic) interaction (Fig. 8iii) (Boone *et al.*, 2007; Guarente, 1993; Mani *et al.*, 2008). The assumption in this case is that the two genes likely act in parallel pathways that converge on a common function or activity. These interpretations do not apply for genetic interactions between two hypomorphic alleles. Because of their residual activity, the linearity of the pathway where they reside is not completely interrupted, any scenario above can be caused by two hypomorphic alleles that either function in the same pathway or the distinct pathways. Other functional information is often needed to

Types of genetic interactions		Implications of interactions
(i) Similar to whichever is more severe	$AB = A$ (or $B$ )	 The same pathway
(ii) Additivity	$AB = A+B$	No genetic interaction
(iii) Synergism	$AB > A+B$	 Converge on a common function

**Fig. 8** Implications of genetic interaction between two null alleles or a null allele and a strong-loss-of-function allele with no activity in the assayed biological process. To determine whether two alleles (*A* and *B*) function in the same pathway, the double mutant (*AB*) homozygous for both alleles is constructed. Comparing the severity of this double mutant with that of single mutants can provide insight into their functional relationship. (i) If the phenotypic severity of the double mutant is similar to that of the single mutant with the more severe phenotype, it implicates that the two genes act in the same pathway; (ii) if the phenotypic severity of the double mutant is the sum of that of the single mutants, it implicates that these two genes have no genetic interaction; (iii) if the phenotypic severity of the double mutant is more than the sum of that of the single mutants, it indicates that the two genes act in parallel pathways that converge on a common function. (For color version of this figure, the reader is referred to the web version of this book.)

reveal the genetic relationship of two hypomorphic alleles. For genes that act in the same pathway, determining their genetic hierarchy often requires comprehensive analysis of the genetic, cell biological, and biochemical information about how genes and their products interact. There are no universal rules to integrate all of this information for ordering genes into biological pathways. Indeed, many strategies that successfully revealed the order of genes are highly context-specific. Below, we discuss several examples as case studies for developing approaches to order genes that regulate a biological process.

## B. Genetic Ordering of Pathways

### 1. Epistatic Analysis for Gene Ordering

As some components in a genetic pathway may play positive regulatory roles and others play negative roles, it is common that genes in the same pathway exhibit the opposite phenotypes when mutated. Epistatic analysis is a powerful way to order these components into a signaling hierarchy. The term “epistatic” was first coined in 1909 by Bateson to describe a masking effect in which an allele at one locus prevents the allele at another locus from exhibiting its phenotypes (Bateson, 1909; Cordell, 2002). Similarly, epistasis defined by molecular geneticists refers to a genetic situation in which the phenotype of a mutation in one gene is masked by the phenotype of the mutation in the other (Avery and Wasserman, 1992). This definition views a phenotype as a qualitative trait; so it is also termed “compositional epistasis” to set it apart from “statistical epistasis” used by population geneticists for quantitative differences of allele-specific effects in a population (Phillips, 2008).

Compositional epistatic analysis is particularly suitable for ordering genes whose mutations cause opposite phenotypes. It has been used in *C. elegans* to successfully construct pathways in various developmental processes such as the development of the vulva (Sternberg and Horvitz, 1989), sex determination (Goodwin and Ellis, 2002), and dauer formation (Thomas *et al.*, 1993). To perform epistatic analysis, double mutants carrying two mutations giving opposite phenotypes are constructed. The mutant phenotype that the double mutant adopts indicates the gene that is epistatic (downstream of) to the other. Two assumptions should be met prior to epistatic analysis. First, the two genes analyzed should be involved in the same pathway. Second, the opposite defective phenotypes should be direct opposite states of a genetic event assayed. For example, in vulval development, there are six VPCs. Normally, only three of these six VPCs give rise to progeny that form the vulva. A mutation in *lin-1*, encoding a transcription factor, causes more than three VPCs to adopt vulval fates, which produces the Muv phenotype. Conversely, mutations in *lin-3*, encoding an inductive cue for vulval formation, cause a reduction in vulval induction, which can lead to the Vulvaless (Vul) phenotype. These opposite phenotypes, Muv and Vul, are two opposite states in the same vulval induction pathway. Thus, epistatic analysis is applicable for ordering these two genes. As the *lin-1;lin-3* double mutant displays a Muv phenotype, *lin-1* is epistatic (downstream) of *lin-3*.

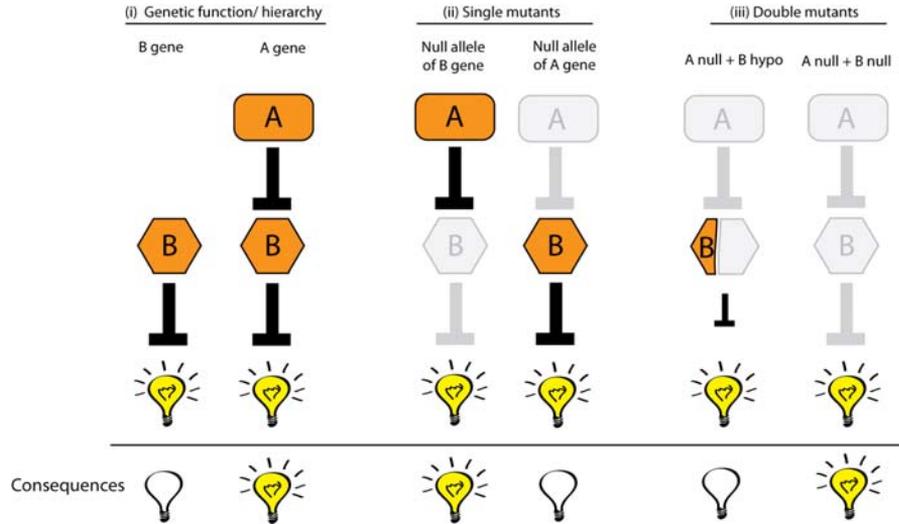
Moreover, since the LIN-3 protein is normally required for VPCs to adopt vulval fates and the LIN-1 protein inhibits the adoption of vulval fates of VPCs, the interactive relationship between them can be inferred as LIN-3 negatively regulates LIN-1 (see Fig. 1).

In cases where mutations involved in a common pathway display the same phenotype, epistatic analysis is not possible. However, some genes may have both *gf* and *lf* alleles that display the opposite phenotypes. Such gain-of-function alleles have proven to be very useful in deducing genetic hierarchies. The *gf* alleles can be either mutagen-induced or artificially engineered. One example using an artificially engineered transgene was a study examining the role of netrin signaling in axon outgrowth. A *lf* mutation in *unc-40*, encoding a netrin receptor, causes defects in axon guidance. Several *lf* mutations in genes involved in actin cytoskeleton regulation also display similar defective phenotypes. To order these genes into the *unc-40* pathway, Gitai *et al.* (2003) generated an artificial *gf* allele of *unc-40* by overexpressing and targeting the UNC-40 intracellular domain to the plasma membrane of the neuron. The engineered transgenic strain displays excessive axon outgrowth. Through epistatic analysis of the double mutants of this artificial *unc-40 gf* allele and the other actin-regulating mutants, they found that these actin-regulating genes are epistatic to *unc-40* because the mutations in these actin-regulating genes suppress the excessive axon outgrowth of the *unc-40 gf* allele. Further, these genes were found to form two bifurcated pathways downstream of *unc-40* as pairwise combinations of these mutations showed synergistic suppression of the *unc-40 gf* phenotype.

## 2. Consideration Regarding Mutants Used for Epistatic Analysis and Limitations of Epistatic Analysis

It is important to use null mutants or hypomorphic alleles with no activity in the biological process being analyzed for epistatic analysis. An appropriate choice of mutants (null or hypomorphic) based on their dosage effects is critical for successful epistatic analysis. Using any mutants with residual activity may lead to a misinterpretation of the results of epistatic analysis, particularly in cases where one tests a gene with a hypomorphic allele that acts downstream of a gene with a null allele. Because the phenotype of the double mutants of these two alleles will be similar to that of the null allele, it will lead to an inaccurate conclusion that the gene with the null allele is epistatic to (downstream of) the gene with the hypomorphic allele (Fig. 9).

An assumption of epistatic analysis is that the genes being ordered function in a linear genetic pathway. However, many pathways regulating biological processes have more complicated topological structures. They are nonlinear and often contain feedback/feedforward loops or autoregulatory elements. Moreover, the genetic hierarchy between the same components is sometimes context-dependent and can change spatially or temporally. Such a complexity limits the applicable scope of epistatic analysis. Fortunately, other tools are also available to *C. elegans* researchers to complement epistatic analysis in constructing gene regulatory pathways.



**Fig. 9** Consideration regarding alleles used for epistatic analysis (i) Normally, gene B negatively regulates a functional state (illustrated as a bulb with light), causing the inactivation of this state (illustrated as a bulb without light). Gene A, however, acts upstream of gene B in this pathway and negatively regulates it, thereby maintaining the functional state. (ii) Complete loss of gene B activity abolishes its negative regulatory effects on the functional state (“light on”). In contrast, complete loss of gene A activity removes its negative regulation on gene B, releasing gene B inhibitory effects on the functional state, therefore inactivating the functional state (“no light”). The single mutants of gene B and A display the opposite phenotypes, which makes epistasis analysis applicable for organizing genes B and A. (iii) When one performs epistasis analysis using the null allele of gene A and the hypomorphic allele for gene B, the constructed double mutant might inactivate the functional state as the hypomorphic allele of gene B preserves some residual inhibitory activity. Thus, the possible phenotype of this double mutant (“no light”) would be similar to that of the single null of gene A (“no light”), leading to a misinterpretation that gene A is epistatic to (downstream of) gene B. Using null alleles of both genes A and B, however, reveals the correct epistatic relationship of gene B downstream of gene A. (For color version of this figure, the reader is referred to the web version of this book.)

### C. Properties of Gene Products for Gene Ordering

The properties of proteins, including conserved function of their homologs/orthologs, site of action, and cellular/subcellular localization, are also useful for ordering genes into pathways. Comparative genomic analysis based on sequence homology and conservation across species can be used to deduce the potential functions of genes identified from screens. This information can provide clues for gene ordering. For example, in the case of two genes with the same phenotype when mutated, if a gene A is predicted to encode a potential transmembrane receptor and gene B encodes a potential cytoplasmic signaling transducer, this suggests that gene A likely acts upstream of gene B in a pathway, a hypothesis that can be further tested with cell biological analysis. If this hierarchical relationship between genes A and B is also conserved in other species, it would make this conclusion more solid. A good example

of this approach is previous work with the TGF- $\beta$  signaling pathway in *C. elegans*. TGF- $\beta$  signaling is involved in diverse developmental processes. A mutation in *daf-4*, encoding a receptor for TGF- $\beta$  superfamily ligands, results in small body size and morphological defects in *C. elegans* male tails. These phenotypes are identical to those of *sma-2*, *sma-3*, and *sma-4* mutants, which initially suggested that these genes were likely involved in the same biological process. To determine the hierarchical relationship between *daf-4* and *sma-2*, *sma-3* and *sma-4*, Savage *et al.* (1996) cloned and sequenced *sma-2*, *sma-3*, and *sma-4*, and found that they encode related proteins homologous to a *Drosophila* TGF- $\beta$  signaling component, *Mad*, suggesting that the three genes might be also required for TGF- $\beta$  signaling in *C. elegans*. This notion was supported by the ensuing genetic analysis revealing that the sites of action of *daf-4* and *sma-2* were in the same cell. The protein sequences of *sma-2*, *sma-3*, and *sma-4* further suggested a cytoplasmic or nuclear localization as no motifs were found to specify extracellular or transmembrane localization. Two possibilities for function arose: (1) *sma-2*, *sma-3*, and *sma-4* might function as cytoplasmic targets downstream of *daf-4*/TGF- $\beta$  signaling, or (2) *sma-2*, *sma-3*, and *sma-4* might act upstream of *daf-4* by regulating *daf-4* expression. The latter possibility was ruled out by showing that functional *daf-4* driven by a heat-shock promoter failed to rescue the defects of *sma-2*, *sma-3*, and *sma-4* mutants. They then concluded that SMA-2, SMA-3, and SMA-4 acted downstream of DAF-4, which was confirmed by later studies showing that SMA-2, SMA-3, and SMA-4 form heterotrimers for propagation of TGF- $\beta$  signaling (Savage-Dunn *et al.*, 2000; Wu *et al.*, 2001).

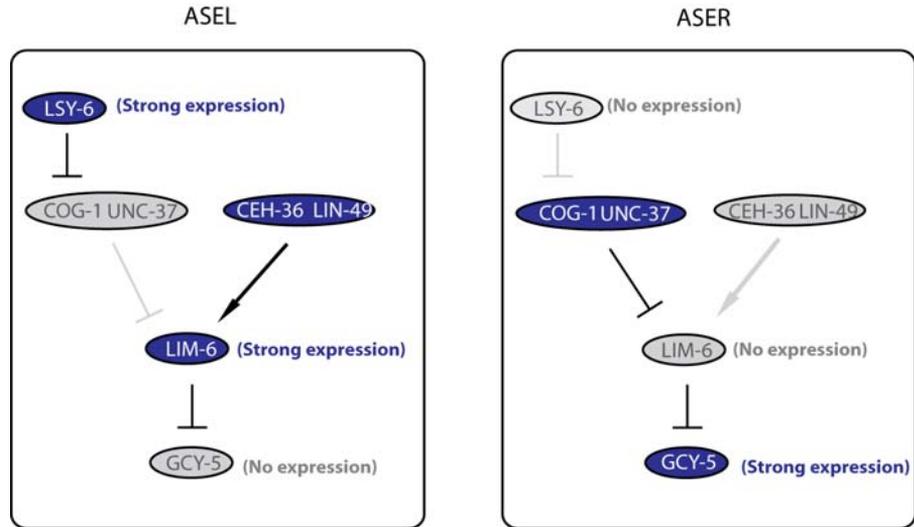
#### D. Localization Dependency of Gene Products for Gene Ordering

The function of a protein often relies on its precise cellular and subcellular localization. In cases where this localization is tightly regulated, mutations in the corresponding genes that genetically interact with these components may alter protein localization. Therefore, examining localization patterns of a gene product in different mutant backgrounds may provide useful information regarding the relative positions of two genes in a pathway or interacting genetic network. This strategy is termed as “molecular epistasis” in some studies. However, as it is so distinct from the definition of “epistasis” discussed above, it would be more precise to term this strategy “localization dependency” to reflect its nature. Our lab (Hagedorn *et al.*, 2009) utilized this strategy to determine the hierarchical relationship for two genes, *ina-1* (encoding an  $\alpha$ -integrin) and *unc-40* (encoding a receptor for the netrin ligand), during anchor cell invasion into the vulval epithelium. Both *ina-1* and *unc-40* mutants are defective in AC invasion and both disrupt F-actin localization at the invasive cell membrane. By tracking GFP fusion proteins, we found that the normal localization pattern of UNC-40 in the AC was disturbed in worms whose *ina-1* activity is depleted by RNAi. In contrast, the localization pattern of INA-1 remained normal in *unc-40* mutants. Based on this result, we concluded that INA-1 acts upstream of UNC-40 functionally in regulating F-actin formation in the AC during AC invasion.

Another example of how this approach can be used to determine the relationships of interacting gene networks involves work on the PAR proteins, which play crucial roles in establishing the anterior–posterior axis of embryos after fertilization (Munro *et al.*, 2004). The hierarchical relationships between PAR proteins are highly spatially dependent. The PAR-3 protein is normally polarized in the anterior of embryos, whereas the PAR-2 protein is enriched in the posterior. In *par-2* mutants, PAR-3 is mislocalized to the posterior, which suggests that PAR-2 excludes PAR-3 from the posterior. This indicates that *par-2* may function upstream of *par-3* in the posterior. Conversely, in the anterior of the embryo, *par-3* is upstream of *par-2* because in *par-3* mutants PAR-2 is aberrantly accumulated in the anterior. In cases like cell polarity, careful localization dependency analysis is crucial in resolving the dynamic relationship between gene products in controlling a biological process.

### E. Gene Expression Dependency for Gene Ordering

Many pathways and networks that control developmental events, such as cell-fate specification (Chang *et al.*, 2004), aging (Lin *et al.*, 2001), and developmental timing regulation (Lee *et al.*, 1993; Reinhart *et al.*, 2000), often involve transcription factors and/or components subject to post-transcriptional regulation. In these pathways and networks, interactions between two components are sometimes manifested by one component regulating expression of another. The hierarchical relationship among these components can be determined by their gene expression dependency. To determine this dependency expression of a full-length GFP reporter transgene carrying one gene (gene A) under control of its endogenous promoter is often examined in the mutant background of the other (gene B). In the case where expression of gene A is reduced in the mutant background of gene B, it suggests that gene B likely acts upstream of gene A as a positive regulator. For example, using this strategy, Johnston and Hobert (2003) placed a microRNA gene, *lisy-6*, into a transcriptional cascade of three homeobox transcription factors (*ceh-36*, *lim-6*, and *cog-1*) that regulate left–right functional asymmetrical expression of a guanyl cyclase (*gcy*) chemoreceptor gene, *gcy-5*, in ASE left (ASEL) neuron, but not in ASE right (ASER) neuron (Chang *et al.*, 2003) (Fig. 10). To order *lisy-6* into this regulatory hierarchy, they analyzed the expression of the transcription factors *ceh-36*, *lim-6*, and *cog-1* in a *lisy-6* mutant background. Expression of *ceh-36* gene was unaffected, but expression of the normally left-expressed *lim-6* gene was lost, suggesting that *lisy-6* acted upstream of *lim-6* in ASEL. Conversely, expression of *cog-1*, a negative regulator of *lim-6*, was upregulated in ASEL, indicating that *lisy-6* acted upstream of *cog-1* as a negative regulator. Consistent with this notion, the removal of *cog-1* activity in a *lisy-6* mutant background resulted in upregulation of *lim-6* expression, suggesting that *lisy-6* acts through *cog-1* to regulate *lim-6* expression.



**Fig. 10** Schematic diagram of the genetic pathway leading to *gcy-5* expression in ASER neuron. Despite bilaterally morphological symmetry, the ASEL and ASER neurons display distinct chemosensory capacities that correlate with the left–right asymmetric expression of three putative sensory receptor genes, *gcy-5*, expressed only in ASER, and *gcy-6* and *gcy-7* (not shown here) expressed only in ASEL (Chang *et al.*, 2003). A cascade consisting of a microRNA and several transcription factors act sequentially to restrict *gcy-5* gene expression to the ASER: a microRNA encoded by *lsy-6* gene, the homeobox transcription factors *cog-1*, *ceh-36*, and *lim-6*, and the transcriptional cofactors *unc-37*/Groucho and *lin-49*. In ASEL and ASER, a tightly balanced antagonistic effect between a repressor (COG-1/UNC-37) and a putative activator (CEH-36/LIN-49) complex regulates expression of *lim-6* (Chang *et al.*, 2003). In ASEL, *lsy-6* acts by targeting a complementary site in the *cog-1* 3' UTR to repress *cog-1* expression. CEH-36/LIN-49 then induces expression of *lim-6*, which subsequently represses expression of *gcy-5*. Thus, ASEL normally does not express *gcy-5*. In ASER, *lsy-6* is not expressed. The raised activity of COG-1/UNC-37 represses expression of *lim-6* by overcoming CEH-36/LIN-49-mediated induction of *lim-6* expression. Consequently, LIM-6-mediated repression of *gcy-5* expression is released, which leads to the differential expression of *gcy-5* in ASEL and ASER. More details about this pathway can be found elsewhere (Hobert, 2006). (For color version of this figure, the reader is referred to the web version of this book.)

## F. Functional Genomic Approaches for Gene Ordering

The functional genomic approaches discussed above are not only effective in identifying genes that regulate a particular biological process, but also useful in ordering these genes, because some functional genomic approaches by their nature reveal the hierarchical relationships between components of pathways. For example, in Section II.B.2., Budovskaya *et al.* (2008) bioinformatically searched for transcription factors that act upstream to control expression of several hundred genes that were found to share similar age-regulated expression changes identified by their DNA microarray analysis. In this study, the direction of the searching strategy was from the age-regulated target genes (downstream) to transcription factors (upstream). Once

the transcription factor ELT-3 was found and its transcriptional regulation on the target genes was experimentally verified, the hierarchical relationship between them was evident. Conversely, in Section II.B.4., Yoo *et al.* (2004) searched for downstream target genes of the known transcription factor LAG-1, and identified and placed five novel targets genes into the LIN-12/Notch signaling pathway.

---

---

---

## IV. Future Outlook

*C. elegans* is a powerful model organism to decode the cellular and molecular mechanisms underlying a variety of biological processes. Looking into the future, we anticipate forward genetic screens to continue to be an important approach for gene identification. More sophisticated genetic screens, such as enhancer screens on genetic backgrounds with tissue-specific gene perturbations, are expected to be carried out to cope with genetic lethality and substantial redundancy. Emerging cytometry-based automated screening techniques and whole genome sequencing for pinpointing mutation lesions will simplify mutant isolation and identification, allowing a more exhaustive and effective interrogation of genetic pathways involved in many biological processes. Technological advances in functional genomics and systems biology are revolutionizing *C. elegans* studies by providing diverse approaches to complement classic genetic screens to dissect genetic pathways/networks at unprecedented scales. Given the wealth of this functional data, we anticipate that strategies involving traditional genetics, cell biology, biochemistry, and functional genomics, will be further refined to effectively construct genetic pathways. The development of new technologies coupled with the established powerful traits of the *C. elegans* model system ensures that this organism will continue to serve an important role at the forefront of biological discovery.

## Acknowledgments

We gratefully acknowledge the support from our laboratory and deeply appreciate helpful input from all of the lab members during the preparation of this manuscript. Research in the Sherwood Lab was supported by a March of Dimes Basil O'Connor Award, Pew Scholars Award, and National Institutes of Health Grant GM079320 to D.R.S.

## References

- Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* **422**, 198–207.
- Ahringer, J. (2006). Reverse genetics. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.47.1, <http://www.wormbook.org>.
- Ahringer, J. (1997). Turn to the worm!. *Curr. Opin. Genet. Dev.* **7**, 410–415.
- Alfonso, A., Grundahl, K., Duerr, J. S., Han, H. P., and Rand, J. B. (1993). The *Caenorhabditis elegans* unc-17 gene: a putative vesicular acetylcholine transporter. *Science* **261**, 617–619.
- Aroian, R. V., Koga, M., Mendel, J. E., Ohshima, Y., and Sternberg, P. W. (1990). The *let-23* gene necessary for *Caenorhabditis elegans* vulval induction encodes a tyrosine kinase of the EGF receptor subfamily. *Nature* **348**, 693–699.

- Ashman, K., Moran, M. F., Sicheri, F., Pawson, T., and Tyers, M. (2001). Cell signalling – the proteomics of it all. *Sci. STKE* **2001**, pe33.
- Avery, L., and Wasserman, S. (1992). Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet.* **8**, 312–316.
- Barr, M. M., DeModena, J., Braun, D., Nguyen, C. Q., Hall, D. H., and Sternberg, P. W. (2001). The *Caenorhabditis elegans* autosomal dominant polycystic kidney disease gene homologs *lov-1* and *pkd-2* act in the same pathway. *Curr. Biol.* **11**, 1341–1346.
- Barr, M. M., and Sternberg, P. W. (1999). A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* **401**, 386–389.
- Bateson, W. (1909). *Mendel's Principles of Heredity*. Cambridge University Press, Cambridge, UK.
- Baugh, L. R., Demodena, J., and Sternberg, P. W. (2009). RNA Pol II accumulates at promoters of growth genes during developmental arrest. *Science* **324**, 92–94.
- Bergamaschi, D., Samuels, Y., O'Neil, N. J., Trigiante, G., Crook, T., Hsieh, J. K., O'Connor, D. J., Zhong, S., Campargue, I., and Tomlinson, M. L., et al. (2003). iASPP oncoprotein is a key inhibitor of p53 conserved from worm to human. *Nat. Genet.* **33**, 162–167.
- Berset, T., Hoier, E. F., Battu, G., Canevascini, S., and Hajnal, A. (2001). Notch inhibition of RAS signaling through MAP kinase phosphatase LIP-1 during *C. elegans* vulval development. *Science* **291**, 1055–1058.
- Boone, C., Bussey, H., and Andrews, B. J. (2007). Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449.
- Boulton, S. J., Gartner, A., Reboul, J., Vaglio, P., Dyson, N., Hill, D. E., and Vidal, M. (2002). Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* **295**, 127–131.
- Boutros, M., and Ahringer, J. (2008). The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.* **9**, 554–566.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- Budovskaya, Y. V., Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., and Kim, S. K. (2008). An *elt-3/elt-5/elt-6* GATA transcription circuit guides aging in *C. elegans*. *Cell* **134**, 291–303.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Chan, S. S., Zheng, H., Su, M. W., Wilk, R., Killeen, M. T., Hedgecock, E. M., and Culotti, J. G. (1996). UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. *Cell* **87**, 187–195.
- Chang, C., Hopper, N. A., and Sternberg, P. W. (2000). *Caenorhabditis elegans* SOS-1 is necessary for multiple RAS-mediated developmental signals. *EMBO J.* **19**, 3283–3294.
- Chang, S., Johnston Jr., R. J., Frokjaer-Jensen, C., Lockery, S., and Hobert, O. (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**, 785–789.
- Chang, S., Johnston Jr., R. J., and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev.* **17**, 2123–2137.
- Cheeseman, I. M., and Desai, A. (2005). A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci. STKE* **2005**, pl1.
- Cheeseman, I. M., Niessen, S., Anderson, S., Hyndman, F., Yates III, J. R., Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18**, 2255–2268.
- Chen, N., and Greenwald, I. (2004). The lateral signal for LIN-12/Notch in *C. elegans* vulval development comprises redundant secreted and transmembrane DSL proteins. *Dev. Cell* **6**, 183–192.
- Chong, H., Vikis, H. G., and Guan, K. L. (2003). Mechanisms of regulating the Raf kinase family. *Cell Signal.* **15**, 463–469.
- Clapham, D. E. (2003). TRP channels as cellular sensors. *Nature* **426**, 517–524.
- Clark, S. G., Stern, M. J., and Horvitz, H. R. (1992). *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* **356**, 340–344.
- Clarke, A. R., Maandag, E. R., van Roon, M., van der Lugt, N. M., van der Valk, M., Hooper, M. L., Berns, A., and te Riele, H. (1992). Requirement for a functional *Rb-1* gene in murine development. *Nature* **359**, 328–330.

- Cordell, H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* **11**, 2463–2468.
- Culetto, E., and Sattelle, D. B. (2000). A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Hum. Mol. Genet.* **9**, 869–877.
- Cusick, M. E., Klitgord, N., Vidal, M., and Hill, D. E. (2005). Interactome: gateway into systems biology. *Hum. Mol. Genet.* **14**(Spec No. 2), R171–R181.
- Dawson, I. A., Roth, S., and Artavanis-Tsakonas, S. (1995). The *Drosophila* cell cycle gene *fizzy* is required for normal degradation of cyclins A and B during mitosis and has homology to the CDC20 gene of *Saccharomyces cerevisiae*. *J. Cell. Biol.* **129**, 725–737.
- Derry, W. B., Putzke, A. P., and Rothman, J. H. (2001). *Caenorhabditis elegans* p53: role in apoptosis, meiosis, and stress resistance. *Science* **294**, 591–595.
- Dimitriadi, M., and Hart, A. C. (2010). Neurodegenerative disorders: insights from the nematode *Caenorhabditis elegans*. *Neurobiol. Dis.* **40**, 4–11.
- Doitsidou, M., Flames, N., Lee, A. C., Boyanov, A., and Hobert, O. (2008). Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nat. Methods* **5**, 869–872.
- Dorfman, M., Gomes, J. E., O'Rourke, S., and Bowerman, B. (2009). Using RNA interference to identify specific modifiers of a temperature-sensitive, embryonic-lethal mutation in the *Caenorhabditis elegans* ubiquitin-like Nedd8 protein modification pathway E1-activating gene *rfl-1*. *Genetics* **182**, 1035–1049.
- Du, W., and Dyson, N. (1999). The role of RBF in the introduction of G1 regulation during *Drosophila* embryogenesis. *EMBO J.* **18**, 916–925.
- Dupuy, D., Bertin, N., Hidalgo, C. A., Venkatesan, K., Tu, D., Lee, D., Rosenberg, J., Svzrikapa, N., Blanc, A., and Carnec, A., *et al.* (2007). Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat. Biotechnol.* **25**, 663–668.
- Fay, D. S., Keenan, S., and Han, M. (2002). *fzr-1* and *lin-35/Rb* function redundantly to control cell proliferation in *C. elegans* as revealed by a nonbiased synthetic screen. *Genes Dev.* **16**, 503–517.
- Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408**, 325–330.
- Friedman, D. B., and Johnson, T. E. (1988). A mutation in the *age-1* gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility. *Genetics* **118**, 75–86.
- Fuhrman, L. E., Goel, A. K., Smith, J., Shianna, K. V., and Aballay, A. (2009). Nucleolar proteins suppress *Caenorhabditis elegans* innate immunity by inhibiting p53/CEP-1. *PLoS Genet.* **5**, e1000657.
- Gao, M. X., Liao, E. H., Yu, B., Wang, Y., Zhen, M., and Derry, W. B. (2008). The SCF FSN-1 ubiquitin ligase controls germline apoptosis through CEP-1/p53 in *C. elegans*. *Cell Death Differ.* **15**, 1054–1062.
- Ge, H., Walhout, A. J., and Vidal, M. (2003). Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet.* **19**, 551–560.
- Gengyo-Ando, K., and Mitani, S. (2000). Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **269**, 64–69.
- Gilleard, J. S., and McGhee, J. D. (2001). Activation of hypodermal differentiation in the *Caenorhabditis elegans* embryo by GATA transcription factors *ELT-1* and *ELT-3*. *Mol. Cell. Biol.* **21**, 2533–2544.
- Gilleard, J. S., Shafi, Y., Barry, J. D., and McGhee, J. D. (1999). *ELT-3*: a *Caenorhabditis elegans* GATA factor expressed in the embryonic epidermis during morphogenesis. *Dev. Biol.* **208**, 265–280.
- Giltsdorf, M., Horn, T., Arziman, Z., Pelz, O., Kiner, E., and Boutros, M. (2010). GenomeRNAi: a database for cell-based RNAi phenotypes. 2009 update. *Nucleic Acids Res.* **38**, D448–D452.
- Gitai, Z., Yu, T. W., Lundquist, E. A., Tessier-Lavigne, M., and Bargmann, C. I. (2003). The netrin receptor UNC-40/DCC stimulates axon attraction and outgrowth through enabled and, in parallel, Rac and UNC-115/AbLIM. *Neuron* **37**, 53–65.

- Goldstein, B., and Hird, S. N. (1996). Specification of the anteroposterior axis in *Caenorhabditis elegans*. *Development* **122**, 1467–1474.
- Goodwin, E. B., and Ellis, R. E. (2002). Turning clustering loops: sex determination in *Caenorhabditis elegans*. *Curr. Biol.* **12**, R111–R120.
- Grad, Y., Aach, J., Hayes, G. D., Reinhart, B. J., Church, G. M., Ruvkun, G., and Kim, J. (2003). Computational and experimental identification of *C. elegans* microRNAs. *Mol. Cell* **11**, 1253–1263.
- Grant, B., and Hirsh, D. (1999). Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. *Mol. Biol. Cell* **10**, 4311–4326.
- Grant, B., Zhang, Y., Paupard, M. C., Lin, S. X., Hall, D. H., and Hirsh, D. (2001). Evidence that RME-1, a conserved *C. elegans* EH-domain protein, functions in endocytic recycling. *Nat. Cell Biol.* **3**, 573–579.
- Grant, B. D., and Wilkinson, H. A. (2003). Functional genomic maps in *Caenorhabditis elegans*. *Curr. Opin. Cell Biol.* **15**, 206–212.
- Greenwald, I. S., Sternberg, P. W., and Horvitz, H. R. (1983). The *lin-12* locus specifies cell fates in *Caenorhabditis elegans*. *Cell* **34**, 435–444.
- Guarente, L. (1993). Synthetic enhancement in gene interaction: a genetic tool come of age. *Trends Genet.* **9**, 362–366.
- Gunsalus, K. C., Ge, H., Schetter, A. J., Goldberg, D. S., Han, J. D., Hao, T., Berriz, G. F., Bertin, N., Huang, J., and Chuang, L. S., *et al.* (2005). Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* **436**, 861–865.
- Gunsalus, K. C., Yueh, W. C., MacMenamin, P., and Piano, F. (2004). RNAiDB and PhenoBlast: web tools for genome-wide phenotypic mapping projects. *Nucleic Acids Res.* **32**, D406–D410.
- Guthrie, C. R., Schellenberg, G. D., and Kraemer, B. C. (2009). SUT-2 potentiates tau-induced neurotoxicity in *Caenorhabditis elegans*. *Hum. Mol. Genet.* **18**, 1825–1838.
- Hagedorn, E. J., Yashiro, H., Ziel, J. W., Ihara, S., Wang, Z., and Sherwood, D. R. (2009). Integrin acts upstream of netrin signaling to regulate formation of the anchor cell's invasive membrane in *C. elegans*. *Dev. Cell* **17**, 187–198.
- Hajdu-Cronin, Y. M., Chen, W. J., Patikoglou, G., Koelle, M. R., and Sternberg, P. W. (1999). Antagonism between G(o)alpha and G(q)alpha in *Caenorhabditis elegans*: the RGS protein EAT-16 is necessary for G(o)alpha signaling and regulates G(q)alpha activity. *Genes Dev.* **13**, 1780–1793.
- Hajnal, A., Whitfield, C. W., and Kim, S. K. (1997). Inhibition of *Caenorhabditis elegans* vulval induction by *gap-1* and by *let-23* receptor tyrosine kinase. *Genes Dev.* **11**, 2715–2728.
- Han, M., Golden, A., Han, Y., and Sternberg, P. W. (1993). *C. elegans lin-45 raf* gene participates in *let-60 ras*-stimulated vulval differentiation. *Nature* **363**, 133–140.
- Hao, J. C., Yu, T. W., Fujisawa, K., Culotti, J. G., Gengyo-Ando, K., Mitani, S., Moulder, G., Barstead, R., Tessier-Lavigne, M., and Bargmann, C. I. (2001). *C. elegans slit* acts in midline, dorsal-ventral, and anterior-posterior guidance via the SAX-3/Robo receptor. *Neuron* **32**, 25–38.
- Harris, P. C., and Torres, V. E. (2009). Polycystic kidney disease. *Annu. Rev. Med.* **60**, 321–337.
- Harris, R., Sabatelli, L. M., and Seeger, M. A. (1996). Guidance cues at the *Drosophila* CNS midline: identification and characterization of two *Drosophila* Netrin/UNC-6 homologs. *Neuron* **17**, 217–228.
- Harris, T. W., Antoshechkin, I., Bieri, T., Blasiar, D., Chan, J., Chen, W. J., De La Cruz, N., Davis, P., Duesbury, M., and Fang, R., *et al.* (2010). WormBase: a comprehensive resource for nematode research. *Nucleic Acids Res.* **38**, D463–D467.
- Hartman, J. L. t., Garvik, B., and Hartwell, L. (2001). Principles for the buffering of genetic variation. *Science* **291**, 1001–1004.
- Hedgecock, E. M., Culotti, J. G., and Hall, D. H. (1990). The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. *Neuron* **4**, 61–85.
- Hill, R. J., and Sternberg, P. W. (1992). The gene *lin-3* encodes an inductive signal for vulval development in *C. elegans*. *Nature* **358**, 470–476.
- Hobert, O. (2006). Architecture of a microRNA-controlled gene regulatory network that diversifies neuronal cell fates. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 181–188.
- Hodgkin, J. (2005). Genetic suppression. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.59.1, <http://www.wormbook.org>.

- Hodgkin, J., and Herman, R. K. (1998). Changing styles in *C. elegans* genetics. *Trends Genet.* **14**, 352–357.
- Hoepfner, D. J., Spector, M. S., Ratliff, T. M., Kinchen, J. M., Granat, S., Lin, S. C., Bhusri, S. S., Conradt, B., Herman, M. A., and Hengartner, M. O. (2004). EOR-1 and EOR-2 are required for cell-specific apoptotic death in *C. elegans*. *Dev. Biol.* **274**, 125–138.
- Hopper, N. A., Lee, J., and Sternberg, P. W. (2000). ARK-1 inhibits EGFR signaling in *C. elegans*. *Mol. Cell* **6**, 65–75.
- Howell, K., Arur, S., Schedl, T., and Sundaram, M. V. (2010). EOR-2 is an obligate binding partner of the BTB-zinc finger protein EOR-1 in *Caenorhabditis elegans*. *Genetics* **184**, 899–913.
- Hsu, V., Zobel, C. L., Lambie, E. J., Schedl, T., and Kornfeld, K. (2002). *Caenorhabditis elegans* lin-45 raf is essential for larval viability, fertility and the induction of vulval cell fates. *Genetics* **160**, 481–492.
- Huang, L. S., and Sternberg, P. W. (1995). Genetic dissection of developmental pathways. *Methods Cell Biol.* **48**, 97–122.
- Hudson, M. L., Kinnunen, T., Cinar, H. N., and Chisholm, A. D. (2006). *C. elegans* Kallmann syndrome protein KAL-1 interacts with syndecan and glypican to regulate neuronal cell migrations. *Dev. Biol.* **294**, 352–365.
- Hwang, B. J., Meruelo, A. D., and Sternberg, P. W. (2007). *C. elegans* EVI1 proto-oncogene, EGL-43, is necessary for Notch-mediated cell fate specification and regulates cell invasion. *Development* **134**, 669–679.
- Igarashi, P., and Somlo, S. (2002). Genetics and pathogenesis of polycystic kidney disease. *J. Am. Soc. Nephrol.* **13**, 2384–2398.
- Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G., and Hedgecock, E. M. (1992). UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. *Neuron* **9**, 873–881.
- Jacks, T., Fazeli, A., Schmitt, E. M., Bronson, R. T., Goodell, M. A., and Weinberg, R. A. (1992). Effects of an Rb mutation in the mouse. *Nature* **359**, 295–300.
- Jansen, G., Hazendonk, E., Thijssen, K. L., and Plasterk, R. H. (1997). Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* **17**, 119–121.
- Jansen, R., Greenbaum, D., and Gerstein, M. (2002). Relating whole-genome expression data with protein-protein interactions. *Genome Res.* **12**, 37–46.
- Jeong, H., Mason, S. P., Barabasi, A. L., and Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature* **411**, 41–42.
- Johnsen, R. C., and Baillie, D. L. (1991). Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans*. *Genetics* **129**, 735–752.
- Johnston, R. J., and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**, 845–849.
- Jorgensen, E. M., and Mango, S. E. (2002). The art and design of genetic screens: *caenorhabditis elegans*. *Nat. Rev. Genet.* **3**, 356–369.
- Kaletta, T., and Hengartner, M. O. (2006). Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discov.* **5**, 387–398.
- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., and Sohrmann, M., *et al.* (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237.
- Kemphues K. (2005). Essential Genes. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.57.1, <http://www.wormbook.org>.
- Kemphues, K. J., Priess, J. R., Morton, D. G., and Cheng, N. S. (1988). Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. *Cell* **52**, 311–320.
- Kennedy, T. E., Serafini, T., de la Torre, J. R., and Tessier-Lavigne, M. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435.
- Kim, H., Rogers, M. J., Richmond, J. E., and McIntire, S. L. (2004). SNF-6 is an acetylcholine transporter interacting with the dystrophin complex in *Caenorhabditis elegans*. *Nature* **430**, 891–896.
- Kim, S. K. (2001). [Http://C. elegans](http://C.elegans): mining the functional genomic landscape. *Nat. Rev. Genet.* **2**, 681–689.

- Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev. Biol.* **70**, 396–417.
- Kimura, K. D., Tissenbaum, H. A., Liu, Y., and Ruvkun, G. (1997). *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* **277**, 942–946.
- Kirienko, N. V., Mani, K., and Fay, D. S. (2010). Cancer models in *Caenorhabditis elegans*. *Dev. Dyn.* **239**, 1413–1448.
- Kornfeld, K., Guan, K. L., and Horvitz, H. R. (1995a). The *Caenorhabditis elegans* gene *mek-2* is required for vulval induction and encodes a protein similar to the protein kinase MEK. *Genes Dev.* **9**, 756–768.
- Kornfeld, K., Hom, D. B., and Horvitz, H. R. (1995b). The *ksr-1* gene encodes a novel protein kinase involved in Ras-mediated signaling in *C. elegans*. *Cell* **83**, 903–913.
- Lackner, M. R., Kornfeld, K., Miller, L. M., Horvitz, H. R., and Kim, S. K. (1994). A MAP kinase homolog, *mpk-1*, is involved in ras-mediated induction of vulval cell fates in *Caenorhabditis elegans*. *Genes Dev.* **8**, 160–173.
- Lakso, M., Vartiainen, S., Moilanen, A. M., Sirvio, J., Thomas, J. H., Nass, R., Blakely, R. D., and Wong, G. (2003). Dopaminergic neuronal loss and motor deficits in *Caenorhabditis elegans* overexpressing human alpha-synuclein. *J. Neurochem.* **86**, 165–172.
- Lauderdale, J. D., Davis, N. M., and Kuwada, J. Y. (1997). Axon tracts correlate with netrin-1a expression in the zebrafish embryo. *Mol. Cell. Neurosci.* **9**, 293–313.
- Lee, J., Jongeward, G. D., and Sternberg, P. W. (1994). *unc-101*, a gene required for many aspects of *Caenorhabditis elegans* development and behavior, encodes a clathrin-associated protein. *Genes Dev.* **8**, 60–73.
- Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.
- Lee, V. M., Goedert, M., and Trojanowski, J. Q. (2001). Neurodegenerative tauopathies. *Annu. Rev. Neurosci.* **24**, 1121–1159.
- Lehner, B., Tischler, J., and Fraser, A. G. (2006). RNAi screens in *Caenorhabditis elegans* in a 96-well liquid format and their application to the systematic identification of genetic interactions. *Nat. Protoc.* **1**, 1617–1620.
- Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., and Hao, T., *et al.* (2004). A map of the interactome network of the metazoan *C. elegans*. *Science* **303**, 540–543.
- Liao, E. H., Hung, W., Abrams, B., and Zhen, M. (2004). An SCF-like ubiquitin ligase complex that controls presynaptic differentiation. *Nature* **430**, 345–350.
- Lin, K., Hsin, H., Libina, N., and Kenyon, C. (2001). Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nat. Genet.* **28**, 139–145.
- Liu, L. X., Spoerke, J. M., Mulligan, E. L., Chen, J., Reardon, B., Westlund, B., Sun, L., Abel, K., Armstrong, B., and Hardiman, G., *et al.* (1999). High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.* **9**, 859–867.
- Lu, X., and Horvitz, H. R. (1998). *lin-35* and *lin-53*, two genes that antagonize a *C. elegans* Ras pathway, encode proteins similar to Rb and its binding protein RbAp48. *Cell* **95**, 981–991.
- Maiato, H., DeLuca, J., Salmon, E. D., and Earnshaw, W. C. (2004). The dynamic kinetochore-microtubule interface. *J. Cell Sci.* **117**, 5461–5477.
- Matus, D. Q., Li, X. Y., Durbin, S., Agarwal, D., Chi, Q., Weiss, S. J., and Sherwood, D. R. (2010). In vivo identification of regulators of cell invasion across basement membranes. *Sci. Signal* **3**, ra35.
- Mani, R., St Onge, R. P., Hartman IV, J. L., Giaever, G., and Roth, F. P. (2008). Defining genetic interaction. *Proc. Natl Acad. Sci. U. S. A.* **105**, 3461–3466.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Metzstein, M. M., Stanfield, G. M., and Horvitz, H. R. (1998). Genetics of programmed cell death in *C. elegans*: past, present and future. *Trends Genet.* **14**, 410–416.
- Miller III, D. M., Niemeyer, C. J., and Chitkara, P. (1993). Dominant *unc-37* mutations suppress the movement defect of a homeodomain mutation in *unc-4*, a neural specificity gene in *Caenorhabditis elegans*. *Genetics* **135**, 741–753.

- Miller, K. G., Emerson, M. D., and Rand, J. B. (1999). Gqalpha and diacylglycerol kinase negatively regulate the Gqalpha pathway in *C. elegans*. *Neuron* **24**, 323–333.
- Mitchell, K. J., Doyle, J. L., Serafini, T., Kennedy, T. E., Tessier-Lavigne, M., Goodman, C. S., and Dickson, B. J. (1996). Genetic analysis of Netrin genes in *Drosophila*: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* **17**, 203–215.
- Munro, E., Nance, J., and Priess, J. R. (2004). Cortical flows powered by asymmetrical contraction transport PAR proteins to establish and maintain anterior-posterior polarity in the early *C. elegans* embryo. *Dev. Cell* **7**, 413–424.
- Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**, 277–283.
- Nakata, K., Abrams, B., Grill, B., Goncharov, A., Huang, X., Chisholm, A. D., and Jin, Y. (2005). Regulation of a DLK-1 and p38 MAP kinase pathway by the ubiquitin ligase RPM-1 is required for presynaptic development. *Cell* **120**, 407–420.
- Nauli, S. M., Alenghat, F. J., Luo, Y., Williams, E., Vassilev, P., Li, X., Elia, A. E., Lu, W., Brown, E. M., and Quinn, S. J., *et al.* (2003). Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat. Genet.* **33**, 129–137.
- Nurrish, S., Segalat, L., and Kaplan, J. M. (1999). Serotonin inhibition of synaptic transmission: Galpha (0) decreases the abundance of UNC-13 at release sites. *Neuron* **24**, 231–242.
- Oltvai, Z. N., and Barabasi, A. L. (2002). Systems biology Life's complexity pyramid. *Science* **298**, 763–764.
- Pazour, G. J., San Agustin, J. T., Follit, J. A., Rosenbaum, J. L., and Witman, G. B. (2002). Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Curr. Biol.* **12**, R378–R380.
- Phillips, P. C. (2008). Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867.
- Piano, F., Gunsalus, K. C., Hill, D. E., and Vidal, M. (2006). *C. elegans* network biology: a beginning. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.118.1, <http://www.wormbook.org>.
- Piano, F., Schetter, A. J., Morton, D. G., Gunsalus, K. C., Reinke, V., Kim, S. K., and Kemphues, K. J. (2002). Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Curr. Biol.* **12**, 1959–1964.
- Pierce, S. B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S. A., Buchman, A. R., Ferguson, K. C., Heller, J., Platt, D. M., and Pasquinelli, A. A., *et al.* (2001). Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* **15**, 672–686.
- Polanowska, J., Martin, J. S., Fisher, R., Scopa, T., Rae, I., and Boulton, S. J. (2004). Tandem immunoaffinity purification of protein complexes from *Caenorhabditis elegans*. *Biotechniques* **36**, 778–780 782.
- Pollock, D. D., and Larkin, J. C. (2004). Estimating the degree of saturation in mutant screens. *Genetics* **168**, 489–502.
- Qiao, L., Lissemore, J. L., Shu, P., Smardon, A., Gelber, M. B., and Maine, E. M. (1995). Enhancers of *glp-1*, a gene required for cell-signaling in *Caenorhabditis elegans*, define a set of genes required for germline development. *Genetics* **141**, 551–569.
- Reboul, J., Vaglio, P., Rual, J. F., Lamesch, P., Martinez, M., Armstrong, C. M., Li, S., Jacotot, L., Bertin, N., and Janky, R., *et al.* (2003). *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* **34**, 35–41.
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906.
- Rhee, S. Y., Dickerson, J., and Xu, D. (2006). Bioinformatics and its applications in plant biology. *Annu. Rev. Plant Biol.* **57**, 335–360.

- Rhiner, C., Gysi, S., Frohli, E., Hengartner, M. O., and Hajnal, A. (2005). Syndecan regulates cell migration and axon guidance in *C. elegans*. *Development* **132**, 4621–4633.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032.
- Rocheleau, C. E., Howard, R. M., Goldman, A. P., Volk, M. L., Girard, L. J., and Sundaram, M. V. (2002). A lin-45 raf enhancer screen identifies EOR-1, EOR-2 and unusual alleles of Ras pathway genes in *Caenorhabditis elegans*. *Genetics* **161**, 121–131.
- Rogers, A., Antoshechkin, I., Bieri, T., Blasiar, D., Bastiani, C., Canaran, P., Chan, J., Chen, W. J., Davis, P., and Fernandes, J., *et al.* (2008). WormBase 2007. *Nucleic Acids Res.* **36**, D612–D617.
- Rual, J. F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., and van den Heuvel, S., *et al.* (2004). Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162–2168.
- Rubin, C., Gur, G., and Yarden, Y. (2005). Negative regulation of receptor tyrosine kinases: unexpected links to c-Cbl and receptor ubiquitylation. *Cell Res.* **15**, 66–71.
- Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., and Fleischmann, W., *et al.* (2000). Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215.
- Rushforth, A. M., Saari, B., and Anderson, P. (1993). Site-selected insertion of the transposon Tc1 into a *Caenorhabditis elegans* myosin light chain gene. *Mol. Cell. Biol.* **13**, 902–910.
- Sato, M., Sato, K., Liou, W., Pant, S., Harada, A., and Grant, B. D. (2008). Regulation of endocytic recycling by *C. elegans* Rab35 and its regulator RME-4, a coated-pit protein. *EMBO J.* **27**, 1183–1196.
- Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sun, C. Y., Baird, S. E., and Padgett, R. W. (1996). *Caenorhabditis elegans* genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components. *Proc. Natl Acad. Sci. U. S. A.* **93**, 790–794.
- Savage-Dunn, C., Tokarz, R., Wang, H., Cohen, S., Giannikas, C., and Padgett, R. W. (2000). SMA-3 smad has specific and critical functions in DBL-1/SMA-6 TGFbeta-related signaling. *Dev. Biol.* **223**, 70–76.
- Schumacher, B., Hanazawa, M., Lee, M. H., Nayak, S., Volkmann, K., Hofmann, E. R., Hengartner, M., Schedl, T., and Gartner, A. (2005). Translational repression of *C. elegans* p53 by GLD-1 regulates DNA damage-induced apoptosis. *Cell* **120**, 357–368.
- Schwab, M., Lutum, A. S., and Seufert, W. (1997). Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell* **90**, 683–693.
- Schwabiuk, M., Coudiere, L., and Merz, D. C. (2009). SDN-1/syndecan regulates growth factor signaling in distal tip cell migrations in *C. elegans*. *Dev. Biol.* **334**, 235–242.
- Sendoel, A., Kohler, I., Fellmann, C., Lowe, S. W., and Hengartner, M. O. (2010). HIF-1 antagonizes p53-mediated apoptosis through a secreted neuronal tyrosinase. *Nature* **465**, 577–583.
- Serafini, T., Kennedy, T. E., Galko, M. J., Mirzayan, C., Jessell, T. M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. *Cell* **78**, 409–424.
- Shen, K., and Bargmann, C. I. (2003). The immunoglobulin superfamily protein SYG-1 determines the location of specific synapses in *C. elegans*. *Cell* **112**, 619–630.
- Sherwood, D. R. (2006). Cell invasion through basement membranes: an anchor of understanding. *Trends Cell Biol.* **16**, 250–256.
- Sherwood, D. R., Butler, J. A., Kramer, J. M., and Sternberg, P. W. (2005). FOS-1 promotes basement-membrane removal during anchor-cell invasion in *C. elegans*. *Cell* **121**, 951–962.
- Shim, J., Sternberg, P. W., and Lee, J. (2000). Distinct and redundant functions of mu1 medium chains of the AP-1 clathrin-associated protein complex in the nematode *Caenorhabditis elegans*. *Mol. Biol. Cell* **11**, 2743–2756.
- Sigrist, S., Jacobs, H., Stratmann, R., and Lehner, C. F. (1995). Exit from mitosis is regulated by *Drosophila* fizzy and the sequential destruction of cyclins A, B and B3. *EMBO J.* **14**, 4827–4838.

- Simmer, F., Moorman, C., van der Linden, A. M., Kuijk, E., van den Berghe, P. V., Kamath, R. S., Fraser, A. G., Ahringer, J., and Plasterk, R. H. (2003). Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions. *PLoS Biol.* **1**, E12.
- Simon, D. J., Madison, J. M., Conery, A. L., Thompson-Peer, K. L., Soskis, M., Ruvkun, G. B., Kaplan, J. M., and Kim, J. K. (2008). The microRNA miR-1 regulates a MEF-2-dependent retrograde signal at neuromuscular junctions. *Cell* **133**, 903–915.
- Simonis, N., Rual, J. F., Carvunis, A. R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Sahalie, J. M., Venkatesan, K., and Gebreab, F., *et al.* (2009). Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. *Nat. Methods* **6**, 47–54.
- Singh, N., and Han, M. (1995). sur-2, a novel gene, functions late in the let-60 ras-mediated signaling pathway during *Caenorhabditis elegans* vulval induction. *Genes Dev.* **9**, 2251–2265.
- Sonnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A. M., Artelt, J., Bettencourt, P., and Cassin, E., *et al.* (2005). Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469.
- Spilker, A. C., Rabilotta, A., Zbinden, C., Labbe, J. C., and Gotta, M. (2009). MAP kinase signaling antagonizes PAR-1 function during polarization of the early *Caenorhabditis elegans* embryo. *Genetics* **183**, 965–977.
- Sternberg, P. W. (2004). Developmental biology: a pattern of precision. *Science* **303**, 637–638.
- Sternberg, P. W., Golden, A., and Han, M. (1993). Role of a raf proto-oncogene during *Caenorhabditis elegans* vulval development. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **340**, 259–265.
- Sternberg, P. W., and Horvitz, H. R. (1986). Pattern formation during vulval development in *C. elegans*. *Cell* **44**, 761–772.
- Sternberg, P. W., and Horvitz, H. R. (1989). The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. *Cell* **58**, 679–693.
- Sternberg, P. W., Yoon, C. H., Lee, J., Jongeward, G. D., Kayne, P. S., Katz, W. S., Lesa, G., Liu, J., Golden, A., and Huang, L. S., *et al.* (1994). Molecular genetics of proto-oncogenes and candidate tumor suppressors in *Caenorhabditis elegans*. *Cold Spring Harb. Symp. Quant. Biol.* **59**, 155–163.
- Styer, K. L., Singh, V., Macosko, E., Steele, S. E., Bargmann, C. I., and Aballay, A. (2008). Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* **322**, 460–464.
- Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
- Sundaram, M. V. (2006). RTK/Ras/MAPK signaling. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.80.1, <http://www.wormbook.org>.
- Sundaram, M., and Han, M. (1995). The *C. elegans* ksr-1 gene encodes a novel Raf-related kinase involved in Ras-mediated signal transduction. *Cell* **83**, 889–901.
- Swaminathan, G., and Tsygankov, A. Y. (2006). The Cbl family proteins: ring leaders in regulation of cell signaling. *J. Cell Physiol.* **209**, 21–43.
- Tan, P. B., Lackner, M. R., and Kim, S. K. (1998). MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31 WH transcription factor complex during *C. elegans* vulval induction. *Cell* **93**, 569–580.
- Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A., and Driscoll, M. (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat. Genet.* **24**, 180–183.
- Tewari, M., Hu, P. J., Ahn, J. S., Ayivi-Guedehoussou, N., Vidalain, P. O., Li, S., Milstein, S., Armstrong, C. M., Boxem, M., and Butler, M. D., *et al.* (2004). Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF-beta signaling network. *Mol. Cell* **13**, 469–482.
- Thomas, J. H., Birnby, D. A., and Vowels, J. J. (1993). Evidence for parallel processing of sensory information controlling dauer formation in *Caenorhabditis elegans*. *Genetics* **134**, 1105–1117.
- Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995). Serial analysis of gene expression. *Science* **270**, 484–487.
- Wagner, A. (2000). Robustness against mutations in genetic networks of yeast. *Nat. Genet.* **24**, 355–361.

- Walhout, A. J., Reboul, J., Shtanko, O., Bertin, N., Vaglio, P., Ge, H., Lee, H., Doucette-Stamm, L., Gunsalus, K. C., and Schetter, A. J., *et al.* (2002). Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline. *Curr. Biol.* **12**, 1952–1958.
- Walhout, A. J., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., and Vidal, M. (2000). Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science* **287**, 116–122.
- Walthall, W. W., and Chalfie, M. (1988). Cell-cell interactions in the guidance of late-developing neurons in *Caenorhabditis elegans*. *Science* **239**, 643–645.
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63.
- Winnier, A. R., Meir, J. Y., Ross, J. M., Tavernarakis, N., Driscoll, M., Ishihara, T., Katsura, I., and Miller III, D. M. (1999). UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in *Caenorhabditis elegans*. *Genes Dev.* **13**, 2774–2786.
- Wu, J. W., Hu, M., Chai, J., Seoane, J., Huse, M., Li, C., Rigotti, D. J., Kyin, S., Muir, T. W., and Fairman, R., *et al.* (2001). Crystal structure of a phosphorylated Smad2. Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-beta signaling. *Mol. Cell* **8**, 1277–1289.
- Wu, Y., and Han, M. (1994). Suppression of activated Let-60 ras protein defines a role of *Caenorhabditis elegans* Sur-1 MAP kinase in vulval differentiation. *Genes Dev.* **8**, 147–159.
- Wu, Y., Han, M., and Guan, K. L. (1995). MEK-2, a *Caenorhabditis elegans* MAP kinase kinase, functions in Ras-mediated vulval induction and other developmental events. *Genes Dev.* **9**, 742–755.
- Xiao, Z. S., and Quarles, L. D. (2010). Role of the polycystin-primary cilia complex in bone development and mechanosensing. *Ann. N. Y. Acad. Sci.* **1192**, 410–421.
- Yoder, B. K., Hou, X., and Guay-Woodford, L. M. (2002). The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *J. Am. Soc. Nephrol.* **13**, 2508–2516.
- Yoo, A. S., Bais, C., and Greenwald, I. (2004). Crosstalk between the EGFR and LIN-12/Notch pathways in *C. elegans* vulval development. *Science* **303**, 663–666.
- Zallen, J. A., Yi, B. A., and Bargmann, C. I. (1998). The conserved immunoglobulin superfamily member SAX-3/Robo directs multiple aspects of axon guidance in *C. elegans*. *Cell* **92**, 217–227.
- Zhen, M., Huang, X., Bamber, B., and Jin, Y. (2000). Regulation of presynaptic terminal organization by *C. elegans* RPM-1, a putative guanine nucleotide exchanger with a RING-H2 finger domain. *Neuron* **26**, 331–343.
- Zheng, Y., Mellem, J. E., Brockie, P. J., Madsen, D. M., and Maricq, A. V. (2004). SOL-1 is a CUB-domain protein required for GLR-1 glutamate receptor function in *C. elegans*. *Nature* **427**, 451–457.
- Zhong, W., and Sternberg, P. W. (2006). Genome-wide prediction of *C. elegans* genetic interactions. *Science* **311**, 1481–1484.
- Ziel, J. W., Hagedorn, E. J., Audhya, A., and Sherwood, D. R. (2009). UNC-6 (netrin) orients the invasive membrane of the anchor cell in *C. elegans*. *Nat. Cell Biol.* **11**, 183–189.

---

---

**PART II**

Molecular Biology and Biochemistry



---

---

## CHAPTER 6

# Transgenesis in *C. elegans*

Vida Praitis<sup>†</sup> and Morris F. Maduro<sup>\*</sup>

<sup>\*</sup>Department of Biology, University of California Riverside, Riverside, California, USA

<sup>†</sup>Biology Department, Grinnell College, Grinnell, Iowa, USA

---

- Abstract
- I. Introduction
- II. Uses for Transgenes in *C. elegans*
  - A. Analysis of Gene Expression
  - B. Analysis of Protein Localization
  - C. Rescue of a Chromosomal Mutation
  - D. Marking Tissues and Cells for Other Manipulations
  - E. Disruption of Gene Activity: RNA Interference
  - F. Mosaic Analysis of Gene Function
  - G. Marking Extrachromosomal Arrays to Probe Gene Regulation
- III. Construction of Transgenes
- IV. Obtaining Transgenic Animals
  - A. Considerations for Marking of Transgenics
  - B. Delivery Methods
  - C. Integration of Extrachromosomal Arrays
  - D. Transgenics in Other Nematode Species
- V. Perspectives: What Lies on the Horizon?
- VI. Summary
- References

---

---

### Abstract

The ability to manipulate the genome of organisms at will is perhaps the single most useful ability for the study of biological systems. Techniques for the generation of transgenics in the nematode *Caenorhabditis elegans* became available in the late 1980s. Since then, improvements to the original approach have been made to address specific limitations with transgene expression, expand on the repertoire of the types of biological information that transgenes can provide, and begin to develop methods to target transgenes to defined chromosomal locations. Many recent, detailed

protocols have been published, and hence in this chapter, we will review various approaches to making *C. elegans* transgenics, discuss their applications, and consider their relative advantages and disadvantages. Comments will also be made on anticipated future developments and on the application of these methods to other nematodes.

---

---

---

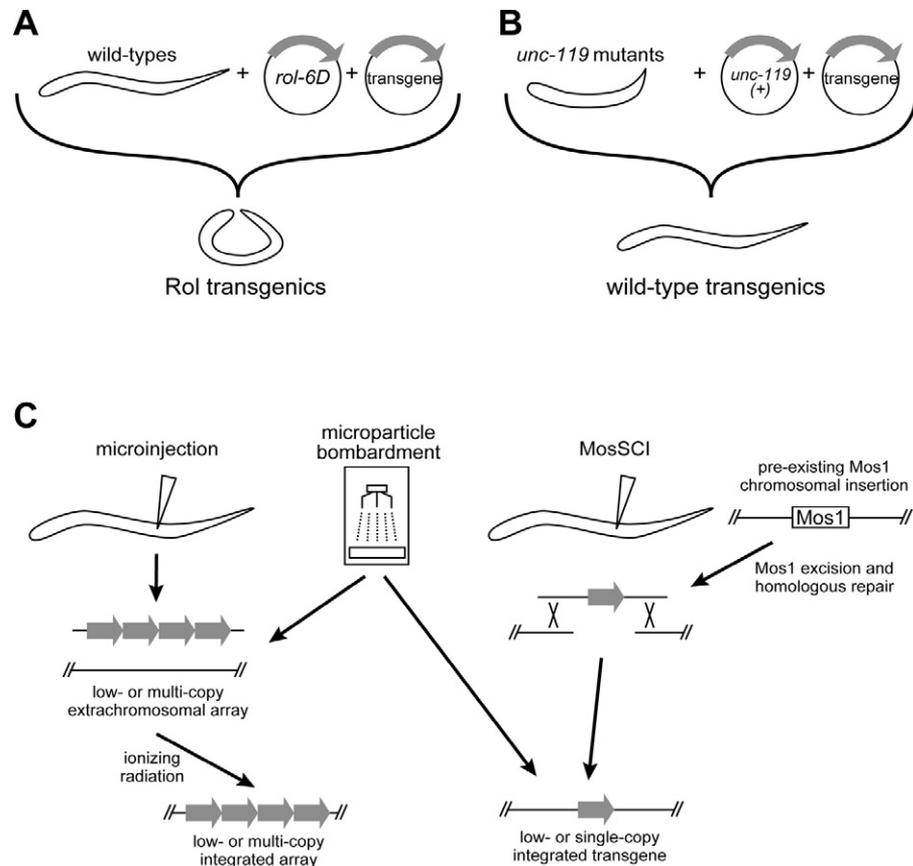
## I. Introduction

Following the generation and characterization of chromosomal mutations (Brenner, 1974), the ability to generate transgenic lines in *C. elegans* (Fire, 1986; Mello *et al.*, 1991; Stinchcomb *et al.*, 1985) opened the system for the rapid genetic characterization of many diverse biological phenomena. The assembly of the genome sequence (*C. elegans* sequencing consortium, 1998) accelerated the rate of gene identification because candidate genes could be identified by first correlating the genetic and physical maps and then transforming easily obtained subclones of the genome into mutants to look for complementation rescue. The advent of green fluorescent protein (GFP) as a reporter seemed destined for this system, as *C. elegans* animals are essentially transparent at all life stages and exhibit little autofluorescence (Chalfie *et al.*, 1994). The discovery of RNA-mediated interference (RNAi) (Fire *et al.*, 1998) expanded further on this set of tools, and the vast majority of work published in the *C. elegans* field uses a combination of all three approaches: genetics, transgenes, and RNAi.

Early approaches for transgenesis in *C. elegans* involved microinjecting DNA into either the hermaphrodite gonad or into unfertilized oocytes for the generation of transgenic animals (Fire, 1986; Mello *et al.*, 1991). In contrast with other systems, *C. elegans* embryos are not used for injection, because it is technically much more challenging and less efficient than gonadal injection, which typically produces many transformed F<sub>1</sub> animals per hermaphrodite. Unlike other systems in which transgenic DNA is generally integrated into chromosomal DNA in single copy (Ringrose, 2009; Ziemienowicz, 2010), *C. elegans* transgenes obtained following microinjection assemble into multicopy extrachromosomal arrays that are transmitted to progeny at 5–95% fidelity (Mello and Fire, 1995). While an extrachromosomal transgene is sufficient or even required for many purposes, arrays can be made to integrate following treatment of a transgene strain with ionizing radiation or chemical mutagenesis (see Evans, 2006).

The ease of producing transgenics in *C. elegans*, and the general reliability of transgene-expression patterns, have permitted rapid characterization of gene expression and often function without the use of *in situ* hybridization or antibodies (Fig. 1). To a first approximation, genes in arrays are expressed similarly to endogenous genes, although the relative expression may be increased due to a higher gene dosage or reduced due to silencing of repetitive sequences (Fire and Waterston, 1989; Kelly *et al.*, 1997; MacMorris *et al.*, 1994; Okkema *et al.*, 1993; Stinchcomb *et al.*, 1985). For many years it was observed that promoters normally active in the germ line fail to

function when present in transgene arrays, whether they are integrated or extrachromosomal. The inclusion of complex DNA in the injection mixtures was found to overcome this problem (Kelly *et al.*, 1997), although such transgene strains require careful maintenance to avoid silencing. Microparticle bombardment, a technique used for many years to make transgenic plant cells (Sanford, 1989), was found to



**Fig. 1** Basic strategies for marking transgenics and delivering DNA to *C. elegans*. (A) Transgenic animals can be marked by an induced gain-of-function phenotype in a wild-type background, such as by the presence of the *rol-6D* allele in the transgene array, or through rescue of a mutant to a wild-type phenotype, as with rescue of *unc-119* mutants (B). (C) Delivery of transgenes is achieved primarily by microinjection, but also by microparticle bombardment and a modification of injection, Mos Single Copy Insertion (MosSCI). Each approach produces a different spectrum of extrachromosomal and/or integrated transgene types. Higher copy number arrays (generated by strategies in A and B) give higher transgene expression, but can undergo silencing (particularly for maternally expressed genes), while lower copy number transgenes (generated by strategies in C) show weaker expression that is less prone to silencing.

be capable of generating transgenes in *C. elegans*, some of which integrate randomly into the genome at low or even single copy (Praitis *et al.*, 2001). These types of transgenes generally overcome the limitations of high-copy arrays and are able to express more efficiently in the germ line.

What has lagged behind in the *C. elegans* field is a robust method for single-copy gene insertions and targeted chromosomal modifications. Such modifications would, by their nature, permit expression of maternal and zygotic genes under the control of endogenous regulatory elements and allow generation of custom-made alleles.

Two general methods to generate homologous recombinants, both of which depend on either microinjection or microparticle bombardment to generate transgenic lines, have been developed in the last few years (Fig. 1C). One approach takes advantage of the excision of a transposable element to create a double-stranded (ds) break in DNA, which can be used to promote gene conversion or direct insertion of transgenic sequences directly into the chromosome (Frokjaer-Jensen *et al.*, 2008; Plasterk and Groenen, 1992; Robert *et al.*, 2009). Genomewide screens that have produced thousands of Tc1 and Mos1 transposon insertion lines have significantly increased the applicability of this approach (Bazopoulou and Tavernarakis, 2009; Boulin and Bessereau, 2007; Duverger *et al.*, 2007; Williams *et al.*, 2005). In a second approach, scaled-up methods for microparticle bombardment have been used to produce integrations targeted at the endogenous locus (Berezikov *et al.*, 2004) and recent work using a positive- and negative-selection strategy promises to dramatically improve the efficiency of this process (Vazquez-Manrique *et al.*, 2010).

Comprehensive protocols for generating transgenics by microinjection and microparticle bombardment are available online, in the *WormMethods* section of *WormBook* and in a variety of other excellent published sources (Evans, 2006; Green *et al.*, 2008; Hope, 1999; Kadandale *et al.*, 2009; Mello and Fire, 1995; Praitis, 2006; Praitis *et al.*, 2001; Rieckher *et al.*, 2009). What follows are brief descriptions of the uses of transgenes in *C. elegans* research, general considerations for constructing transgenes and delivering them to *C. elegans*, an assessment of methods in other nematodes, and a brief discussion of what future developments may lie ahead.

---

---

---

## II. Uses for Transgenes in *C. elegans*

### A. Analysis of Gene Expression

The most frequent use of transgenes in *C. elegans* is for the assessment of endogenous gene-expression patterns of protein-coding genes. The simplest approach for making a transcriptional reporter is to clone the 5' regulatory sequence from a gene of interest, fuse it to a reporter gene whose activity can be easily assayed, and include a 3'UTR, usually that of the *unc-54* gene (Fire *et al.*, 1990). Because of

the transparency of the animal at all life stages, the reporter of choice is GFP or other fluorescent proteins such as the GFP variants YFP and CFP (Miller *et al.*, 1999), or the red fluorescent proteins dsRed and its faster-folding, monomeric variant, mCherry (Shaner *et al.*, 2004).

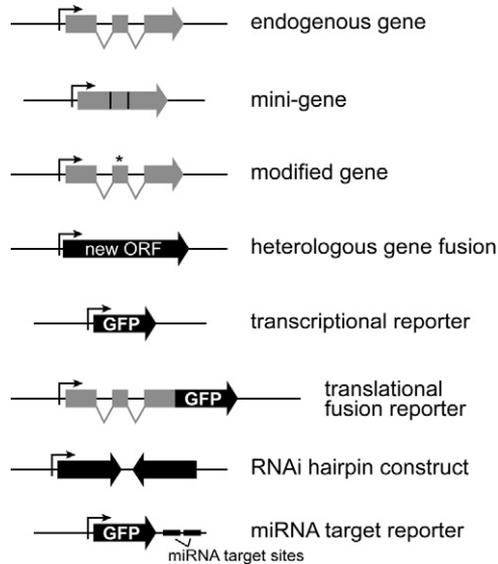
The most important consideration in constructing a transcriptional reporter is the amount of predicted regulatory sequence to include, as this will be the primary determinant of transcriptional regulation of most genes. Because the genome sequence of *C. elegans* is completely known, researchers can examine the physical map around an uncharacterized gene on WormBase and select as large a region as is practical, typically some 3–10 kbp or to the next upstream gene (Dupuy *et al.*, 2004b; Mounsey *et al.*, 1999). Comparison of noncoding sequences in orthologous genes across sequenced genomes has also been helpful in identifying key regulatory elements (Elemento and Tavazoie, 2005; Kuntz *et al.*, 2008). If it is found that the nearest upstream gene is a very short distance (100–400 bp) from the start of the gene of interest, it is possible that the two genes are in an operon (Zorio *et al.*, 1994), in which case the promoter sequences will lie upstream of the most 5' transcript. Regulatory sequences can also be found in introns, so transcriptional fusions may need to include these sequences (Okkema *et al.*, 1993). Common methods of constructing reporters will be described later.

To aid in determining the timing and tissue- or cell-specificity of gene expression, it is useful to include sequences that direct the transgene product to the nucleus. Both the SV40 nuclear-localization signal (NLS) and the coding sequence for a histone have been used to concentrate signals in nuclei to facilitate cell identification (Fire *et al.*, 1990; Strome *et al.*, 2001). The histone tags have the advantage that they stay with chromosomes during mitosis. When combined with other transgenic lines for which expression patterns are well characterized, a precise cell-expression pattern can be determined. The advent of new software combined with four-dimensional imaging using confocal microscopy has made this type of analysis technically simpler and more sophisticated (Murray *et al.*, 2008).

## B. Analysis of Protein Localization

Where the subcellular localization of a protein is being studied, the transgene can be engineered to carry most (or all) of the coding region for the gene of interest, tagged to a reporter construct (Fig. 2). To be assured that function of the protein is not affected by the reporter (Prasher *et al.*, 1992), it is wise to design constructs where the tag is inserted in different positions in the coding sequence. To ensure that the construct is functioning like the endogenous protein, the transgene should be assayed for its ability to rescue the mutant phenotype, if a mutant is available, or for the anticipated behavior following ectopic overexpression.

Finally, because of the possibility that the 3'UTR of a gene might be under control of a micro-RNA (miRNA), inclusion of the gene's native 3'UTR may be required for the construct to reflect the expression of endogenous protein. Predictions of miRNA



**Fig. 2** Schematic showing examples of different types of transgenes (not meant to be exhaustive).

binding sites (Lall *et al.*, 2006) found in WormBase may be informative as to whether or not consideration should be given to possible post-transcriptional regulation in the design of a reporter fusion. Where it is desired to test only the effect of a 3'UTR on gene regulation, sensor transgenes carrying the particular 3'UTR can be tested for responsiveness to miRNA regulation (Wightman *et al.*, 1993). Expression patterns of miRNA genes can be determined with reporter fusions to GFP by using sequences upstream of the mature miRNA as regulatory element (Hayes *et al.*, 2006).

Researchers wishing to know whether the expression of a gene has been studied should first check WormBase (Table I), where expression patterns carried out by gene-specific or genomewide expression studies (Hunt-Newbury *et al.*, 2007) are available. Information on WormBase is often not completely up-to-date and so a literature search should always be performed at the same time. It should be anticipated that a documented expression pattern might not have considered the particular stage, tissue, or condition that is of interest. Hence, the investigator may simply wish to obtain a previously constructed reporter strain, at least for comparison purposes, from either the authors that produced them or from the *Caenorhabditis* Genetics Center (CGC) at the University of Minnesota (Table II). Additional expression information may exist in the form of *in situ* hybridization data published by the Kohara laboratory in Japan (Kohara, 2001), which is accessible in the Nematode Expression Pattern Database (Table I).

**Table I**

Internet links (current as of May, 2011)

Website	Web host	Method/notes
<a href="http://www.wormbase.org">http://www.wormbase.org</a>	WormBase	Contains information about <i>C. elegans</i> genes, including sequences
<a href="http://wiki.wormbase.org/index.php/Cosmids/YACs">http://wiki.wormbase.org/index.php/Cosmids/YACs</a>	WormBase	Information about obtaining <i>C. elegans</i> clones
<a href="http://www.wormbook.org/chapters/www_transformationmicroinjection/transformationmicroinjection.html">http://www.wormbook.org/chapters/www_transformationmicroinjection/transformationmicroinjection.html</a>	WormBook – WormMethods	<i>C. elegans</i> microinjection. Excellent step-by-step instructions on microinjection and microparticle bombardment procedures
<a href="http://www.wormbook.org/chapters/www_reportergenefusions/reportergenefusions.html">http://www.wormbook.org/chapters/www_reportergenefusions/reportergenefusions.html</a>	WormBook – WormMethods	An excellent description of techniques to generate reporter gene fusions
<a href="http://www.wormbook.org/chapters/www_transgenic/transgenic.html">http://www.wormbook.org/chapters/www_transgenic/transgenic.html</a>	WormBook – WormMethods	Considerations for generation of transgenes that express in the germ line
<a href="http://worfdb.dfci.harvard.edu/">http://worfdb.dfci.harvard.edu/</a>	ORFeome	Source of <i>C. elegans</i> ORFs
<a href="http://www.geneservice.co.uk/products/clones/Celegans_Prom.jsp">http://www.geneservice.co.uk/products/clones/Celegans_Prom.jsp</a>	Promoterome	The library contains 6000 predicted promoters, available from Source Bioscience
<a href="http://wormbase.org/db/searches/expr_search">http://wormbase.org/db/searches/expr_search</a>	WormBase	Expression pattern search tool. Can be used to identify promoters active in particular cells or tissues
<a href="http://nematode.lab.nig.ac.jp/db2/index.php">http://nematode.lab.nig.ac.jp/db2/index.php</a>	Nematode Expression Pattern Database (NEXTDB)	Contains <i>in situ</i> expression patterns for a large number of genes
<a href="http://www.cbs.umn.edu/CGC/">http://www.cbs.umn.edu/CGC/</a>	<i>Caenorhabditis</i> Genetics Center at University of Minnesota	Source for many of the strains used in transgenesis experiments
<a href="http://sites.google.com/site/jorgensenmossci/Home">http://sites.google.com/site/jorgensenmossci/Home</a>	Jorgensen Lab, Utah	Mos Single Copy Insertion (MosSCI). Detailed protocol on plasmid construction and screening methods
<a href="http://www.addgene.org">http://www.addgene.org</a>	Addgene	Source for many of the plasmids described here, including for MosSCI
<a href="http://www.faculty.ucr.edu/~mmaduro/int.html">http://www.faculty.ucr.edu/~mmaduro/int.html</a>	Maduro Lab, UC Riverside, CA	Summary of integration techniques using gamma rays, chemical mutagenesis, or UV treatment
<a href="http://www.med.yale.edu/mbb/koelle/protocols/protocol_integrating_array.html">http://www.med.yale.edu/mbb/koelle/protocols/protocol_integrating_array.html</a>	Koelle Lab, Yale School of Medicine, New Haven, CT	Step-by-step integration protocol using gamma rays or X-rays
<a href="http://www.addgene.org/pgvec1?f=c&amp;cmd=showcol&amp;colid=1">http://www.addgene.org/pgvec1?f=c&amp;cmd=showcol&amp;colid=1</a>	Addgene	Links to documentation for Fire Lab plasmids

### C. Rescue of a Chromosomal Mutation

The *C. elegans* system is powerful primarily because of its genetics. For recessive mutations, transgenes carrying the wild-type version of a gene should be able to complement the mutation. For dominant mutations, a transgene carrying the dominant allele should be able to confer a similar phenotype onto an otherwise wild-type strain. Both of these strategies are used for marking transgenes in transformation

experiments. A discussion of markers commonly used for making transgenic lines is included below.

In one variant of complementation of a chromosomal mutation, the transgene is used to rescue an animal with a phenotype resulting from treatment with RNAi. In this case, the transgene is engineered to have resistance to the RNAi effect. When the transgene is combined with tissue-specific or altered promoters, this strategy allows for the assessment of genes that act in multiple tissues or at different stages of development (Green *et al.*, 2008).

Introduction of transgenic sequences is also useful for functional characterization of a gene (Fig. 2). For example, a transgenic construct carrying an altered version of a gene can be assayed to determine if it rescues some or all functions provided by the wild-type gene product (i.e., a sufficiency assay). Introduction of predicted orthologs or paralogs, under the control of a *C. elegans* promoter, can determine conservation of functional domains or gene products. Transgenic constructs containing altered or entirely different regulatory sequences can be used to examine the consequences of ectopic or reduced expression. An altered transgene can be designed to test the function of a particular splice isoform. The transgenes can also be fused to sequences that target them to specific subcellular locations or cause them to be secreted. Several sequences are known that provide subcellular targeting, which includes nuclear localization (SV40 NLS or histone), membrane targeting, secretion, or mitochondrial import (Fire *et al.*, 1990; Portereiko and Mango, 2001; Strome *et al.*, 2001 and Fire Lab Vector information; Addgene). Each of these strategies permits the researcher to manipulate gene activity in order to better characterize the function of a gene of interest.

#### D. Marking Tissues and Cells for Other Manipulations

The rich variety of existing transgenes allows investigators to mark tissues so that they are more easily followed for live cell imaging, to characterize the effect of environmental or genetic manipulation on development of particular cell types or substructures, or to follow cells in a non-natural context, such as after dissociating embryonic blastomeres. A large number of well-characterized promoters can be searched indirectly by expression in particular tissues, stages, or cells on WormBase. The use of reporters in combinations allows the detection of multiple expression patterns in the same animal, an analysis that is particularly useful for determining lineage-specific expression. Reporters of differing absorption/emission spectra can be used, such as the combination of CFP and YFP, or mCherry with GFP. With mCherry and GFP using standard TRITC and FITC filter sets, the two reporter signals show very little overlap. With CFP and YFP, specific filter sets are used to prevent cross-detection (Miller *et al.*, 1999). Signals in strains expressing all four fluorescent proteins can be discerned because of the behavior of each fluorescent protein in each optimal filter set (Table II). This analysis permits both the deeper understanding of mutant phenotypes and the expression patterns of newly characterized gene products.

**Table II**

Cross-detection of popular fluorescent reporters in common filter sets

Fluorophore	Appearance in filter set			
	TRITC (Chroma 31002)*	YFP (Chroma 41029)*	GFP (Omega XF100-2)*	CFP (Chroma 31053)*
mCherry	Red	Faint red	Not visible	Not visible
dsRed	Red	Orange	Faint orange	Not visible
YFP	Faint red	Green	Green	Not visible
GFP	Not visible	Green	Green	Green
CFP	Not visible	Not visible	Green	Cyan

\* Specifications of the various filters can be found on the manufacturer's websites (<http://www.chroma.com>; <http://www.omegafilters.com>).

### E. Disruption of Gene Activity: RNA Interference

Extrachromosomal arrays have been used to generate dsRNA *in vivo* that can elicit RNAi. Because of the ability of RNAi to spread among tissues (systemic RNAi), expression of the RNAi construct does not need to occur throughout the entire animal. The constructs can consist of separate sense and antisense RNA transgenes, or constructs expressing a single hairpin (stem-loop) construct (Fig. 2). Expressing hairpin constructs within neurons has been effective for RNAi knockdown of genes that might be more difficult to achieve by feeding-based RNAi (Johnson *et al.*, 2005; Tavernarakis *et al.*, 2000). One difficulty with hairpin constructs is that the DNA constructs are unstable in *E. coli*; this limitation may be overcome by using stem-loops with introns in the loop portion, or the use of *E. coli* strains that are more tolerant to such structures (e.g., SURE2 cells; Stratagene).

### F. Mosaic Analysis of Gene Function

Researchers often need to test the requirement of a gene within the context of a subset of its normal expression. This may be done to avoid a requirement for the function of a gene in an earlier developmental stage, or to test if gene function is cell-autonomous. Restricted expression of a gene product can be achieved using a variety of techniques, including creating mosaics through loss of extrachromosomal arrays carrying a gene of interest or by fusion of the gene of interest to a specific promoter. A number of other strategies that promise to permit even more sophisticated analysis of tissue-specific gene expression have also been recently developed (Table III).

The classic approach to making mosaic animals in *C. elegans* is to use extrachromosomal arrays as surrogate chromosomal free duplications, which experience mitotic loss within a single animal at a low frequency ( $0.1 \times 10^{-3}$  to  $5 \times 10^{-3}$  loss per cell division) (Hedgecock and Herman, 1995; Lackner *et al.*, 1994; Miller *et al.*, 1996; Yochem and Herman, 2005). Extrachromosomal arrays have an advantage over free duplications because the researcher can determine their composition. To

**Table III**

Mosaic analysis of gene function

Technique	Applications	Considerations	References
Loss of extrachromosomal arrays	Permits expression of gene in a lineage-specific manner. When introduced into a mutant background, permits analysis of lineage-specific presence or loss of gene activity	Marker may not always correlate with gene activity. Expression levels of gene altered due to silencing, overexpression, or perdurance of gene product. Lineage analysis can be difficult and may not be specific enough to limit expression to a small number of tissues	Hedgecock and Herman (1995); Lackner <i>et al.</i> (1994); Miller <i>et al.</i> (1996); Yochem and Herman (2005)
Expression of gene under control of tissue-specific promoters	Permits expression of genes in a very specific set of tissues, cells, or developmental stages	Limited by availability of specific, well-characterized promoters. Use of non-native regulatory elements may produce inappropriate levels of gene product. Requires a new construct for each gene of interest	
Addition of long 3'UTR that alters gene product stability	Permits temperature-sensitive regulation of gene expression	Some background gene expression in the off state. Requires work in a nonsense-mediate decay mutant background	
FLP-recombinase gene activation	Sensitive spatial and temporal control of gene expression. Allows for use of endogenous promoters and other regulatory elements in gene of interest. Creates a set of strains that can be used with different constructs	Change in gene expression due to FLP activation is not reversible. Time delay associated with FLP expression and recombination	Davis <i>et al.</i> (2008); Voutev and Hubbard (2008)
Controlled expression of heat-shock sensitivity	Sensitive spatial and temporal control of gene expression. Relatively rapid changes in gene expression levels. Creates a set of strains that can be used with different constructs.	Need to work in <i>hsf-1(lof)</i> background. Use of non-native regulatory elements may produce inappropriate levels of gene product. Heat-shock response does not allow for sustained gene expression	
Reconstituting gene activity from two components	Sensitive spatial and temporal control of gene expression. Used for cell-specific labeling and killing of cells. Sets of strains can be combined in different ways	Limited to genes or processes that can be reconstituted from two components	Chelur and Chalfie (2007); Zhang <i>et al.</i> (2008)

(Continued)

**Table III** (Continued)

Technique	Applications	Considerations	References
Cell-specific delivery of heat shock	Sensitive spatial and temporal control of gene expression using a focused laser microbeam	Requires laser apparatus and ability to identify cells. Care is required to avoid damaging the induced cell(s)	Stringham <i>et al.</i> (1992)
Temperature-sensitive <i>mec-8</i> -dependent splicing	Permits controlled, temperature-sensitive regulation of gene expression, including RNAi-sensitivity	Perdurance of <i>mec-8</i> activity can make precise regulation difficult. Splicing event requires low-doses of MEC-8. Need to work in <i>mec-8</i> background	Calixto <i>et al.</i> (2010)
Selective depolarization of cells by light stimulation (“optogenetics”)	Activation of transgene-driven light-sensitive proteins such as channelrhodopsin-2 (ChR2) (Nagel <i>et al.</i> , 2003).	Light activation can be delivered broadly, as only cells expressing ChR2 will become depolarized. Light-sensitive channels that respond to different wavelengths can be used simultaneously	Stirman <i>et al.</i> (2011)
Tissue-specific RNAi sensitivity	Permits specific loss of gene activity in a subset of cells	RNAi effectiveness can be variable	Qadota <i>et al.</i> (2007)

identify which cells/tissues have inherited the array, cells carrying the wild-type copy of a gene could be identified by tagging the gene with GFP, by including a ubiquitous reporter such as *sur-5::GFP* (Yochem and Herman, 2005), by including rescue of *ncl-1*, whose function can be scored cell-autonomously (Hedgecock and Herman, 1995), or by using a nuclear-localized GFP::LacI to mark LacO sequences present in the array (Gonzalez-Serricchio and Sternberg, 2006) (discussed below). By referring back to the *C. elegans* lineage (Sulston *et al.*, 1983), the researcher can conclude which cell(s) lost the array in a particular animal, and, if this loss includes tissues of interest, conclusions can be made about cell-autonomous and cell nonautonomous functions. Finally, arrays can be specifically lost in the maternal germ line, so that progeny animals can be produced that lack both maternal and zygotic contributions of the gene (Hunter and Kenyon, 1996). While this approach has been immensely powerful, using array loss to examine tissue-specific gene expression does have limitations, which include the sometimes-complex lineage analysis required to understand emerging phenotypes.

A second technique for examining tissue-specific expression relies on the use of tissue-specific promoters linked to one’s gene of interest (Table III). While this technique has also significantly contributed to our understanding of tissue-specific gene expression, this analysis can be restricted by the limited availability of well-characterized promoters.

A number of strategies for temporally and spatially controlling gene expression have been recently developed (Table IV). In general, these techniques depend on

**Table IV**

Markers used to identify transgenics

Marker	Plasmid	Notes
<i>rol-6(su1006)</i>	pRF4	Confers a dominant right-handed Roller phenotype to animals (Kramer <i>et al.</i> , 1990). Male Rollers do not mate well. Plasmid available from most <i>C. elegans</i> laboratories
<i>unc-119</i> rescue	pDP#MM016B or <i>Cb-unc-119(+)</i> in transgene plasmid	Rescues uncoordinated <i>unc-119</i> mutants to a wild-type phenotype (Maduro and Pilgrim, 1995; Maduro and Pilgrim, 1996). Mutant strain <i>unc-119(ed4)</i> available from the <i>Caenorhabditis</i> Genetics Center (CGC). For use with microparticle bombardment <i>unc-119(+)</i> is usually included in the transgene plasmid. For MosSCI, the <i>C. briggsae</i> gene is inserted into the targeting vector. Plasmid is available from the Maduro lab (University of California, Riverside, CA)
<i>lin-15</i> rescue	pEKL15	Rescues temperature-sensitive multivulva (Muv) phenotype of <i>lin-15(n765)</i> (Clark <i>et al.</i> , 1994). The <i>lin-15(+)</i> plasmid pEKL15 is available from the Horvitz laboratory (Massachusetts Institute of Technology, Cambridge, MA). Strains harboring <i>lin-15(n765)</i> are available from the CGC
<i>pha-1</i> rescue	pBX or pC1	Rescues larval lethality of <i>pha-1(ts)</i> mutants to a wild-type phenotype. Strain is maintained at 15°C, used at 25°C. Mutant strain <i>pha-1(e2123)</i> is available from the CGC (Kramer <i>et al.</i> , 1990)
<i>dpy-20</i> rescue	pMH86	Rescues strong Dumpy phenotype of <i>dpy-20(ts)</i> mutants to a wild-type phenotype. Mutant strain <i>dpy-20(e1282ts)</i> available from the CGC (Clark <i>et al.</i> , 1995). The plasmid is available from the Han lab (University of Colorado, Boulder, CO)
<i>spe-26</i> rescue	pJV145	Rescues <i>spe-26(hc138ts)</i> (H. Smith and S. Ward, personal communication; Praitis, 2006)
Puromycin resistance	pBCN21-R4R3 or pBCN22-R4R3	Plasmids confer resistance to puromycin (Semple <i>et al.</i> , 2010)
G418/neomycin resistance	pdestDD04Neo, pdestRG5271Neo, pdestRG5273Neo	Resistance to G418 (neomycin) (Giordano-Santini <i>et al.</i> , 2010)
<i>myo-2::mCherry</i>	pCFJ90	Expresses mCherry in pharynx muscle. Plasmid available from Addgene
<i>elt-2::NLS::GFP::lacZ</i>	pJM66	Expresses GFP in intestinal nuclei (Fig. 3D). Plasmid available from McGhee Lab (University of Calgary, AB)
<i>sur-5::GFP</i>	pTG96	Expresses GFP in all nuclei. Plasmid available from the Han lab (University of Colorado, Boulder, CO)
<i>let-858::GFP</i>	pBK48.1	Expresses GFP in all nuclei. Plasmid available from Kelly Lab (Emory University, Atlanta, GA)
<i>unc-119::GFP</i>	pDP#MMUGF12	Expresses GFP throughout nervous system and in some head muscles (Fig. 3E). GFP, YFP, CFP, and mCherry versions of this reporter are available from the Maduro lab (University of California, Riverside, CA)

controlling gene structure, gene product stability, or gene activity using either tissue-specific promoters or by controlling the presence of inducers or repressors of gene expression. One technique takes advantage of the well-characterized FLP recombinase system to activate a gene of interest in specific tissues or at specific times of development. In this strategy, the regulatory element is separated from the gene of interest by a sequence that is cleaved upon tissue-specific expression of FLP recombinase, thereby activating the gene (Davis *et al.*, 2008; Voutev and Hubbard, 2008). Until recently, the available methods for delivering transgenes, which tend to be higher copy number and contain rearrangements, limited the effectiveness of this technique for temporal or tissue-specific knockdown (as opposed to activation) of gene activity. A second technique that permits control of gene expression depends on reconstituting gene activity from two gene components, each under the control of a specific promoter. When the two elements are combined, as would be expected in a small number of specific cells, gene activity is restored (Chelur and Chalfie, 2007; Zhang *et al.*, 2008). This technique is limited to genes or processes that can be reconstituted from two components. A third technique requires altering the 3'UTR of a gene and taking advantage of temperature sensitivity of nonsense-mediated decay of RNA products to promote gene stability or decay (Drake *et al.*, 2003). However, the effects on gene-expression levels are not always absolute. A fourth technique depends on rescuing a heat-shock deficient *hsf-1(sy441)* mutant in a cell-specific manner by controlled expression of wild-type *hsf-1*, permitting expression of a heat-shock inducible promoter linked to one's gene of interest in only those cells (Bacaj and Shaham, 2007). This technique is limited by the temperature-sensitivity of the process and by the transient nature of the heat-shock response. A fifth method takes advantage of the observation that MEC-8 is required to splice *mec-2* intron 9, thereby regulating the expression of *mec-2* splice variants. By creating a transgene carrying the *mec-2* intron 9 sequence upstream of a gene of interest in a *mec-8(u218ts)* strain, one can regulate expression of the gene using temperature shifts. This technique was used to control expression of the RNAi gene *rde-1* to create a line with temperature-dependent RNAi (Calixto *et al.*, 2010). Potential limitations of this technique include the need to work in a *mec-8* strain, as well as the relative stability of MEC-8 and dose-sensitivity of the splicing event. Despite some limitations, each of these techniques offers researchers valuable tools for selective expression of their gene of interest.

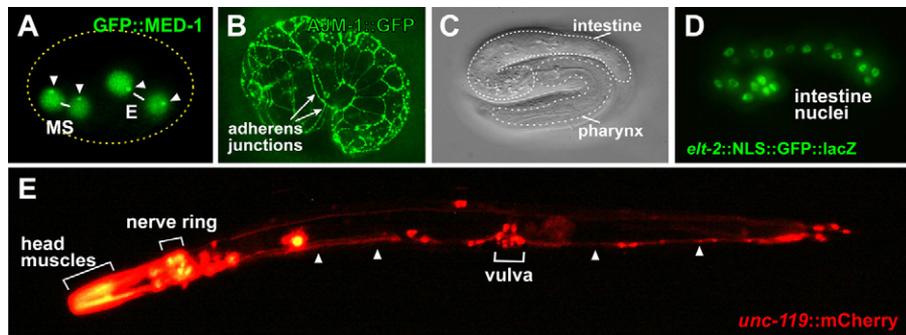
Knockdown of gene function in specific tissues can also be used to examine gene activity. The cell-autonomous requirement for RDE-1 function in RNAi can be exploited by providing wild-type function of *rde-1* in a tissue-specific manner to an *rde-1* mutant strain. Animals then treated with RNAi to a gene of interest will experience knockdown only in cells carrying RDE-1 function (Qadota *et al.*, 2007). While this strategy has been used effectively for some genes, the strength of the RNAi response can be variable.

In summary, the *C. elegans* researcher has a large set of techniques that can be used to understand the role of a gene in a specific cell or developmental process, bypass requirements at specific stages, or examine the consequences of ectopic gene expression. The specific gene studied, the hypothesis being tested, and the

limitations and advantages of each strategy will dictate which technique is best for a given application.

### G. Marking Extrachromosomal Arrays to Probe Gene Regulation

The interaction of the *E. coli* LacI protein with *lacO* lactose operator sequences was exploited as a method for marking chromosomes in yeast (Belmont and Straight, 1998) and has been used as a marker for transgenes in *C. elegans* as well (Gonzalez-Serricchio and Sternberg, 2006). Use of the LacI/LacO systems has also been used to label extrachromosomal arrays to study gene regulation (Fig. 3A). In such experiments, a GFP-tagged endogenous transcription factor is expressed in the presence of an extrachromosomal array that carries a promoter that contains its target *cis*-regulatory sites. The factor will interact with the many copies of the target promoter in the array, producing a subnuclear spot. LacI tagged with a different marker can label *lacO* sequences in the same target array, allowing an independent means by which to verify interaction of the GFP-tagged factor with the array (Carmi *et al.*, 1998). Researchers have also used the GFP::LacI/LacO system to demonstrate that transgenes move to different locations in the nucleus depending on whether they are active or inactive in a given cell or tissue (Meister *et al.*, 2010).



**Fig. 3** Examples of types of transgenes and their expression patterns. (A) Expression of a chromosomally-integrated *med-1::GFP::MED-1* translational reporter in the early embryo, showing nuclear GFP expression in the daughters of the blastomeres MS and E (Maduro *et al.*, 2002). Due to the presence of a separate extrachromosomal array carrying a transcriptional *lacZ* reporter for the MED-1 target gene *end-3*, the GFP::MED-1 localizes to subnuclear spots representing the extrachromosomal array (arrowheads) in each nucleus. (B) Expression of a translational fusion of the adherens junction marker *ajm-1* in mid-embryogenesis. GFP becomes localized to adherens junctions, giving an outline of epidermal cells (Koppen *et al.*, 2001). (C) DIC image of a late embryo, just prior to hatching, with the pharynx and intestine indicated. (D) Expression of an *elt-2::NLS::YFP::lacZ* reporter transgene in the same embryo as in (C) localized to intestinal nuclei (and excluded from nucleoli). (E) A *C. elegans* adult hermaphrodite showing expression of an *unc-119::mCherry* transcriptional reporter throughout the nervous system (including the nerve ring, neurons around the vulva, and the ventral nerve cord indicated by arrowheads) and in head muscles (Maduro and Pilgrim, 1995). The head muscle expression has been overexposed. Anterior is to the left. A *C. elegans* embryo is approximately 50  $\mu\text{m}$  long, while adults are approximately 1mm long. (See color plate.)

---

---

### III. Construction of Transgenes

An excellent description of the many considerations for construction of plasmid reporters can be found in Boulin *et al.* (2006); Mounsey *et al.* (1999), which we have updated here. There do not appear to be any sequence requirements for the stable inheritance of arrays in *C. elegans* (Mello *et al.*, 1991), as DNA from plasmids or phage, for example, appears to be incorporated into arrays. Hence, standard molecular biology techniques can be used to construct most transgenes. When segments of wild-type DNA are needed, polymerase chain reaction (PCR) can be used to amplify directly from genomic DNA, or larger clones, such as a cosmids or fosmids, can be ordered and used for either direct subcloning or PCR. Information on ordering clones is available on WormBase (Table I). Researchers are cautioned that some larger clones are unstable when propagated in bacteria or yeast, such that a particular isolate of a clone could be missing regions of DNA. When working with these constructs it is always advisable to check that sequences have not been lost.

For simple reporter fusions of zygotically expressed genes, it is usually sufficient to clone a suitable upstream promoter fragment (3–10 kbp is a good start, without taking sequences from the neighboring gene upstream) along with a small part of the coding region. The fragment is cloned into one of the available GFP vectors (generated by the laboratory of Andrew Fire). These vectors supply a useful polylinker, synthetic introns to increase expression, and a 3'UTR from the *unc-54* gene. Variants are available that encode other fluorescent proteins (YFP or CFP), include a nuclear localization signal (NLS), or are a fusion to both GFP and lacZ. Other vectors use a histone H2B coding sequence as a more effective means to localize GFP to nuclei. A number of useful vectors as well as additional documentation from the Fire lab can be obtained from Addgene (<http://www.addgene.org>). Where an investigator hypothesizes sequence requirements that necessitate a much larger context for the reporter (e.g., tens of kilobasepairs), manipulations can be performed using recombination in yeast or fosmids (Dolphin and Hope, 2006; Tursun *et al.*, 2009; Zhang *et al.*, 2008).

Other applications of transgenes, such as the fusion of a promoter to a different downstream sequence, will require approaches unique to each application (Fig. 2). Additional resources available to the research community can simplify cloning or allow rapid scaling-up of construct production. For example, it is now possible to use the Gateway recombination cloning system to fuse promoters from the “promoterome” library into a suitable reporter. For making novel fusions of promoters to different coding regions, clones from the promoterome can be combined with clones from the ORFeome (Dupuy *et al.*, 2004a; Hope *et al.*, 2004; Reece-Hoyes *et al.*, 2005).

For expression of heterologous coding regions, it may be cost-efficient for an investigator to order an open reading frame to be synthesized *de novo*. Custom gene synthesis can now be achieved for a relatively low cost per base pair. This would also allow engineering for efficient expression in *C. elegans*, such as by the introduction of short introns, or the selection of codons that are optimized for maximal gene

expression (Duret, 2000; Okkema *et al.*, 1993). Hence, some labs may consider that particular manipulations, such as modification of protein coding regions, might be best achieved by direct synthesis, considering the time and/or number of manipulations that would otherwise be necessary.

Expression of genes in the germ line or early embryo can be less straightforward than expression in other tissues or stages of development. The repetitive nature of conventional transgenes results in germ-line silencing (Kelly *et al.*, 1997). Conventional extrachromosomal arrays are compatible with maternal (germ line) expression for only a small subset of GFP reporters (Fire *et al.*, 2006). There are several approaches for achieving expression of maternal transgenes. One is to use conventional arrays, but to coinject genomic DNA that has been digested with a restriction enzyme that leaves blunt ends (Kelly *et al.*, 1997). This approach appears to achieve expression of maternal transgenes because the arrays are made complex and less prone to silencing. In some instances, maternal expression can be achieved by using a promoter and 3'UTR that seem to be compatible with expression from a multicopy array, such as from *glh-2* (Bessereau *et al.*, 2001). More reliable approaches for germ-line or maternal expression use microparticle bombardment or MosSCI (Fig. 1B), both of which deliver fewer copies of the transgene, which makes them less prone to silencing. Both of these techniques require special consideration for the plasmids that carry the transgenes, as described below.

## ==== IV. Obtaining Transgenic Animals

### A. Considerations for Marking of Transgenics

It is usual practice to mark the presence of a transgene by a convenient marker that can be scored visually in larvae or adults, to facilitate identification following transgene delivery, and when transformants are obtained, during crosses or screening for integrants (Figs. 1A,B, Table IV). Transgenes that confer a readily detectable change in phenotype from nontransgenic animals (e.g., rescue of a visible mutation or very bright GFP reporter), may not need a coinjection marker unless a positive control for the injection process is desired. During the process that gives rise to conventional transgene arrays, recombination among the injected plasmids (if present at high enough relative concentrations) will almost always ensure that multiple, separate plasmids become incorporated into the same array. For microparticle bombardment, the marker is often but not always included in the same plasmid as the transgene, because the low copy number of the resulting insertions makes it less likely that both will become integrated. For Mos-directed chromosomal insertion, the marker and transgene must both be included in between the flanking homology segments.

Simple transgenes in a wild-type background can be marked with the plasmid pRF4, which encodes the *su1006* allele of the *rol-6* gene, also called *rol-6D* (Fig. 1A) (Mello *et al.*, 1991). pRF4 induces an obvious right-handed Roller (Rol) phenotype.

Unfortunately, the effect greatly reduces the mating efficiency of males, which can make crosses more difficult. Some chromosomal integrants of *rol-6D*-marked transgenes show a much weaker Rol phenotype as heterozygotes, or they can be combined with mutations in some genes (e.g., *dpy-11*), which can suppress the Rol effect.

Where a wild-type phenotype is desired from the transgenic animals, it is convenient to start with a strain carrying a recessive mutation and use rescue of the mutation as the transgene marker (Fig. 1B). Markers commonly used are rescue of *pha-1*, *dpy-20*, and *unc-119*. Loss of *pha-1* is lethal, but the allele used is temperature-sensitive (ts), so that animals are propagated at 15°C and selected for transgenics at 25°C (Granato *et al.*, 1994). Loss of *dpy-20* results in a viable dumpy (Dpy) phenotype, but Dpy adults are more difficult to inject, so a ts allele is used (Clark *et al.*, 1995). Until such transgenes are integrated, maintenance requires propagation at 25°C, which may make downstream genetics more challenging (e.g., if a transgene were to be crossed into another ts mutant background). *unc-119* mutants do not form dauer larvae, an alternative larval stage that allows worms to survive prolonged starvation. As a result, non-Unc-119 transgenics can be identified from large populations since they are viable after starvation. For MosSCI or microparticle bombardment, in which a very small fraction of animals becomes transgenic, rescue of *unc-119* has been the most frequently used marker, although a number of other markers have been used successfully (Praitis, 2006). Use of the more compact *C. briggsae* homolog of *unc-119* is convenient as it facilitates cloning of the transgene and marker into the same plasmid. Inclusion of the *unc-119* marker into transgenic constructs has also been made simpler by a recent modification of recombineering techniques (Ferguson and Fisher, 2009; Zhang *et al.*, 2008).

As an alternative to using rescue of a mutation, transgenes can be marked by the presence of a GFP reporter to *myo-2* (Okkema *et al.*, 1993), *elt-2* (Fig. 3D) (Fukushige *et al.*, 1998), *sur-5*, or *let-858* (Kelly *et al.*, 1997; Yochem *et al.*, 1998). Access to a dissecting microscope equipped with a fluorescent lamp and appropriate filters or a fluorescent worm sorter are necessary for identifying and maintaining lines carrying extrachromosomal arrays. Other transgene markers include antisense-*unc-22*, which imposes a twitching paralysis, and selection for resistance to antibiotics (Fire *et al.*, 1991; Giordano-Santini *et al.*, 2010; Semple *et al.*, 2010).

As a final consideration, expression of one gene on a single array may be precluded by the presence of the second gene. In such cases, it may be desirable to obtain separate transgene reporters, and combine the two strains together. If this is done, the researcher may wish to consider different strategies for marking the presence of either transgene. For example, if both are *rol-6D* marked transgenes, it may be difficult to identify double-transgenics or to even mate them together. In such cases, rescue of *unc-119* and *rol-6D* could be used to make separate transgenes, and then the two strains can be combined by crossing rescued *unc-119* transgenics to the *rol-6D* strain that is homozygous for *unc-119*. The double transgenics will be Rol non-Unc.

## B. Delivery Methods

The *C. elegans* germ line is the target organ for microinjection. It contains a syncytium of germ-line nuclei sharing a common cytoplasm (Klass *et al.*, 1976). Researchers have two choices for delivery of transgenes to *C. elegans*, microinjection or microparticle bombardment (Fig. 1C).

Microinjection is typically the first technique tried, as it requires only a small number of animals, and F<sub>1</sub> animals are scored within a few days after injection. An inverted microscope setup with differential interference contrast (DIC) or similar optics, a needle puller, and a micromanipulator are necessary. Pressurized nitrogen, delivered through a regulator with a foot pedal controller, is usually used to force injection mixtures through the needle. Alternatively, other lower-cost arrangements are possible. Laboratories performing *Drosophila* microinjection may have similar setups that can be used. A detailed protocol for *C. elegans* transformation using microinjection can be found online from the WormMethods section of WormBook and from other published sources (Evans, 2006; Kadandale *et al.*, 2009; Mello and Fire, 1995). For laboratories that desire low copy number transgenes, for example, to avoid toxicity or to achieve maternal expression, injection can be modified by the inclusion of digested genomic yeast or nematode DNA (Kelly *et al.*, 1997). Alternatively, bombardment or MosSCI, both of which yield low copy integrants, can be used. When stable lines are required, extrachromosomal transgenic arrays can be integrated using chemical mutagens or radiation, as described below.

Microparticle bombardment requires more time initially, as a large number of starting *unc-119* mutant animals are required, and there is usually a 10–14-day post-treatment wait time before active screening for transformants begins. However, the chief advantages of the technique are that both integrants and extrachromosomal arrays are obtained in the same procedure, the technique relies on a selection that yields only the most stable lines, and it requires little technical expertise. Access to a Biolistic PDS-1000 Helium Microparticle Bombardment machine or other delivery device and several consumables are needed for this procedure. Laboratories may find access to such a machine if there is a nearby facility that performs plant cell transformations. Detailed descriptions of the microparticle bombardment procedure are available in WormBook or in other published sources (Evans, 2006; Green *et al.*, 2008; Jackstadt *et al.*, 1999; Praitis *et al.*, 2001; Rieckher *et al.*, 2009; Wilm *et al.*, 1999).

The delivery of transgenes for directed chromosomal insertion using Mos transposition (Fig. 1C) is really just a special case of direct microinjection into *unc-119* mutant animals, as the desired transgenics do not require infinite passage of the extrachromosomal transgene array. The injected plasmids serve primarily as the chromosomal repair source (the “targeting plasmid”) and to activate Mos transposition in the germ line. To distinguish *bona fide* chromosomal insertions from transmission of an extrachromosomal array, several reporter plasmids are coinjected simultaneously. In one version of the approach, chromosomal insertions are rescued for the *unc-119* phenotype but fail to express the other coinjected transformation

markers, including a plasmid encoding the temperature-sensitive, dominant negative selection marker *twk-18(cn110)*. The plasmids are microinjected simultaneously in a manner similar to that for conventional arrays, but the wait time following microinjection is about 10–14 days. Some 25–50 animals are needed for injection, and the equipment is identical to that needed for normal extrachromosomal arrays (Frokjaer-Jensen *et al.*, 2008). Hence, laboratories that are already established for regular microinjection will find it easier to work with Mos-directed insertions if their goal is to obtain low copy number arrays.

Researchers wishing to try both microparticle bombardment and MosSCI may consider constructing their transgene into a targeting vector for MosSCI, as the resultant plasmid can then be used directly for bombardment, MosSCI or a conventional multicopy array transgene, as all can be delivered to *unc-119* mutants. It is worth noting that as MosSCI insertions are targeted to predetermined locations, researchers may wish to consider which location (and corresponding targeting vector) they will use if there is a later need to combine transgenes into one strain. There are currently two locations, on chromosomes II and IV (Frokjaer-Jensen *et al.*, 2008), though it is anticipated that additional targeting loci will become available over time.

### C. Integration of Extrachromosomal Arrays

Transgenes carried on extrachromosomal arrays can be integrated into a chromosomal location, which eliminates mitotic and meiotic loss of the array (Evans, 2006). Spontaneous integration of extrachromosomal arrays has been observed by many investigators, which may be more likely to be seen in large populations propagated for many generations, especially if there is a selective advantage to the integrants. Otherwise, spontaneous integration is rare enough that it is not convenient to expect it to occur for any given transgene. Hence, most investigators use chemical mutagens or ionizing radiation (gamma rays or ultraviolet light) to induce integration of an array into a chromosome. This is usually done by mutagenizing a small starting population of animals, establishing several hundred single F<sub>1</sub> animals, and testing F<sub>2</sub> progeny individually for 100% transmission of the transgene to subsequent generations (Evans, 2006). Coinjection of oligonucleotides can also stimulate integration of arrays (Mello *et al.*, 1991), and integration is observed if oocyte nuclei are directly injected (Fire, 1986). However, neither of these approaches appears to be in wide use. Once integrated, it is usually no longer necessary to follow a transgene by the coinjection marker. This may simplify subsequent genetic manipulations and permit combining multiple transgenes into a single strain.

For all integrated transgenic lines, strains should be backcrossed several times to eliminate background changes to the genome introduced by the integration treatment. It is also important to examine phenotypes and expression patterns in several integrated lines to be assured that results are not dependent on the site of integration or any linked background mutations.

## D. Transgenics in Other Nematode Species

With the recent availability of genome sequences for other nematodes, researchers may wish to perform gene manipulations in these species. The closely related *C. briggsae* is the most frequent choice for comparative work, perhaps because it is hermaphroditic like *C. elegans* (most others are male–female), and because its genome sequence is of very high quality (Stein *et al.*, 2003). Like *C. elegans*, *C. briggsae* can be made transgenic by microinjection with the use of *rol-6D* or rescue of mutants. Mutants in *unc-119* have been made available that permit the use of microparticle bombardment (Zhao *et al.*, 2010). There are Mos insertions that have been made in *C. briggsae* but the MosSCI approach is still being developed (Marie Delattre, personal communication). In the more distant nematode *Pristionchus pacificus*, also a hermaphroditic species, transgenics can be made, though with some difficulty (Schlager *et al.*, 2009), using an adaptation of the protocol for using complex arrays in *C. elegans* (Kelly *et al.*, 1997). Routine transgenesis in male–female nematode species has not been developed, although in principle coinjection of a dominant marker that does not affect male mating, such as a GFP reporter, could be used to mark transgenics. The *rol-6D* phenotype compromises male mating, which would make maintenance of homozygous transgenic integrants more difficult. As well, the basis for identification of transgenics in bombardment and MosSCI – *unc-119* rescue – would be impossible in male–female species because *unc-119* blocks male mating. Microparticle bombardment may be the best possibility for transgenesis in other species, if a system can be devised to identify rare transgenics. One promising breakthrough, a transformation strategy that depends on conferring drug resistance, will likely make it simpler to generate transgenics in a large number of nematode species (Giordano-Santini *et al.*, 2010; Semple *et al.*, 2010).

---

---

---

## V. Perspectives: What Lies on the Horizon?

While recent research has added to an already rich suite of applications for transgenesis in *C. elegans*, there are some technologies, used in other systems, which are still being developed or refined in worms. One essential technique for studying gene function is a simple, reproducible method for knocking out, tagging, or otherwise manipulating a gene at its endogenous locus. Several recently developed methods for creating lines carrying homologous integrations, using negative/positive selection after microparticle bombardment or MosSCI, promise to make homologous, targeted modifications the standard in *C. elegans* (Frokjaer-Jensen *et al.*, 2008; Vazquez-Manrique *et al.*, 2010).

A second essential technique that would permit more sophisticated analysis of gene function is one that promotes or prevents gene expression in a precise spatial or temporal pattern, similar to the *Drosophila* GAL4 system. In principle, several recently developed techniques described in the mosaic analysis section of this chapter could achieve this aim (Table III). Particularly promising are those techniques that create strains that confer specific expression or inhibition of any

appropriately constructed transgene or RNAi construct. Once a toolbox of these strains has been created, it could be used by anyone in the community who requires a specific expression pattern for their gene of interest. Another exciting possibility is that the FLP recombinase system, which has been used to induce gene expression (Davis *et al.*, 2008; Voutev and Hubbard, 2008), could be used for tissue-specific elimination of gene expression. This seems a likely future development, given that single-gene insertions and homologous recombination techniques have been developed (Frokjaer-Jensen *et al.*, 2008; Vazquez-Manrique *et al.*, 2010).

A variation on mosaic analysis that would be useful for researchers is a technique that permits tissue- or spatially restricted knockdown of gene activity for use in a genetic screen. This technique is crucial for any researcher studying a process that requires genes essential at earlier developmental time points or for other processes. Emerging techniques that restrict RNAi sensitivity to a small set of cells or to specific developmental stages could, in theory, permit these types of genomewide screens (Calixto *et al.*, 2010; Qadota *et al.*, 2007).

Another technology emerging in other systems that may have applications in *C. elegans* is the use of zinc finger nucleases. These are heterodimers consisting of engineered C<sub>2</sub>H<sub>2</sub> zinc finger arrays expressed as fusions to the nuclease domain of the restriction enzyme *FokI* (Kim *et al.*, 2010). These enzymes are capable of single site-specific cleavage of DNA, which in theory can result, as with transposon excision, in imprecise repair of the break (i.e., generating a mutation) or incorporation of sequences from a transgene repair template. At least one such enzyme has been found to function somatically in *C. elegans* (Carroll *et al.*, 2008; Morton *et al.*, 2006), raising the hope that if germ-line expression can be achieved, it may be possible to cause germ-line site-specific chromosome modification.

---

---

---

## VI. Summary

The creation of transgenic strains is one of the most important tools for analysis of gene function. Two different delivery methods for *C. elegans* transgenesis, microinjection and microparticle bombardment, have been developed. From these basic methods, a plethora of techniques have emerged that permit analysis of gene expression and function in a range of key cellular and developmental processes. The future holds promise for even greater precision and sophistication in experimentally manipulating gene expression in *C. elegans*.

## References

- Bacaj, T., and Shaham, S. (2007). Temporal control of cell-specific transgene expression in *Caenorhabditis elegans*. *Genetics* **176**, 2651–2655.
- Bazopoulou, D., and Tavernarakis, N. (2009). The NemaGENETAG initiative: Large scale transposon insertion gene-tagging in *Caenorhabditis elegans*. *Genetica* **137**, 39–46.

- Belmont, A., and Straight, A. (1998). *In vivo* visualization of chromosomes using lac operator–repressor binding. *Trends Cell Biol.* **8**, 121–124.
- Berezikov, E., *et al.* (2004). Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40.
- Bessereau, J. L., *et al.* (2001). Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74.
- Boulin, T., and Bessereau, J. (2007). Mos1-mediated insertional mutagenesis in *Caenorhabditis elegans*. *Nat. Protoc.* **2**, 1276–1287.
- Boulin, T., *et al.* (2006). Reporter Gene Fusions. In (The *C. elegans* Research Community, ed.). WormBook. doi/10.1895/wormbook.1.108.1.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- Calixto, A., *et al.* (2010). Conditional gene expression and RNAi using MEC-8-dependent splicing in *C. elegans*. *Nat. Methods* .
- Carmi, I., *et al.* (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. *Nature* **396**, 168–173.
- Carroll, D., *et al.* (2008). Gene targeting in *Drosophila* and *Caenorhabditis elegans* with zinc-finger nucleases. *Methods Mol. Biol.* **435**, 63–77.
- C. elegans Genome Sequencing Consortium(1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2028.
- Chalfie, M., *et al.* (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Chelur, D. S., and Chalfie, M. (2007). Targeted cell killing by reconstituted caspases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2283–2288.
- Clark, D. V., *et al.* (1995). Molecular cloning and characterization of the *dpy-20* gene of *Caenorhabditis elegans*. *Molec. Gen. Genet.* **247**, 367–378.
- Clark, SG., *et al.* (1994). The *Caenorhabditis elegans* locus *lin-15*, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. *Genetics* **137**, 987–997.
- Davis, M., *et al.* (2008). Gene activation using FLP recombinase in *C. elegans*. *PLoS Genet.* **4**, e1000028.
- Dolphin, C., and Hope, I. (2006). *Caenorhabditis elegans* reporter fusion genes generated by seamless modification of large genomic DNA clones. *Nucleic Acids Res.* **34**, e72.
- Drake, J., *et al.* (2003). Oxidative stress precedes fibrillar deposition of Alzheimer’s disease amyloid beta-peptide (1-42) in a transgenic *Caenorhabditis elegans* model. *Neurobiol. Aging* **24**, 415–420.
- Dupuy, D., *et al.* (2004 a). A first version of the *Caenorhabditis elegans* promoterome. *Genome Res.* **14**, 2169–2175.
- Dupuy, D., *et al.* (2004 b). A first version of the *Caenorhabditis elegans* promoterome. *Genome Res.* **14**, 2169–2175.
- Duret, L. (2000). tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet.* **16**, 287–289.
- Duverger, Y., *et al.* (2007). A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res.* **35**, e11.
- Elemento, O., and Tavazoie, S. (2005). Fast and systematic genome-wide discovery of conserved regulatory elements using a non-alignment based approach. *Genome Biol.* **6**, R18.
- Evans, T. C. (2006) Transformation and Microinjection. In (The *C. elegans* Research Community, ed.). Wormbook, Vol. doi/10.1895/wormbook.1.108.1.
- Ferguson, A., and Fisher, A. (2009). Retrofitting ampicillin resistant vectors by recombination for use in generating *C. elegans* transgenic animals by bombardment. *Plasmid* **62**, 140–145.
- Fire, A. (1986). Integrative transformation of *Caenorhabditis elegans*. *EMBO J.* **5**, 2673–2680.
- Fire, A., *et al.* (1991). Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. *Development* **113**, 503–514.
- Fire, A., *et al.* (2006). Unusual DNA structures associated with germline genetic activity in *Caenorhabditis elegans*. *Genetics* **173**, 1259–1273.
- Fire, A., *et al.* (1990). A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**, 189–198.

- Fire, A., and Waterston, R. H. (1989). Proper expression of myosin genes in transgenic nematodes. *EMBO J.* **8**, 3419–3428.
- Fire, A., *et al.* (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- Frokjaer-Jensen, C., *et al.* (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383.
- Fukushige, T., *et al.* (1998). The GATA-factor *elt-2* is essential for formation of the *Caenorhabditis elegans* intestine. *Dev. Biol.* **198**, 286–302.
- Giordano-Santini, R., *et al.* (2010). An antibiotic selection marker for nematode transgenesis. *Nat. Methods* **7**, 721–723.
- Gonzalez-Serricchio, A., and Sternberg, P. (2006). Visualization of *C. elegans* transgenic arrays by GFP. *BMC Genet.* **7**, 36.
- Granato, M., *et al.* (1994). *pha-1*, a selectable marker for gene-transfer in *C. elegans*. *Nucl. Acids Res.* **22**, 1762–1763.
- Green, R., *et al.* (2008). Expression and imaging of fluorescent proteins in the *C. elegans* gonad and early embryo. *Methods Cell Biol.* **85**, 179–218.
- Hayes, G. D., *et al.* (2006). The *mir-84* and *let-7* paralogous microRNA genes of *Caenorhabditis elegans* direct the cessation of molting via the conserved nuclear hormone receptors NHR-23 and NHR-25. *Development* **133**, 4631–4641.
- Hedgecock, E. M., and Herman, R. K. (1995). The *ncl-1* gene and genetic mosaics of *Caenorhabditis elegans*. *Genetics* **141**, 989–1006.
- Hope, I. (1999). *C. elegans: A Practical Approach*. Oxford University Press, Oxford, U.K.
- Hope, I., *et al.* (2004). Feasibility of genome-scale construction of promoter::reporter gene fusions for expression in *Caenorhabditis elegans* using a multisite gateway recombination system. *Genome Res.* **14**, 2070–2075.
- Hunt-Newbury, R., *et al.* (2007). High-throughput *in vivo* analysis of gene expression in *Caenorhabditis elegans*. *PLoS Biol.* **5**, e237.
- Hunter, C., and Kenyon, C. (1996). Spatial and temporal controls target *pal-1* blastomere-specification activity to a single blastomere lineage in *C. elegans* embryos. *Cell* **87**, 217–226.
- Jackstadt, P., *et al.* (1999). Transformation of nematodes via ballistic DNA transfer. *Mol. Biochem. Parasitol.* **103**, 261–266.
- Johnson, N., *et al.* (2005). Heritable and inducible gene knockdown in *C. elegans* using Wormgate and the ORFeome. *Gene* **359**, 26–34.
- Kadandale, P., *et al.* (2009). Germline transformation of *Caenorhabditis elegans* by injection. *Methods Mol. Biol.* **518**, 123–133.
- Kelly, W. G., *et al.* (1997). Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238.
- Kim, J. S., *et al.* (2010). Genome editing with modularly assembled zinc-finger nucleases. *Nat. Methods* **7**, 91 author reply 91-2.
- Klass, M., *et al.* (1976). Development of the male reproductive system and sexual transformation in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **52**, 1–18.
- Kohara, Y. (2001). [Systematic analysis of gene expression of the *C. elegans* genome]. *Tanpakushitsu Kakusan Koso* **46**, 2425–2431.
- Koppen, M., *et al.* (2001). Cooperative regulation of *AJM-1* controls junctional integrity in *Caenorhabditis elegans* epithelia. *Nat. Cell Biol.* **3**, 983–991.
- Kramer, J. M., *et al.* (1990). The *Caenorhabditis elegans* *rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. *Molec. Cell Biol.* **10**, 2081–2089.
- Kuntz, S., *et al.* (2008). Multigenome DNA sequence conservation identifies Hox *cis*-regulatory elements. *Genome Res.* **18**, 1955–1968.
- Lackner, M. R., *et al.* (1994). A MAP kinase homolog, *mpk-1*, is involved in ras-mediated induction of vulval cell fates in *Caenorhabditis elegans*. *Genes Dev.* **8**, 160–173.

- Lall, S., *et al.* (2006). A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr. Biol.* **16**, 460–471.
- MacMorris, M., *et al.* (1994). Analysis of the VPE sequences in the *Caenorhabditis elegans* *svit-2* promoter with extrachromosomal tandem array-containing transgenic strains. *Mol. Cell Biol.* **14**, 484–491.
- Maduro, M., and Pilgrim, D. (1995). Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**, 977–988.
- Maduro, M., and Pilgrim, D. (1996). Conservation of function and expression of *unc-119* from two *Caenorhabditis* species despite divergence of non-coding DNA. *Gene* **183**, 77–85.
- Maduro, M. F., *et al.* (2002). Dynamics of a developmental switch: Recursive intracellular and intranuclear redistribution of *Caenorhabditis elegans* POP-1 parallels Wnt-inhibited transcriptional repression. *Dev. Biol.* **248**, 128–142.
- Meister, P., *et al.* (2010). The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes Dev.* **24**, 766–782.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Mello, C. C., *et al.* (1991). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Miller, D. M., *et al.* (1999). Two-color GFP expression system for *C. elegans*. *Biotechniques* **26**, 914–918 920-1.
- Miller, L. M., *et al.* (1996). Mosaic analysis using a *ncl-1* (+) extrachromosomal array reveals that *lin-31* acts in the Pn.p cells during *Caenorhabditis elegans* vulval development. *Genetics* **143**, 1181–1191.
- Morton, J., *et al.* (2006). Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16370–16375.
- Mounsey, A., *et al.* (1999). Gene expression patterns. In “*C. elegans*: A Practical Approach,” (and I. A. Hope, ed.), pp. 181–199. Oxford University Press, Oxford.
- Murray, J. I., *et al.* (2008). Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat. Methods* **5**, 703–709.
- Nagel, G., *et al.* (2003). Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13940–13945.
- Okkema, P., *et al.* (1993). Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* **135**, 385–404.
- Plasterk, R. H., and Groenen, J. T. (1992). Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. *EMBO J.* **11**, 287–290.
- Portereiko, M. F., and Mango, S. E. (2001). Early morphogenesis of the *Caenorhabditis elegans* pharynx. *Dev. Biol.* **233**, 482–494.
- Praitis, V. (2006). Creation of transgenic lines using microparticle bombardment methods. *Methods Mol. Biol.* **351**, 93–107.
- Praitis, V., *et al.* (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Prasher, D. C., *et al.* (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229–233.
- Qadota, H., *et al.* (2007). Establishment of a tissue-specific RNAi system in *C. elegans*. *Gene* **400**, 166–173.
- Reece-Hoyes, J., *et al.* (2005). A compendium of *Caenorhabditis elegans* regulatory transcription factors: A resource for mapping transcription regulatory networks. *Genome Biol.* **6**, R110.
- Rieckher, M., *et al.* (2009). Transgenesis in *Caenorhabditis elegans*. *Methods Mol. Biol.* **561**, 21–39.
- Ringrose, L. (2009). Transgenesis in *Drosophila melanogaster*. *Methods Mol. Biol.* **561**, 3–19.
- Robert, V., *et al.* (2009). Mos1 transposition as a tool to engineer the *Caenorhabditis elegans* genome by homologous recombination. *Methods* **49**, 263–269.
- Sanford, J. (1989). Biolistic plant transformation. *Physiologia Plantarum* **79**, 206–209.
- Schlager, B., *et al.* (2009). Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode *Pristionchus pacificus*. *Genesis* **47**, 300–304.

- Semple, J. I., *et al.* (2010). Rapid selection of transgenic *C. elegans* using antibiotic resistance. *Nat. Methods* **7**, 725–727.
- Shaner, N. C., *et al.* (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotech.* **22**, 1567–1572.
- Stein, L. D., *et al.* (2003). The genome sequence of *Caenorhabditis briggsae*: A Platform for Comparative Genomics. *PLoS Biol.* **1**, E45.
- Stinchcomb, D. T., *et al.* (1985). Extrachromosomal DNA Transformation of *Caenorhabditis elegans*. *Molec. Cell Biol.* **5**, 3484–3496.
- Stirman, JN., *et al.* (2011). Real-time multimodal optical control of neurons and muscles in freely behaving *Caenorhabditis elegans*. *Nat. Methods* **8**, 153–158.
- Stringham, E. G., *et al.* (1992). Temporal and spatial expression patterns of the small heat shock (hsp16) genes in transgenic *Caenorhabditis elegans*. *Mol. Biol. Cell* **3**, 221–233.
- Strome, S., *et al.* (2001). Spindle dynamics and the role of gamma-tubulin in early *Caenorhabditis elegans* embryos. *Mol. Biol. Cell* **12**, 1751–1764.
- Sulston, J. E., *et al.* (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
- Tavernarakis, N., *et al.* (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat. Genet.* **24**, 180–183.
- Tursun, B., *et al.* (2009). A toolkit and robust pipeline for the generation of fosmid-based reporter genes in *C. elegans*. *PLoS One* **4**, e4625.
- Vazquez-Manrique, R., *et al.* (2010). Improved gene targeting in *C. elegans* using counter-selection and Flp-mediated marker excision. *Genomics* **95**, 37–46.
- Voutev, R., and Hubbard, E. (2008). A “FLP-Out” system for controlled gene expression in *Caenorhabditis elegans*. *Genetics* **180**, 103–119.
- Wightman, B., *et al.* (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855–862.
- Williams, D., *et al.* (2005). Characterization of *Mos1*-mediated mutagenesis in *Caenorhabditis elegans*: A method for the rapid identification of mutated genes. *Genetics* **169**, 1779–1785.
- Wilm, T., *et al.* (1999). Ballistic transformation of *Caenorhabditis elegans*. *Gene* **229**, 31–35.
- Yochem, J., *et al.* (1998). A new marker for mosaic analysis in *Caenorhabditis elegans* indicates a fusion between *hyp6* and *hyp7*, two major components of the hypodermis. *Genetics* **149**, 1323–1334.
- Yochem, J., and Herman, R. K. (2005). Genetic Mosaics. In (T. C. e. R. Community, ed.). WormBook, Vol. doi/10.1895/wormbook.1.58.1.
- Zhang, Y., *et al.* (2008). A simplified, robust, and streamlined procedure for the production of *C. elegans* transgenes via recombineering. *BMC Dev. Biol.* **8**, 119.
- Zhao, Z., *et al.* (2010). New tools for investigating the comparative biology of *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Genetics* **853**–863.
- Ziemienowicz, A. (2010). Plant transgenesis. *Methods Mol. Biol.* **631**, 253–268.
- Zorio, D. A., *et al.* (1994). Operons as a common form of chromosomal organization in *C. elegans*. *Nature* **372**, 270–272.

---

---

## CHAPTER 7

# RNA Processing in *C. elegans*

**J. Jason Morton and Thomas Blumenthal**

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado, USA

---

Abstract

- I. Introduction
- II. RNA Processing in Other Organisms
  - A. Eukaryotic Pre-mRNA Processing
  - B. Operons
- III. RNA Processing in *C. elegans*
  - A. Intron Splicing
  - B. Trans-Splicing
  - C. Operons
  - D. 3' End Formation

Acknowledgments

References

---

---

---

### Abstract

In *Caenorhabditis elegans*, newly transcribed RNA is processed in several novel ways. Although introns are removed by a canonical spliceosome, they have evolved several specialized features that reflect the differences in the way they are recognized and the way they are spliced. *C. elegans* introns are unusually short, in part because they have no specific branch-point sequences and contain minimal polypyrimidine tracts. Instead, their 3' splice site is characterized by a highly conserved consensus sequence, which alone may be sufficient to position all spliceosomal elements at the 3' end of the intron. Many RNA molecules are also trans-spliced: a capped 22 nt RNA leader is donated by one of a family of specialized snRNPs and spliced to an unpaired 3' splice site, usually just upstream of the start codon. The RNA upstream of this splice site, the outtron, is removed during trans-splicing and presumably degraded, making the identification of the transcriptional start site problematic. Transcripts from approximately 70% of all genes are trans-spliced. Trans-splicing has enabled the evolution of operons – multigene clusters in which a single upstream promoter drives the transcription of a polycistronic pre-mRNA. The *C. elegans*

genome contains more than 1000 such operons. The polycistronic pre-mRNA is processed into individual gene-encoding mRNAs by coordinated upstream 3' end formation and downstream trans-splicing. An intercistronic RNA sequence, the Ur element, plays a key role in specifying downstream trans-splicing.

---

---

---

## I. Introduction

During and after transcription, pre-mRNA is processed into mature mRNA. In *C. elegans*, several different types of canonical RNA processing occur. Shortly after transcription begins, the nascent RNA is capped with a 7-methylguanosine nucleotide. As transcription proceeds, introns are recognized and removed from the pre-mRNA by spliceosomes, which contain the snRNAs and associated splicing factors found in other metazoans. Cleavage at the 3' end of the RNA is executed by orthologs of the mammalian CPSF and CstF complexes.

Additionally, novel RNA-processing events in *Caenorhabditis elegans* result from gene structures and chromosomal gene arrangements not found in other model organisms. Many nascent RNA molecules are trans-spliced: a capped 22 nt RNA leader sequence is spliced to an acceptor site near the 5' end of the pre-mRNA. Trans-splicing has permitted the evolution of operons throughout the nematode genome, since a single pre-mRNA encoding multiple genes can be processed into single-gene units by coordinated 3' end formation and trans-splicing. Downstream genes in an operon are uniquely identified by a special family of spliced-leader sequences. This chapter will examine each of these processing activities with emphasis on aspects of gene expression that are special to *C. elegans* and other nematodes.

---

---

---

## II. RNA Processing in Other Organisms

Many of the RNA features and processing events studied in *C. elegans* have also been examined in other organisms. Most eukaryote RNA contains introns, and their removal is conducted by a conserved cellular machinery that has been extensively studied. Likewise, the basic mechanism of 3' end formation is highly conserved among the eukaryotes in which it has been studied. This section reviews what is known about these processing events in these other systems. The subsequent section describes the differences in common processing events and the unique events observed in *C. elegans* and other nematodes.

### A. Eukaryotic Pre-mRNA Processing

As pre-mRNA is transcribed, its introns are identified and removed by a large, dynamic RNP complex, the spliceosome. This complex is composed of five core small nuclear RNAs (snRNAs) – U1, U2, U4, U5, and U6 – as well as hundreds of

accessory proteins (reviewed in Rino and Carmo-Fonseca, 2009; Soller, 2006). Intron splicing is a process that is highly conserved among metazoans. Intron borders and internal features are identified by short, partially degenerate consensus sequences, to which individual components of the spliceosome have affinity. Coordinated binding by multiple factors results in intron marking and subsequent spliceosome formation. The vertebrate 5' splice site is identified by the consensus sequence AG/GURAGU, while the 3' splice site is characterized by the consensus YAG/N (Soller, 2006; Wahl *et al.*, 2009) (Fig. 1a). An internal binding site, the branch point, with a consensus sequence YURAY, is located somewhere within the final third of the intron. Finally, a polypyrimidine tract extends 10–12 nt upstream from the 3' splice site. While these splice site consensus sequences are conserved in the fission yeast *Schizosaccharomyces pombe*, they are slightly different in the budding yeast *Saccharomyces cerevisiae* (Fig. 1b). In this organism, the 5' splice site consensus sequence is NN/GUAUGU, and the branch-point consensus sequence is UACUAAC. There is usually no polypyrimidine tract, although the sequence immediately upstream of the 3' YAG/N consensus splice site is often pyrimidine-rich (Kuhn and Käufer, 2003). In all organisms, introns containing these splice-site consensus sequences are called GT–AG introns.

The splicing reaction happens in multiple steps: First, the U1 snRNP binds to the 5' splice site of the intron via base-pairing between the U1 snRNP sequence 3'-UC/CAUUCA-5' and the 5' splice-site consensus 5'-AG/GURAGU-3' (Aebi *et al.*, 1987; Séraphin *et al.*, 1988). The two subunits of the U2 snRNP auxiliary factor (U2AF), U2AF<sup>65</sup> and U2AF<sup>35</sup>, bind the polypyrimidine tract and the AG nucleotides of the 3' splice site, respectively (Merendino *et al.*, 1999; Wu *et al.*, 1999; Zorio and Blumenthal, 1999a). The SF1/BBP protein binds the branch point of the intron. Next, the U2 snRNP can base pair with this same sequence, displacing SF1/BBP from its binding site. Finally, the U1 and U2 snRNPs recruit the U4/U6•U5 tri-snRNP, whose arrival stimulates additional rearrangements and the formation of the catalytically active spliceosome (Wahl *et al.*, 2009).

Splicing proceeds via two transesterification reactions. In the first step, the spliceosome orients the intron so that the 2' hydroxyl of the branch-point adenosine can attack the phosphodiester bond at the 5' splice site, producing an upstream exon with a free 3' hydroxyl group and a branch-point lariat intron attached to the downstream exon. In the next step of splicing, the spliceosome positions the free

	<u>5' splice site</u>	<u>branch point</u>	<u>polypyrimidine tract</u>	<u>3' splice site</u>
a. Vertebrate	AG/GURAGU	YURAY	YYYYYYYYYYYYYAG/N	
b. <i>S. cerevisiae</i>	NN/GUAUGU	UACUAAC		YAG/N
c. <i>C. elegans</i>	AG/GURAGU	none		UUUUCAG/R

**Fig. 1** A comparison of the consensus sequences found in a typical (a) vertebrate intron with those found in (b) *S. cerevisiae* and (c) *C. elegans*. The 5' and 3' splice sites are demarcated by slashes (/).

3' hydroxyl group of the upstream exon so it can attack the phosphodiester bond at the 3' splice site, linking the exons together and releasing the lariat intron for debranching and degradation (Wachtel and Manley, 2009).

The process of mRNA 3' end formation is highly conserved among eukaryotes. Several groups of proteins are involved, and research into this process, along with the related process of transcription termination, is rapidly progressing. There are scores of proteins involved in 3' cleavage and transcription termination. Of these proteins, two different complexes appear to play crucially important roles and have been studied most thoroughly, usually in mammalian systems: CPSF (cleavage and polyadenylation specificity factor) and CstF (cleavage stimulation factor). In yeast, the protein constituents of these complexes associate somewhat differently, and the complexes are CF1A and CPF (reviewed in Mandel *et al.*, 2008).

In vertebrates, The CPSF complex catalyzes cleavage of the RNA at a site ~30 nt downstream of the AAUAAA polyadenylation signal to which it binds (Fitzgerald and Shenk, 1981). CPSF contains five proteins: CPSF-160, CPSF-100, CPSF-73, CPSF-30, and Fip1 (Kaufmann *et al.*, 2004; Mandel *et al.*, 2008). CPSF-160 recognizes and binds directly to the AAUAAA. It also may link the CPSF complex to the CstF complex. CPSF-73 is the endonuclease that cleaves the RNA (Mandel *et al.*, 2006).

The CstF complex contains three polypeptides: CstF-77, CstF-64, and CstF-50 (Mandel *et al.*, 2008). CstF-77 acts as a scaffold around which the complex assembles (Legrand *et al.*, 2007). It also interacts with CPSF-160, effectively linking the CPSF and CstF complexes. The CstF-64 binds to a less-well-defined GU-rich region close to, but downstream of, the cleavage site on the RNA. The CstF-50 subunit binds to the C-terminal domain of the largest subunit of RNA polymerase, thereby physically linking the 3' end formation and transcriptional machineries (McCracken *et al.*, 1997).

Evidence from ChIP experiments indicates that the complexes travel along the transcribed DNA with RNA polymerase (Glover-Cutter *et al.*, 2008). They scan the nascent RNA for their recognition sequences, and, when both AAUAAA and a downstream GU-rich sequence are found in the proper orientation, 3' end formation can occur. Some variability is tolerated at the polyadenylation signal. A survey in humans and mice showed that 70% of all polyadenylation signals contained the canonical sequence AAUAAA. The only variant commonly seen was AUUAAA (15%), indicating that there is little plasticity in CPSF's ability to recognize this 3' processing signal in mammals (Mandel *et al.*, 2008).

## B. Operons

An operon is a cluster of genes transcribed together from a common upstream promoter. In the operons of bacteria and archaea (Brown *et al.*, 1989; Jacob and Monod, 1961), genes encoding enzymes in the same metabolic pathway are often located next to each other on the chromosome. A common upstream

promoter deploys RNA polymerase to transcribe all of these genes at once, making a polycistronic mRNA that is translated by repeated translation termination at the 3' ends of upstream coding sequences followed by reinitiation by ribosomes at the 5' start site of downstream coding sequences (McCarthy and Gualerzi, 1990).

In eukaryotes there are several types of gene clusters that resemble operons. First, in several species of yeast, genes involved in similar functions can sometimes be found in close proximity (Wong and Wolfe, 2005). Although the genes in this arrangement, called a metabolic gene cluster, do not contain a common promoter, their organization resembles that of an operon in other respects (reviewed in Osbourn and Field, 2009). Second, dicistronic gene clusters resembling short operons have been reported in plants, flies, and even mammals (reviewed in Blumenthal, 1998, 2004). Intron splicing of the pre-mRNA and 3' end formation of the downstream transcript occur normally, and the dicistronic transcript is exported for translation, which may occur via internal ribosomal entry sites between the coding sequences. Often these genes encode metabolically related products. This latter type of polycistronic cluster appears to be present, at least occasionally, in *C. elegans* (e.g., *tin-9.2* and *exos-4.1*), but it has not been studied there. Multigene transcription units of a third type are prevalent in trypanosomes (Muhich and Boothroyd, 1988), although these have not been termed operons because there is no evidence for their transcriptional regulation. Indeed, often these transcriptional units extend the length of an entire chromosome (reviewed in Clayton, 2002). The vast majority of *C. elegans* operons (to be discussed in a later section) are processed like those of trypanosomes, by a concerted process of 3' end formation and trans-splicing just downstream. Many animal phyla, including most or all nematodes, have operons of this type.

---

---

---

### III. RNA Processing in *C. elegans*

#### A. Intron Splicing

##### 1. *C. elegans* Introns

The splicing machinery and the process of intron removal have been highly conserved in *C. elegans*. Most *C. elegans* genes have multiple introns, defined by the core GT-AG splice sites (The *C. elegans* Sequencing Consortium, 1998). Intron splicing occurs cotranscriptionally and is directed by the well-characterized spliceosomal snRNAs U1, U2, U4, U5, and U6, along with their associated proteins (Thomas *et al.*, 1990). In many species including plants, flies, and vertebrates, there exists a related, but distinct, group of introns, the removal of which is catalyzed by a spliceosome, called the U12 type, with some different components (reviewed in Will and Lührmann, 2005). However, no U12-type introns or minor spliceosomal component genes are present within the *C. elegans* genome, and it appears that all introns have evolved to undergo splicing by the U2-type major spliceosome (Burge *et al.*, 1998; Sheth *et al.*, 2006).

## 2. Noncanonical Splice Sites

As is the case in other metazoans, there are also a few examples of introns with GC–AG splice sites in the *C. elegans* genome (Farrer *et al.*, 2002; Sheth *et al.*, 2006). They are processed in the same manner as the GT–AG introns, indicating that the splicing machinery can also recognize introns with these borders. Presumably the surrounding nucleotides are able to direct the U1 snRNP into position in spite of the mismatched cytosine. Indeed, it has been observed that GC–AG introns typically have 5′ splice sites that, in all other regards, more closely complement the splice site recognition sequence on the U1 snRNP. Although many of these introns appear to be constitutively spliced, some have been shown to be alternatively spliced. In these cases, the weaker 5′ splice site consensus created by the mismatched cytosine is thought to influence splice site selection (Farrer *et al.*, 2002).

Other instances of splicing from noncanonical splice sites have been reported in *C. elegans*, and additional examples can be found by inspection of sequences annotated as splice sites in the genome. For example, it has been reported that introns with mutations in the 3′ splice site – such as AA, AT, GG, or TG – are nonetheless spliced at these mutant sites, albeit less efficiently than when the wild-type sequence is present (Aroian *et al.*, 1993). Also, there are numerous instances of use of non-AG 3′ splice sites in wild-type *C. elegans* genes (unpublished observations; WormBase, WS210). Furthermore, when genes containing a Tc1 transposon insertion are transcribed, the transposon RNA is often spliced out from 5′ splice sites as varied as TT or AT, and from 3′ splice sites like GG, TG, AC, or GC (Rushforth *et al.*, 1993; Rushforth and Anderson, 1996). The mechanism by which these reactions occur is not understood. However, it is clear from all of these findings that the cellular machinery responsible for splice site recognition in *C. elegans* is significantly less stringent than it is in other studied systems.

## 3. Intron Properties

These examples of noncanonical splicing indicate that, in spite of the extensive conservation of the intron-splicing mechanism among higher eukaryotes, some intron sequences and splicing factors have acquired specialized features in *C. elegans*. In general, introns in this organism have an A+U content of about 70%, while the average A+U content found in the exons flanking them is about 54%. Conversely, there is little elevation in the A+U content of yeast and mammalian introns (Blumenthal and Steward, 1997; Csank *et al.*, 1990). Studies in *C. elegans* and in plants have shown that steep transitions in A+U content assist in defining intron borders, and that artificial introns with high C+G content are not efficiently spliced (Conrad *et al.*, 1995; Goodall and Filipowicz, 1989).

Introns in *C. elegans* tend to be shorter than those in other metazoans. While some introns within the *C. elegans* genome are well over 1 kb long, the median intron length is only 65 nt, and the most common intron size is 47 nt. By comparison, the

most common intron size in flies is 59 nt, and in humans most introns are several kilobases long. It is clear that the *C. elegans* splicing machinery has evolved to process small introns. Studies have shown that short *C. elegans* introns cannot be processed by mammalian spliceosomes (Ogg *et al.*, 1990), in which a length of approximately 80 nt is necessary for efficient splicing to occur (Wieringa *et al.*, 1984). The unusually short length of the introns in *C. elegans* may be, in part, explained by how they are recognized and processed.

#### 4. Splice Site Recognition

The *C. elegans* 5' splice site is defined by the canonical metazoan consensus sequence AG/GURAGU, indicating that it is recognized by the U1 snRNP, as occurs in other eukaryotes (Blumenthal and Steward, 1997) (Fig. 1c). The 3' splice site consensus sequence, however, is much more extensive than those found in vertebrate or yeast introns. In *C. elegans*, there is a very highly conserved sequence UUUUCAG/R at the boundary between the intron and the next exon (Csank *et al.*, 1990; Sheth *et al.*, 2006; Spieth and Lawson, 2006). This 3' splice-site sequence is not efficiently recognized by the mammalian spliceosome (Kay *et al.*, 1987). All *C. elegans* introns also lack ANY branch-point consensus sequence (YURAY in mammals), even though worms do encode a SF1/BBP protein containing the conserved domain that recognizes this sequence (Blumenthal and Thomas, 1988; Mazroui *et al.*, 1999). Additionally, the *C. elegans* U2 RNA has an antisense branch-point sequence identical to that found in mammalian U2 (Thomas *et al.*, 1990). The only other apparent information content of *C. elegans* intron 3' splice sites is a peak of adenosines at -16 to -18 nt from the 3' splice site, which presumably serves as the site of branching. Such a variably positioned branch-point adenosine is not without precedent. It has been found that when possible branch-point adenosines are removed from mammalian introns, splicing proceeds through nearby alternative cryptic branch points (Ruskin *et al.*, 1985). A study in plants has also shown that several different adenosines in the last third of an intron can act as branch points during splicing (Goodall and Filipowicz, 1989). This may also be the case in nematodes, whose introns bear many similarities to those of plants. In fact, it has been observed in *C. elegans* that mutation of putative branch-point adenosines does not affect 3' splice site choice (Conrad *et al.*, 1993b). Finally, as has been described in plants, *C. elegans* introns often lack the polypyrimidine tract found immediately upstream of the 3' splice-site (Csank *et al.*, 1990; Goodall and Filipowicz, 1989; Spieth and Lawson, 2006). Generally, the short uridine stretch characterized as part of the UUUUCAG/R 3' splice site serves as the only polypyrimidine tract (Blumenthal and Thomas, 1988).

These modified intron features are indicators of some important differences in the *C. elegans* splicing mechanism. Both U2AF<sup>65</sup> and U2AF<sup>35</sup>, shown to recognize the polypyrimidine tract and 3' splice site in most metazoans, have been identified in *C. elegans* (Zorio *et al.*, 1997; Zorio and Blumenthal, 1999b). An early examination of the UUUUCAG/R 3' splice site showed that mutation of the uridine residues immediately upstream of the 3' splice site reduced proper splice site recognition

(Zhang and Blumenthal, 1996). In subsequent studies, it was determined that U2AF<sup>65</sup> recognized these residues in place of a lengthy polypyrimidine tract, while U2AF<sup>35</sup> interacted specifically with the terminal CAG/R (Hollins *et al.*, 2005; Zorio and Blumenthal, 1999a). Since the remainder of *C. elegans* splicing machinery has been conserved with respect to other metazoans, it is thought that, after U2AF binding, intron splicing proceeds canonically.

Comparison of branch point and 3' splice site recognition in different phyla suggests that they each recognize the intron border by different mechanisms, but using orthologous proteins. In the yeast *Saccharomyces cerevisiae*, SF1/BBP tightly binds the branch point, while there appears to be little consensus to direct U2AF binding to the polypyrimidine tract, which is often missing altogether (Rutz and Séraphin, 1999; Wang *et al.*, 2008). The AG at the 3' splice site may not be recognized by any protein, since U2AF<sup>35</sup> is not present in this organism. In mammals there are relatively weak consensus sequences for both U2AF<sup>65</sup> and BBP/SF1, so splice-site recognition may be more combinatorial (Berglund *et al.*, 1998). In worms, there is a tight binding consensus for both U2AF subunits right at the 3' splice site, but not for BBP/SF1. We hypothesize that U2AF recognizes the UUUUCAG/R at the 3' splice site, and then the branch-point adenosine is chosen by its proximity to SF1/BBP, which is perhaps already bound to U2AF, as it is in the fission yeast, *Schizosaccharomyces pombe* (Huang *et al.*, 2002). The fact that some *C. elegans* introns contain noncanonical 3' splice sites may indicate that the uridine tract in the UUUUCAG/R is sufficient to direct the splicing machinery to the correct site of splicing (Zhang and Blumenthal, 1996). This flexibility exists because U2AF<sup>65</sup> and U2AF<sup>35</sup> work in combination to specify the 3' splice site of an intron and direct U2 snRNP binding (Wu *et al.*, 1999; Zorio and Blumenthal, 1999a).

Intron length can also influence 3' splice site choice in *C. elegans* (Zhang and Blumenthal, 1996). In constructs containing 48 nt introns with the engineered 3' splice site UUCAA/AAG, splicing occurred with equal frequency after either the CAA or the AAG. However, when the intron length was increased to either 171 nt or 283 nt, splicing occurred predominantly after G. Furthermore, analysis of the splice sites of 139 *C. elegans* introns suggests that there are actually two classes of introns: frequent short introns, with an average length of 52 nt; and rarer long introns, with an average length of 551 nt. Both classes have the same 5' splice site consensus sequence but differ in the amount of variability (Fields, 1990).

## 5. Paired Splice Sites

The 3' ends of some *C. elegans* introns contain partially duplicated sequences, paired splice sites. Similar arrangements, which lead to a process called alternative tandem splicing, have been described in other organisms (Dou *et al.*, 2006; Hiller *et al.*, 2004, 2007). In mammals, the 3' splice site of such an intron contains the sequence NAGNAG. In this case, U2AF<sup>35</sup> can bind either of these sequences, demarcating two potential 3' splice points. In worms, a similar phenomenon occurs fairly frequently (Sullivan and Blumenthal, unpublished observations). In these



**Fig. 2** Paired 3' splice site consensus sequences. The upstream 3' splice site has no consensus other than a terminal AG. The downstream 3' splice site consists of the canonical *C. elegans* 3' consensus sequence TTTTCAG, except with a G in place of the T at -7 due to the preponderance of 6 bp spacing between the two splice sites. In introns, such sites are commonly observed separated by multiples of three bases. At trans-splice sites, splice site separation is not constrained by reading frame. (For color version of this figure, the reader is referred to the web version of this book.)

cases there is a good match to the UUUUCAG/R consensus, which U2AF recognizes as the intron border region. However, only some of the splicing actually occurs at this sequence. The rest of it occurs at a short distance upstream at a much weaker match to the consensus, often even at a non-AG site (Fig. 2). Documented cases of this phenomenon generally show 6, 9, or 12 nucleotides between the two splice sites, presumably because splicing at other nearby sites results in a frameshift and would not be detected due to RNA degradation by nonsense-mediated decay. Indeed, this phenomenon also occurs at trans-splice sites (see below) upstream of the translation initiation codon, and here the spacing between the sites does not occur in multiples of 3 (Blumenthal, unpublished).

## B. Trans-Splicing

### 1. Components of the Trans-Splicing Reaction

During transcription of many *C. elegans* genes, the 5' ends of their pre-mRNAs are replaced with a 22 nt trimethylguanosine-capped RNA leader sequence by trans-splicing (Krause and Hirsh, 1987). The mRNAs of approximately 70% of all genes in the *C. elegans* genome are trans-spliced (Blumenthal, 2005). Trans-splicing is closely related to intron splicing and utilizes the same U2, U4, U5, and U6 snRNAs found in the spliceosome, as well as their associated proteins. The U1 snRNP, however, does not play a role in this process (Hannon *et al.*, 1991). Instead, a different snRNP, the SL snRNP, is required for this reaction (Bruzik *et al.*, 1988; Thomas *et al.*, 1988).

After transcription, some snRNA molecules are transported from the nucleus to mature in the cytoplasm at a large protein assembly, the survival of motor neurons (SMN) complex (Patel and Bellini, 2008). Here, a heteroheptamer ring of Sm proteins is assembled around its Sm-binding site, RAUUUUGR. During assembly, the arginine residues of the D1 and D3 Sm proteins can be symmetrically dimethylated by a protein arginine methyltransferase, such as PRMT-5. It is thought that this modification increases the stability of the Sm protein heptamer by enhancing Sm affinity to the SMN complex. After addition of the Sm ring, the 7-methylguanosine cap of the snRNA is hypermethylated into a 2,2,7-trimethylguanosine

cap. Finally, the 3' end of the snRNA is trimmed, and the molecule is transported back into the nucleus as a mature snRNP. It is then processed with other snRNPs in the Cajal bodies, from which it can be released to participate in splicing.

## 2. The SL snRNP

Evidence indicates that the SL RNA molecules also mature by this process. The SL1 RNA (the first *C. elegans* SL RNA to be characterized) is transcribed by RNA polymerase II from a gene cluster (*rrs-1*) on chromosome V, containing ~110 tandem repeats of this gene and the gene encoding 5S rRNA in the opposite orientation (Ferguson *et al.*, 1996; Krause and Hirsh, 1987). The SL1 RNA primary transcript is 105 nt long. Early characterization of the SL1 RNA locus showed that it contains the same distal and proximal transcriptional promoter elements seen at the U1, U2, U4, and U5 snRNA loci (Thomas *et al.*, 1990). Although SL1 RNA nuclear export and Sm assembly have never been examined, *C. elegans* does have an ortholog of the SMN gene (Miguel-Aliaga *et al.*, 1999). Furthermore SL1 RNA contains a TMG cap and is bound by Sm proteins (Bruzik *et al.*, 1988; Thomas *et al.*, 1988; Van Doren and Hirsh, 1988), features also found in most spliceosomal snRNPs. Finally, Sm-containing SL1 snRNPs have been observed in the nucleus (MacMorris and Blumenthal, unpublished observations). These similarities suggested that the SL1 snRNP could play a role in trans-splicing analogous to the roles played by the spliceosomal U snRNPs in cis-splicing (Bruzik *et al.*, 1988). A subsequent *in vitro* study employing extract made from the nematode *Ascaris lumbricoides* (which also uses an SL snRNP to trans-splice its pre-mRNA) showed that synthetic SL RNA worked for trans-splicing following extract-mediated addition of the TMG cap and the Sm proteins (Maroney *et al.*, 1990).

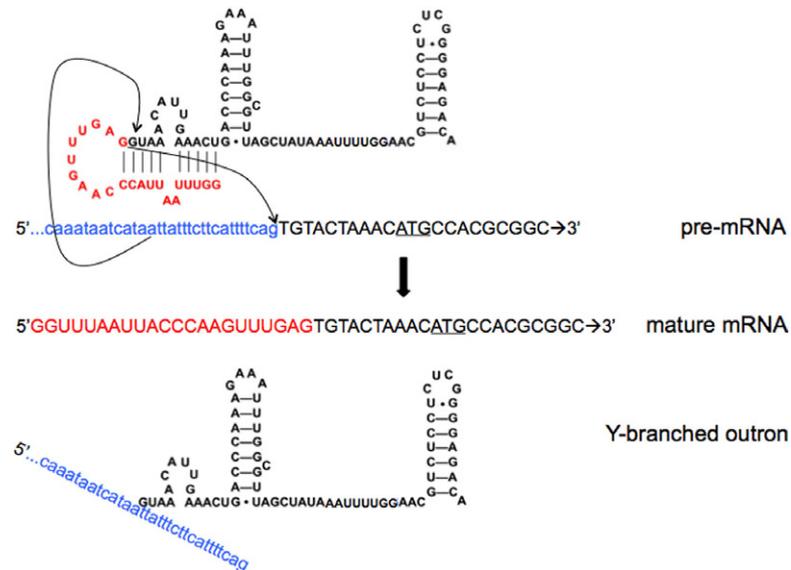
In its mature form, an SL snRNP is thought to be composed of three stem-loops and an Sm-binding domain (Bruzik *et al.*, 1988) (Fig. 3a). The first stem-loop contains the leader sequence, which is spliced onto an RNA molecule during trans-splicing. The nucleotide sequence at this splice site (AG/GUAAAC) closely mirrors the 5' splice site consensus sequence in introns (AG/GURAGU). Several studies have examined the effects of mutations incorporated into this leader sequence. Initial *in vitro* studies of trans-splicing in *Ascaris* extract indicated that the leader sequence could be extensively modified without abolishing trans-splicing (Maroney *et al.*, 1991). Later, it was discovered that the expression of a synthetic SL1 RNA transgene could rescue the embryonic lethality induced by deletion of the native *rrs-1* SL RNA gene locus (Ferguson *et al.*, 1996). Most deletions engineered into the leader sequence of this SL1 RNA transgene resulted in constructs incapable of rescuing this lethality (Ferguson *et al.*, 1996; Xie and Hirsh, 1998). The severity of these deletions was probably due to removal of a necessary promoter element from the leader sequence (Hannon *et al.*, 1990), because no transgenic SL1 RNA was transcribed from these constructs *in vivo*. Very small deletions or substitutions just upstream of the splice site were less detrimental. When SL1 RNA transcription was driven by the U2 promoter, essential features of the spliced leader could be more



Because there is no contiguous upstream 5' splice site, U2 cannot associate with a bound U1 snRNP. Instead, it interacts with the 5' splice site on the SL snRNP to initiate spliceosome formation and trans-splicing (Maroney *et al.*, 2000).

The 5' splice site is located in the first stem-loop of the SL RNA and is consequently base-paired in a manner similar to the pairing that occurs between the 5' splice site of an intron and the U1 snRNP during cis-splicing. In this case, the leader sequence of the SL1 snRNP was proposed to act as a chimeric molecule, with the upstream exon capable of interacting with downstream sequence in the same RNA molecule to initiate its own splicing (Bruzik *et al.*, 1988). However, it was subsequently shown that this base pairing is not required for trans-splicing *in vitro* (Maroney *et al.*, 1991).

Trans-splicing is also dependent on the U4–U5–U6 tri-snRNP (Maroney *et al.*, 1996, 2000). This complex interacts with the SL snRNP and the U2 snRNP at the 3' splice site to form an active spliceosome, although the mechanism of this process is poorly understood. Presumably the trans-splicing reaction occurs analogously to the cis-splicing reaction, so a branch-point adenosine upstream of the 3' splice site on the pre-mRNA attacks the 5' splice site of SL snRNP, forming a Y-branched molecule and freeing the leader sequence (Fig. 4). Such a Y-branched structure has been



**Fig. 4** Products of the trans-splicing reaction. The spliced leader of the SL1 snRNP is in red. Sequences found in the outon of the pre-mRNA (from *rsp-3*, in this case) are in blue. During the trans-splicing reaction, an upstream adenosine in the outon attacks the 5' splice site of the SL1 snRNP. The 3' end of the SL1 spliced leader then attacks the 3' splice site on the pre-mRNA to form a capped, trans-spliced RNA molecule. The outon is freed from the pre-mRNA but remains attached to the remainder of the snRNP by the 2'-5' phosphodiester bond at its branch-point adenosine, forming a Y-branched molecule. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)

identified on Northern blots, and treatment with debranching enzyme eliminated these branched RNAs (Bektesh and Hirsh, 1988). The 3' hydroxyl of the freed leader then attacks the phosphodiester bond at the 3' splice site, covalently attaching itself to the downstream RNA as a spliced leader and becoming the first exon of the mature mRNA. This results in a trans-spliced RNA molecule, complete with the TMG-capped spliced leader transferred from the SL snRNP (Liou and Blumenthal, 1990; Thomas *et al.*, 1988; Van Doren and Hirsh, 1990). This trans-spliced mRNA molecule is exported from the nucleus for translation. Because eukaryotic mRNA does not typically contain a TMG cap, specialized isoforms of the eukaryotic initiation factor eIF4E have evolved to recognize the TMG cap and initiate translation from these mRNA molecules (Keiper *et al.*, 2000; Wallace *et al.*, 2010).

#### 4. Properties of the Outron

The RNA between the transcriptional start site and the 3' splice site used in trans-splicing is the outtron (Conrad *et al.*, 1991). Several studies have characterized the features of an outtron that mark pre-mRNAs for trans-splicing. An intron lacking a 5' splice site can act as an outtron when placed in the 5'UTR of a synthetic construct (Conrad *et al.*, 1991). Furthermore, if a 5' splice site was constructed upstream of the 3' splice site used in trans-splicing, RNA produced from the construct was cis-spliced at the trans-splice site. The outtron, in effect, became an exon and an intron (Conrad *et al.*, 1993a). Several constructs in which increasingly long artificial outtrons, each composed only of AU-rich sequence and a 3' splice site, were placed upstream of a gene were found to be trans-spliced efficiently if the outtron was at least 51 nt in length (Conrad *et al.*, 1995). These studies were interpreted to mean that an outtron contains no sequences or features necessary for trans-splicing, although a loose consensus could have been inadvertently contained in the synthetic sequence. In addition, there appears to be a minimal length requirement, as has been observed for introns. Apparently, a 3' splice site preceded by an adequately sized AU-rich sequence with no upstream 5' splice site is sufficient to promote trans-splicing. However, a recent bioinformatic analysis of information content of *C. elegans* outtrons has identified a very weak consensus sequence, UUUUCUUU, termed the Ou element, centered about 50 nt upstream from the trans-splice site (Graber *et al.*, 2007). The functional significance of this motif remains to be investigated.

The removal and subsequent destruction of the 5' section of many pre-mRNA molecules by trans-splicing has complicated the task of determining the transcriptional start site of trans-spliced genes. Like cis-splicing, trans-splicing is a very efficient process (Blumenthal and Steward, 1997; Cramer *et al.*, 2001). Also, most *C. elegans* genes do not have obvious TATA boxes to indicate the approximate location of transcription start sites. For these reasons, the transcriptional start site is known for only a handful of trans-spliced genes (e.g., Kramer *et al.*, 1990; Park and Kramer, 1990). Traditional RNA-end identification methods, such as 5' RACE, SAGE, TEC-RED, and even primer extension can identify the presence and location of the spliced leader on mRNA molecules, but, since trans-splicing most likely happens

before transcription of pre-mRNA is even complete, there is very little outtron-containing pre-mRNA available from which a transcriptional start site can be identified.

Recent experiments have shown that nested primers can be used to determine outtron lengths and approximate transcriptional start sites by RT-PCR (Morton and Blumenthal, 2011). In addition, when chromatin immunoprecipitation (ChIP) with an antibody recognizing RNA polymerase II was performed on starved *C. elegans* larvae, deep sequencing of the precipitated DNA products revealed accumulated polymerase upstream of many genes, possibly paused at promoters (Baugh *et al.*, 2009). Likewise, *C. elegans* ChIP with an antibody recognizing the histone variant H2A.Z (HTZ-1), followed by hybridization to DNA microarrays, indicated that this type of histone is also often found just upstream of genes, possibly marking active promoters (Whittle *et al.*, 2008). The genomic sites of these polymerase and HTZ-1 peaks correspond closely to the putative start sites of several outtrons analyzed by RT-PCR (Table I). Finally, minimum outtron lengths can be estimated from occasional ESTs representing untrans-spliced RNA. Data from these multiple measures of outtron length indicate that average outtrons are around 300 nt, although examples of shorter and significantly longer outtrons have also been observed (Morton and Blumenthal, 2011).

## 5. Additional Components of the SL1 snRNP

At the conclusion of the trans-splicing reaction, the outtron becomes attached to the 78 nt 3' portion of the SL snRNP as a Y-branched molecule, analogous to the lariat byproduct associated with intron splicing (Bektesh *et al.*, 1988). This molecule is subsequently debranched and degraded, although the Sm proteins associated with this branched molecule may be recycled. The fate of the components of the truncated snRNP is an interesting question because, unlike the U snRNPs used in cis-splicing, the SL snRNPs required for trans-splicing are consumed in each reaction.

**Table I**  
Estimated outtron lengths

Gene	Outtron (RT-PCR)	Longest EST	ChIP-seq
rsp-3	247–255	All trans-spliced	250
rps-3	288–310	All trans-spliced	310
vha-6	363–506	All trans-spliced	No peak
Y37E3.8	263–321	All trans-spliced	270
col-13	57–100	65	85
pas-3	82–146	164	No peak
idh-2	302–347	304	310

*Note:* These genes have all been analyzed by RT-PCR, using forward primers in the outtron located successively upstream of the gene's trans-splice site. The transcriptional start site was determined to be in the region between the last forward primer from which a product could be obtained and the next forward primer upstream of this site. ESTs were taken from Wormbase, WS210. ChIP-seq data is from Baugh *et al.* (2009), analyzed after being loaded as a custom track on the UCSC genome browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>). A peak of polymerase was not identified upstream of all genes.

Analysis of *Ascaris* extract has identified two proteins, SL95p and SL30p, which specifically associate with the SL snRNP and are necessary for *in vitro* trans-splicing (Denker *et al.*, 1996). Additional examination of their function has shown that SL95p binds both the SL snRNP and SL30p, and that this complex can associate with SF1/BBP, effectively bringing the SL snRNP 5' splice site into proximity with the pre-mRNA 3' splice site (Denker *et al.*, 2002). Recently, orthologs of SL95p and SL30p have been identified in *C. elegans* (MacMorris *et al.*, 2007). SNA-2 and SNA-1 have also been shown to associate with the SL1 snRNP. RNAi-mediated knockdown of *sna-2* is lethal, as is the deletion of the *sna-2* gene. Deletion of *sna-1* produces cold-sensitive sterility. However, the *sna-1* mutant animals are still capable of SL1 trans-splicing, indicating that these proteins may not perform the function ascribed to them in *Ascaris*.

Furthermore, some SNA-2 is found bound to a SNA-1 paralog, SUT-1. Deletion of SUT-1 also leads to cold-sensitive sterility. This heterodimer does not associate with any SL snRNP, but can associate with one of the family of several recently discovered nematode-specific snRNPs, called SmY 1-12 (Jones *et al.*, 2009; MacMorris *et al.*, 2007). These SmY snRNPs are thought to fold into two stem-loops, flanking an Sm-binding site (Fig. 3b). It has been proposed that stem-loop nucleotides of these SmY snRNPs might base pair with complementary sequences in the stem-loops of the SL snRNPs. These interactions may aid in recycling Sm proteins from spent SL snRNPs after trans-splicing (MacMorris *et al.*, 2007).

## 6. Trans-Splicing in Other Species

Trans-splicing is not restricted to *C. elegans*. Soon after the discovery of the SL1 spliced leader and trans-splicing in *C. elegans*, RNA containing identical spliced leader sequence was discovered in nematodes from several other genera (Bektesh *et al.*, 1988). Indeed, further analysis showed that closely related variants of this spliced leader can be found in all but one of the five major clades of the nematode phylum (Guiliano and Blaxter, 2006; Pettitt *et al.*, 2008). The high degree of conservation observed throughout most of the phylum indicates that trans-splicing in nematodes probably arose in a common ancestor. Interestingly, however, a representative of the most basal nematode clade has trans-splicing, but its multiple spliced leaders appear unrelated to SL1 of the other four clades.

SL-type trans-splicing has also been observed in several other phyla, including platyhelminthes (Davis, 1997), rotifers (Pouchkina-Stantcheva and Tunnacliffe, 2005), cnidaria (Stover and Steele, 2001), the primitive chordates *Ciona intestinalis* (Vandenbergh *et al.*, 2001) and oikopleura (Ganot *et al.*, 2004), and trypanosomes (Sutton and Boothroyd, 1986), as well as several others including some but not all arthropods (Douris *et al.*, 2010). Although all of these organisms attach a spliced leader to their RNA, the leader sequences, as well as the SL RNAs that donate them, are unrelated to those found in nematodes. Additionally, these widely divergent phyla are evolutionarily separated from each other by phyla in which trans-splicing

has not been observed, so it is very likely that trans-splicing has arisen independently numerous times (Douris *et al.*, 2010).

## 7. A Second Class of Spliced Leaders

Shortly after the discovery of trans-splicing in *C. elegans* and the characterization of the SL1 RNA, another spliced leader was found at the 5' ends of mRNA molecules transcribed from specific genes (Huang and Hirsh, 1989). This second spliced leader, designated SL2, was found to be transcribed from several snRNA genes scattered throughout the genome. It is also capped with trimethylguanosine, contains an Sm protein binding site, and is predicted to fold into a structure closely resembling that of SL1 RNA (Fig. 3c). Soon after, several additional SL RNA genes were identified through characterization of their novel spliced leaders on various mRNA molecules (Ross *et al.*, 1995) (Kuwabara *et al.*, 1992). Further analysis indicated that these additional SL RNA genes are variants of SL2 RNA. Eleven SL2 variants, SL2 to SL12, are encoded by 18 genes scattered throughout the genome (MacMorris *et al.*, 2007).

An early examination of SL2 trans-spliced genes showed that the SL2 spliced leader is exclusively attached to pre-mRNAs transcribed from genes located in downstream positions in operons (discussed in the following section) (Spieth *et al.*, 1993). It is thought that the SL2 snRNP trans-splicing reaction occurs analogously to that of the SL1 snRNP, although systematic studies of its mechanism have not been reported. *In vivo* analysis using a marked SL2 RNA construct has shown that the sequence of the first 20 nt of the spliced leader can be altered without a significant drop in trans-splicing efficiency (Evans and Blumenthal, 2000). In contrast, the primary sequences of stem II and stem-loop III are necessary for trans-splicing activity and/or specificity. Unlike the SL1 snRNP, the SL2 snRNP does not associate with SNA-1 or SNA-2. It can potentially base-pair with the stem-loops of SmY snRNPs, an observation that provides additional support for the idea that the SmY snRNPs have a role in Sm recycling (MacMorris *et al.*, 2007). It has also been found that SL2 RNA overexpression will partially rescue the lethality resulting from deletion of the *rrs-1* (SL1 RNA) locus (Ferguson *et al.*, 1996). Thus, when forced, SL2 snRNP can function in place of SL1. Clearly, however, the two classes of SL snRNP are not completely interchangeable, since the SL2 snRNP normally does not donate a spliced leader to pre-mRNAs transcribed from nonoperon genes or first genes in operons (Hillier *et al.*, 2009; Spieth *et al.*, 1993).

Like the SL1 spliced leader, the SL2 spliced leader can be found in other species of nematode (Blumenthal, 2005; Evans *et al.*, 1997; Lee and Sommer, 2003). All of these nematode species are relatively closely related to *C. elegans*, indicating that this second spliced leader evolved at some point subsequent to the divergence of the rhabditid group of nematodes (Guiliano and Blaxter, 2006). Furthermore, while the SL1 spliced leader is largely invariant throughout most of the nematode phylum, SL2 sequence varies considerably between different species. In all cases, it is used to trans-splice downstream genes in operons. However, operons have been discovered

in nematode species outside the Rhabditina clade DITTO that use SL1 or SL1 variants to trans-splice downstream gene mRNAs (Liu *et al.*, 2009).

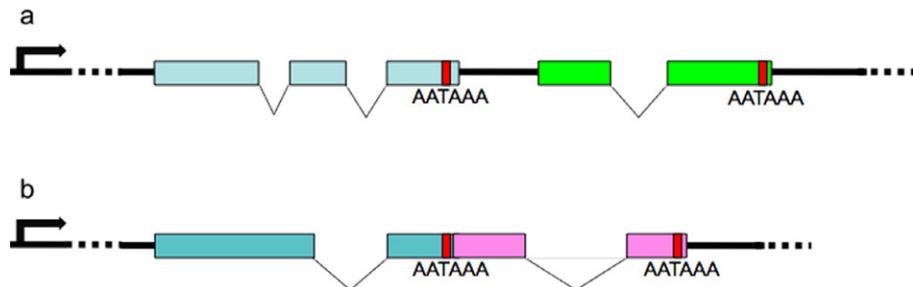
## C. Operons

### 1. Operon Architecture

The *C. elegans* genome contains more than 1000 operons, in which several closely clustered genes are transcribed from a single, upstream promoter, and the resulting polycistronic pre-mRNA is cut into monocistronic units by 3' end formation and trans-splicing, before the individual mRNAs exit the nucleus for translation (Blumenthal, 2005; Nilsen, 1994; Spieth *et al.*, 1993). Unlike bacterial operons, these operon genes usually encode metabolically unrelated products.

### 2. Discovery of *C. elegans* Operons

The discovery of operons in *C. elegans* (Spieth *et al.*, 1993) (Fig. 5a) was facilitated by identification of the SL2 spliced-leader attached to only a subset of mRNAs. As SL2 trans-spliced mRNAs were identified, it was observed that they were always transcribed from genes located in the same orientation and immediately downstream of other genes, usually with only 100 bp separating the 3' end of the



**Fig. 5** Typical *C. elegans* operons. A single promoter (arrow) drives expression of both genes, in which exons are depicted as colored boxes and introns are shown as chevrons. Each gene contains an independent 3' end formation signal (AATAAA). (a) An SL2-type operon. An intergenic region (ICR) averaging approximately 100 nt (shown as a bold line) separates the genes within the operon. As 3' end formation occurs in the upstream gene, the downstream gene is trans-spliced with an SL2 spliced leader and the ICR is excised as a Y-branched molecule on the used SL2 snRNP. (b) An SL1-type operon. This type of operon contains no ICR and trans-splicing of the downstream genes is conducted by the SL1 snRNP. Studies have shown that 3' end formation of the upstream gene can occur by the canonical mechanism, in which case RNA cleavage destroys the trans-splice site of the downstream gene. Alternately, trans-splicing of the downstream gene can inhibit canonical 3' end formation of the upstream gene. New evidence suggests, however, that downstream trans-splicing in this type of operon may simultaneously act as an alternate 3' end processing pathway in the upstream gene, allowing both genes in the operon to be expressed. (For color version of this figure, the reader is referred to the web version of this book.)

upstream gene and the site of SL2 trans-splicing at the 5' end of the downstream gene. Furthermore, attempts to identify upstream promoters conferring SL2 specificity to transcripts from these downstream genes were unsuccessful (Blumenthal, 1995). Ultimately, the identification of cDNA containing the last exon of an upstream gene followed by a short intercistronic region (ICR) and the downstream gene indicated that both genes were transcribed together from a common upstream promoter (Spieth *et al.*, 1993).

However, the strongest evidence for operons comes from the almost perfect correlation between SL2 trans-splicing and the location of the trans-splice site at a downstream position in a tight gene cluster (Blumenthal *et al.*, 2002). For this reason, these operons have been named SL2-type operons. Recently this evidence has been extended to the entire genome by analysis of the transcriptome by deep sequencing (Allen *et al.*, 2011). This analysis demonstrates that SL1 and SL2 trans-splicing occur on different genes and that frequent SL2 trans-splicing is almost perfectly correlated with downstream genes spaced about 100 bp from the 3' end of a gene upstream.

Around 15% of the genes in *C. elegans* are arranged in operons (Blumenthal and Gleason, 2003). These operons are concentrated within the central region of the autosomes and there is a paucity of operons on the X chromosome (Blumenthal *et al.*, 2002). Each operon contains from two to eight genes, usually transcribed from a single upstream promoter, although some operons also contain internal promoters (Huang *et al.*, 2007). Although operons in nematodes do share some similarities with bacterial operons, they are evolutionarily unrelated. Instead, they evolved independently, presumably after the development of trans-splicing. Trans-splicing can isolate translatable single-gene units from a polycistronic pre-mRNA. These capped and spliced monocistronic mRNAs can then be exported from the nucleus for translation. The RNA transcribed from the first genes in operons is not always trans-spliced, but when trans-splicing does occur, it is to SL1, as occurs in transcripts from genes not in operons.

### 3. Functional Relationships of Operon Genes

Often the genes within a single *C. elegans* operon appear functionally unrelated, although numerous instances of related genes being expressed in a single operon have been observed. For example, an operon has been identified that contains genes encoding U2AF<sup>35</sup> and SF1/BBP, proteins involved in 3' splice-site recognition, along with a third gene, *cyn-13*, the human ortholog of which is also spliceosome-associated. Indeed, groupings similar to this one have been found to occur more frequently than expected by chance (Blumenthal and Gleason, 2003). It has also been observed that genes expressed preferentially in the female germ line occur in operons significantly more frequently than expected by chance (Reinke and Cutter, 2009). These tend to be genes encoding proteins involved in universal processes such as gene expression (the basic machinery of transcription, RNA degradation, splicing, and translation) or mitochondrial function. Conversely, genes encoding proteins expressed

only in certain tissues or developmental stages (such as transcription factors and collagens) are rarely found in operons (Blumenthal and Gleason, 2003). Thus, operons may have evolved in nematodes not to coordinate expression of similarly regulated genes, but to ensure universal expression of continuously needed gene products (Blumenthal, 2004). It has been proposed that in *C. elegans* operons serve to allow rapid response to global signals (Blumenthal and Gleason, 2003). In support of this idea, it has recently been shown that operons serve to respond to a need for growth following starvation or during development (Zaslaver *et al.*, 2011).

Several studies have shown that the transcripts from genes in operons are not always present in equal amounts (Blumenthal, 2005; Cutter *et al.*, 2009). This could be due to a variety of causes. For example the individual mRNAs might have different stabilities; processing sites might have differential or even regulated efficiencies, resulting in failure to make stable downstream mRNAs or even termination of transcription. Recently, it has been demonstrated that some operons contain internal promoters (Huang *et al.*, 2007; Whittle *et al.*, 2008). Hybrid operons feature an upstream promoter, which drives expression of all genes in the cluster, but they also contain an additional internal promoter, located within the operon. Presumably, such a system allows transcription of all genes from the promoter at the 5' end of the complex, while allowing transcription of specific genes in the operon during various periods of alternative metabolic requirements. Recent transcriptome evidence suggests that the promoters within the operons tend to be differentially regulated compared to those at the operon 5' ends (Allen *et al.*, 2011).

#### 4. Elements Controlling SL2 Trans-Splicing

The mechanism that controls SL2 specificity is incompletely understood, although intensive study of the process since its discovery has resolved some questions. During processing, a polycistronic pre-mRNA must be divided into separate genes, each competent for nuclear export and translation. At the 3' end of every gene, several signals have been identified that signal the RNA processing machinery to cleave the nascent RNA from the polymerase and polyadenylate it (Mandel *et al.*, 2008). Key among these signals is the polyadenylation signal, AAUAAA (Wickens and Stephenson, 1984). The machinery of 3' processing will be discussed in a later section. Typically, 3' processing in some way signals the RNA polymerase to terminate transcription (Rosonina *et al.*, 2006), and this must be prevented at 3' end formation sites within operons. There are currently two models to describe how termination might occur: the allosteric model and the torpedo model. According to the allosteric model, interaction with the 3' processing machinery induces a conformational shift in the RNA polymerase, resulting in weaker binding to DNA, reduced processivity, and termination (Logan *et al.*, 1987). The torpedo model proposes that the uncapped RNA produced downstream of the 3' cleavage site is subject to rapid 5'→3' degradation (Connelly and Manley, 1988; Kim *et al.*, 2004; West *et al.*, 2004). The exonuclease responsible for this eventually catches up to the polymerase, knocking it off the DNA. A hybrid termination model combines features of both of these

mechanisms (Luo *et al.*, 2006). Whatever the mechanism, within operons, this termination signal either never occurs or is inhibited. Polymerase continues to transcribe downstream genes and only terminates after the last gene in the operon (Haenni *et al.*, 2009).

Several features of the operon help carry out the processing. The nucleotides between the polyadenylation site of an upstream gene and the trans-splice site of a downstream gene constitute the ICR. Typically, an ICR is around 100 nt long, although examples of significantly longer ICRs have been observed (Blumenthal *et al.*, 2002). This region and the nearby upstream 3' end processing signals have been extensively examined for sequence elements and motifs that enable RNA processing machinery to suspend transcription termination and correctly SL2 trans-splice downstream genes (Huang *et al.*, 2001). An upstream polyadenylation site is not necessary to specify downstream SL2 trans-splicing, although it does appear to play some role in favoring SL2 over SL1 (Kuersten *et al.*, 1997; Spieth *et al.*, 1993). Removal of the AAUAAA still allows SL2 trans-splicing, but SL1 trans-splicing increases.

The fact that the AAUAAA is somehow involved and the strong tendency for the ICR to be just about 100 bp long suggested the 3' end formation machinery might be involved in the specification of SL2 trans-splicing. It is possible that the AAUAAA specifies downstream transcription termination, and the short ICR ensures that downstream trans-splicing occurs before termination occurs (Spieth *et al.*, 1993). Also, analysis of the SL2 snRNP showed that it coimmunoprecipitates with a protein component of the 3' end processing machinery, named for its mammalian ortholog-cleavage stimulation factor, 64kd subunit (CstF-64) (Evans *et al.*, 2001). Importantly, residues in the third stem-loop, necessary for SL2 trans-splicing specificity (Evans and Blumenthal, 2000), were also required for CstF-64 interaction.

Analysis of the sequence within the ICR identified the uridylyate-rich (Ur) element (Huang *et al.*, 2001). It was initially thought to have no consensus sequence and be identifiable only based on its nucleotide content. The mutation of this element resulted in complete loss of downstream RNA. If 3' end formation was prevented, downstream RNA did accumulate, but if the Ur element was absent, only SL1 trans-splicing was observed. Further *in vivo* analysis showed that, in an operon construct driven by a heat shock promoter, uncapped RNA extending from the 5' end of the Ur element through the downstream gene could be detected (Liu *et al.*, 2003), indicating that this element likely transiently stabilized downstream pre-mRNA to allow SL2 trans-splicing to occur. This stability could be recapitulated by the replacement of the Ur element with MS2 coat protein binding sites combined with expression of the MS2 coat protein.

Given the established relationship between the SL2 snRNP and CstF64, it was initially hypothesized that the Ur element served as the binding site for this processing factor. However, recent *in vitro* mutational analysis of both the 3' end formation signals and Ur element, coupled with a bioinformatic analysis of trans-splicing control elements have shown that CstF64 actually binds to a U-rich sequence just downstream from the 3' cleavage site. The Ur element is a separate feature

(Graber *et al.*, 2007). Subsequent study has shown that the Ur element contains a highly conserved hairpin of variable sequence, followed closely by a UAYYUU consensus sequence (Lasda *et al.*, 2010). Both of these elements are required for downstream SL2 trans-splicing *in vitro*. Interestingly, the UAYYUU sequence is complementary to the sequence surrounding the donor splice site on the SL2 snRNP, indicating that the same type of interaction may occur here as occurs between the U1 snRNP and the 5' splice site during cis-splicing, except that in trans-splicing, the splice site is on the snRNP and the site that binds it is on the pre-mRNA.

These observations have produced the following current, albeit incomplete, model of SL2 trans-splicing to downstream genes in operons. Transcription proceeds through the first gene, and 3' end formation factors are recruited to cleave and polyadenylate the mRNA. The polymerase continues to transcribe the operon, but termination is inhibited, possibly by CstF-facilitated recruitment of the SL2 snRNP to the Ur element. This protects the uncapped RNA downstream from this cleavage event from degradation and inhibits the torpedoing of the polymerase. The SL2 snRNP then interacts with the trans-splice site of the downstream gene, initiating spliceosome formation. Once trans-splicing occurs, the downstream RNA is capped and safe from degradation so the polymerase can continue to transcribe without premature termination.

As noted previously, there are several different species of SL2 spliced leader (MacMorris *et al.*, 2007; Ross *et al.*, 1995). Any additional specificity rendered by these variant SL2s has not been discovered, although they could be involved in spatial or temporal regulation of gene expression.

## 5. SL1-Type Operons

There is a second type of operon in *C. elegans*, the SL1-type operon (Fig. 5b). This class of operon does not contain a canonical 100 nt ICR between its genes. Instead, the polyadenylation signal of the first gene immediately precedes the trans-splice site of the second gene; that is there is no ICR at all. Initially, the overlap between these two sites led researchers to suggest that trans-splicing of the downstream gene would leave a free 3' end on the upstream mRNA that could then be polyadenylated, instructed by the polyadenylation signal that is present just upstream of the trans-splice site. However, experiments with one SL1-type operon suggested that, at least in this case, formation of the 3' end of the upstream mRNA is not dependent on trans-splicing of the downstream mRNA (Williams *et al.*, 1999). It was concluded that any given pre-mRNA from an SL1-type operon could give rise to either a functional mRNA for the upstream gene or the downstream gene, but not both, since 3' cleavage for the upstream mRNA would destroy the trans-splice site of the downstream mRNA. Conversely, trans-splicing of the downstream mRNA would leave a branched upstream mRNA, which might or might not allow formation of the upstream mRNA. This would result in a rather interesting regulatory situation.

Recently, processing of this type of operon has been reexamined. As the results of various transcriptional deep-sequencing projects have been reported, it has been

noted that polyadenylation of the upstream mRNA in these operons almost always occurs precisely at the AG of the trans-splice site, even though polyadenylation on G residues is extremely rare in *C. elegans* (Blumenthal, unpublished observations.) This observation suggests strongly that trans-splicing of the downstream gene mRNA is actually the process responsible for 3' end formation of the upstream gene mRNA. Upstream 3' end cleavage by trans-splicing would represent a novel RNA-processing mechanism.

Incidental observations indicate that this mechanism could be employed occasionally even in canonical SL2 operons. Indeed, during the initial characterization of operons in *C. elegans*, RNA that terminated not at the polyadenylation sequence of an upstream gene but at the downstream trans-splice site was detected (Spieth *et al.*, 1993). Furthermore, in cases where the upstream gene in an operon does not contain a well-recognized polyadenylation signal, 3' end formation often fails to occur. When this happens, trans-splicing of the downstream gene and polyadenylation of the upstream gene occurs at the 3' end generated at the downstream trans-splice site, effectively creating a long 3' UTR of the upstream gene (Liu *et al.*, 2003).

## 6. How Widespread are Operons?

Although operons in metazoans were first discovered in *C. elegans*, they are by no means unique to this group of nematodes or indeed to the nematode phylum. They are also present in *C. briggsae*, as well as several other species in the rhabditid group. Evidence for operons exists for more distantly related nematodes, including *Brugia malayi* and *Ascaris suum* (Guiliano and Blaxter, 2006; Liu *et al.*, 2010; ). Although operons have not yet been identified in the nematode species most distantly related to *C. elegans*, such as *Trichinella spiralis*, the presence of a spliced leader in this species makes their existence possible (Pettitt *et al.*, 2008). It is reasonable to suppose that operons evolved shortly after the arrival of trans-splicing and are universal among nematodes. Once formed, operons are, of necessity, extremely difficult to lose, evolutionarily. Since downstream genes are now separated from their promoters, their duplication or transposition is less likely to result in a functional gene (Qian and Zhang, 2008). Additionally, since processing of the polycistronic RNA resulting from operons requires trans-splicing, out-of-frame start codons can evolve in the DNA sequences upstream of a trans-splice site (Blumenthal, 2005). For this reason, studies have concluded that, once formed, operons are not easily lost through evolution. However, even in closely related species, such as *C. elegans* and *C. briggsae*, separated by about 100 million years of evolution, operon content and arrangement are notably different (Qian and Zhang, 2008). When more distantly related nematodes were also examined, these authors documented numerous additional instances of operon loss and rearrangement.

Outside of the nematode phylum, operons have been reported in several additional phyla. It appears that in all of these instances trans-splicing has arisen independently (from cis-splicing), and that operons subsequently arose in the genome. In at least two cases, arthropods and chordates, it appears that trans-splicing and possibly

operons arose relatively recently, since most branches of these phyla do not have these features (Douris *et al.*, 2010).

#### D. 3' End Formation

##### 1. Canonical 3' Processing Factors in Worms

In *C. elegans*, orthologs of all members of each of the 3' end formation complexes can be found, although some differences in their functionality have been proposed. For instance, it has been observed that the polyadenylation signal in *C. elegans* is more variable than seen in other organisms. Only around 50% of all genes contain the canonical AAUAAA sequence. Instead, many different variations of this sequence are employed, including AAUGAA, UAUAAA, GAUAAA, CAUAAA, UAUGAA, and AUUAAA (Hajarnavis *et al.*, 2004). Additionally, roughly 7% of genes have no recognizable signal related to AAUAAA at all (Blumenthal and Steward, 1997; Salisbury *et al.*, 2006). This sequence variability indicates that CPSF-160 may recognize this signal differently in *C. elegans* than it does in higher metazoans, even though the protein sequence is reasonably well conserved. Furthermore, analysis of 3' cleavage sites in *C. elegans* ESTs has shown that RNA cleavage typically occurs ~20 nucleotides downstream of the polyadenylation signal (Hajarnavis *et al.*, 2004). Studies examining the mechanisms behind these differences have not been performed.

Similar to CPSF-160, the CstF-64 protein seems to have also become specialized, although its exact role remains unclear. While the mammalian CstF-64 protein contains 597 amino acids, the *C. elegans* ortholog contains only 336 amino acids. Conservation between the two proteins extends only through the N-terminal region (approximately the initial 200 amino acids.) The entire C-terminal section of the *C. elegans* CstF-64 aligns very poorly to its mammalian counterpart (Salisbury *et al.*, 2006), an indication that this section of the protein may have evolved a specialized function in nematodes. Furthermore, bioinformatic analysis has indicated that the CstF-64 binding site on the mRNA is not the canonical GU-rich sequence ~25 nt downstream from the cleavage site, as is found in mammals, but rather a U-rich region just 3' to the site of cleavage (Graber *et al.*, 2007).

##### 2. Additional Proteins Implicated in 3' Processing

Recently, several additional proteins have been implicated in 3' end formation in *C. elegans* (Cui *et al.*, 2008). The genes encoding these proteins were discovered in a screen for suppressors of the Muv phenotype of an unusual deletion/insertion allele in the *lin-15* operon that revealed several 3' end formation and transcription termination proteins. The screen was conducted in a *C. elegans* strain in which a genomic fragment containing the 3' end and downstream region of the nonoperon gene H18N23.2 had been transposed into the third exon of the *lin-15B* gene. Since *lin-15B* is in an operon, upstream of *lin-15A*, insertion of this site not only inactivates

*lin-15B*, but also terminates transcription upstream of *lin-15A*. The combined absence of both of these gene products results in an easily identifiable synthetic phenotype, characterized by several vulvae protruding from the side of the mutant animals. When these animals were subjected to a genome-wide RNAi screen, the multivulva phenotype was suppressed by the inactivation of genes responsible for (among other things) 3' end formation or transcription termination.

In addition to identifying several of the genes producing components of the CPSF and CstF complexes, this screen implicated the genes *cids-1*, *cids-2*, *nrd-1*, and *rsp-6* in these processes. The *cids-1* gene is an ortholog of the *S. cerevisiae* gene *RTT103*. In yeast, Rtt103p is a member of a complex, also containing Rat1p and Rai1p, that have a role in transcription termination (Kim *et al.*, 2004). The *cids-2* gene is a paralog of *cids-1* and has no similarity to any other known gene. Even within *C. elegans*, similarity between the genes is restricted to a single region known to interact with the C-terminal domain of RNA polymerase II, the CID motif. The *nrd-1* gene is an ortholog of the yeast gene *NRD1*, known to play a role in the termination of small nuclear RNAs and of cryptic unstable transcripts (CUTs) (Arigo *et al.*, 2006; Thiebaut *et al.*, 2006). It also contains a CID motif, but has not been previously linked with *cids-1* or *cids-2*. Finally, *rsp-6* encodes the SR protein SRp20, known to have a role in alternative splicing (de la Mata and Kornblihtt, 2006; Longman *et al.*, 2000).

Additional analysis indicated that both CIDS-1 and CIDS-2 are involved in 3' end cleavage, but the specific nature of their roles is still unclear. Likewise, NRD-1 may help stimulate 3' end cleavage, although its role in terminating CUTs may have also been a factor in its identification as a suppressor of 3' end formation and transcription termination. Finally, SRp20 functions in transcription termination without affecting 3' end cleavage, although its mechanism of action remains undetermined (Cui *et al.*, 2008).

### 3. A Summary of RNA Processing in *C. elegans*

As in other eukaryotes, RNA processing in *C. elegans* occurs primarily during transcription. Intron splicing is conducted by the canonical spliceosomal machinery, but the 3' splice-site consensus sequence is markedly different from that found in other splicing organisms. Additionally, the splicing machinery is occasionally capable of splicing at noncanonical sites. These features have allowed *C. elegans* introns to be unusually short, when compared with the introns of other organisms.

*C. elegans* RNA is also processed in several other unique ways. The 5' region of approximately 70% of transcripts is replaced by a common 22 nt capped leader sequence through trans-splicing, catalyzed by the specialized SL1 snRNP. The discarded 5' piece of RNA, the outtron, is removed and degraded during this process, making it difficult to identify the promoter and transcriptional start site for these *C. elegans* genes. Trans-splicing has also facilitated the evolution of operons, since this process allows a polycistronic pre-mRNA to be divided into several capped monocistronic transcripts.

As upstream genes in an operon are transcribed, CPSF and CstF stimulate 3' end formation at the end of each gene-encoding section. RNA polymerase termination, usually coupled with 3' end formation, is suspended. It is possible that this is mediated by the CstF64 interaction with the SL2 snRNP, which is used specifically to trans-splice downstream genes in operons. Downstream SL2 specificity appears to be ensured by the Ur element in the ICR, which is capable of base pairing with the SL2 snRNP after its arrival during upstream 3' end formation. The RNA duplex may be sufficient to protect the downstream RNA from degradation until trans-splicing places a new cap on its 5' end. It is also possible that recently discovered proteins CIDS-1 or CIDS-2 may play some role in coordinating 3' end formation, trans-splicing, and transcription termination in *C. elegans* operons.

## Acknowledgments

The authors thank Peg MacMorris and Erika Lasda for their careful reading of this manuscript and their numerous helpful suggestions.

## References

- Aebi, M., Hornig, H., and Weissmann, C. (1987). 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. *Cell* **50**, 237–246.
- Allen, M. A., Hillier, L. W., Waterston, R. H., and Blumenthal, T. (2011). A global analysis of trans-splicing. *Genome Res.* **21**, 255–264.
- Arigo, J. T., Eyler, D. E., Carroll, K. L., and Corden, J. L. (2006). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell* **23**, 841–851.
- Aroian, R. V., Levy, A. D., Koga, M., Ohshima, Y., Kramer, J. M., and Sternberg, P. W. (1993). Splicing in *Caenorhabditis elegans* does not require an AG at the 3' splice acceptor site. *Mol. Cell Biol.* **13**, 626–637.
- Baugh, L. R., Demodena, J., and Sternberg, P. W. (2009). RNA Pol II accumulates at promoters of growth genes during developmental arrest. *Science* **324**, 92–94.
- Bektesh, S., Van Doren, K., and Hirsh, D. (1988). Presence of the *Caenorhabditis elegans* spliced leader on different mRNAs and in different genera of nematodes. *Genes Dev.* **2**, 1277–1283.
- Bektesh, S. L., and Hirsh, D. I. (1988). *C. elegans* mRNAs acquire a spliced leader through a trans-splicing mechanism. *Nucleic Acids Res.* **16**, 5692.
- Berglund, J. A., Abovich, N., and Rosbash, M. (1998). A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.* **12**, 858–867.
- Blumenthal, T. (1998). Gene clusters and polycistronic transcription in eukaryotes. *Bioessays* **20**, 480–487.
- Blumenthal, T. (1995). Trans-splicing and polycistronic transcription in *Caenorhabditis elegans*. *Trends Genet.* **11**, 132–136.
- Blumenthal, T., and Thomas, J. (1988). Cis and trans mRNA splicing in *C. elegans*. *Trends Genet.* **4**, 305–308.
- Blumenthal, T. (2004). Operons in eukaryotes. *Briefings Funct. Genomics Proteomics* **3**, 199–211.
- Blumenthal, T. (2005). Trans-Splicing and Operons (2005). In (WormBook, ed.). The *C. elegans* Research Community (Wormbook, doi/10.1895/wormbook.1.5.1). Available at: <http://www.wormbook.org>.
- Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., Chiu, W. L., Duke, K., and Kiraly, M., et al. (2002). A global analysis of *Caenorhabditis elegans* operons. *Nature* **417**, 851–854.

- Blumenthal, T., and Gleason, K. S. (2003). *Caenorhabditis elegans* operons: Form and function. *Nat. Rev. Genet.* **4**, 112–120.
- Blumenthal, T., and Steward, K. (1997). RNA processing and gene structure. In *C. elegans II* (D.L. Riddle, ed.), pp. 117–145. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Brown, J. W., Daniels, C. J., and Reeve, J. N. (1989). Gene structure, organization, and expression in archaeobacteria. *Crit. Rev. Microbiol.* **16**, 287–337.
- Bruzik, J. P., Van Doren, K., Hirsh, D., and Steitz, J. A. (1988). Trans splicing involves a novel form of small nuclear ribonucleoprotein particles. *Nature* **335**, 559–562.
- Burge, C. B., Padgett, R. A., and Sharp, P. A. (1998). Evolutionary fates and origins of U12-type introns. *Mol. Cell* **2**, 773–785.
- Clayton, C. E. (2002). Life without transcriptional control? From fly to man and back again. *EMBO J.* **21**, 1881–1888.
- Connelly, S., and Manley, J. L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2**, 440–452.
- Conrad, R., Lea, K., and Blumenthal, T. (1995). SL1 trans-splicing specified by AU-rich synthetic RNA inserted at the 5' end of *Caenorhabditis elegans* pre-mRNA. *RNA* **1**, 164–170.
- Conrad, R., Liou, R. F., and Blumenthal, T. (1993a). Conversion of a trans-spliced *C. elegans* gene into a conventional gene by introduction of a splice donor site. *EMBO J.* **12**, 1249–1255.
- Conrad, R., Liou, R. F., and Blumenthal, T. (1993b). Functional analysis of a *C. elegans* trans-splice acceptor. *Nucleic Acids Res.* **21**, 913–919.
- Conrad, R., Thomas, J., Spieth, J., and Blumenthal, T. (1991). Insertion of part of an intron into the 5' untranslated region of a *Caenorhabditis elegans* gene converts it into a trans-spliced gene. *Mol. Cell Biol.* **11**, 1921–1926.
- Cramer, P., Srebrow, A., Kadener, S., Werbajh, S., de la Mata, M., Melen, G., Nogués, G., and Kornblihtt, A. R. (2001). Coordination between transcription and pre-mRNA processing. *FEBS Lett.* **498**, 179–182.
- Csank, C., Taylor, F. M., and Martindale, D. W. (1990). Nuclear pre-mRNA introns: Analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. *Nucleic Acids Res.* **18**, 5133–5141.
- Cui, M., Allen, M. A., Larsen, A., MacMorris, M., Han, M., and Blumenthal, T. (2008). Genes involved in pre-mRNA 3'-end formation and transcription termination revealed by a *lin-15* operon Muv suppressor screen. *Proc. Natl Acad. Sci. U. S. A.* **105**, 16665–16670.
- Cutter, A. D., Dey, A., and Murray, R. L. (2009). Evolution of the *Caenorhabditis elegans* genome. *Mol. Biol. Evol.* **26**, 1199–1234.
- Davis, R. E. (1997). Surprising diversity and distribution of spliced leader RNAs in flatworms. *Mol. Biochem. Parasitol.* **87**, 29–48.
- Denker, J. A., Maroney, P. A., Yu, Y. T., Kanost, R. A., and Nilsen, T. W. (1996). Multiple requirements for nematode spliced leader RNP function in trans-splicing. *RNA* **2**, 746–755.
- Denker, J. A., Zuckerman, D. M., Maroney, P. A., and Nilsen, T. W. (2002). New components of the spliced leader RNP required for nematode trans-splicing. *Nature* **417**, 667–670.
- Dou, Y., Fox-Walsh, K. L., Baldi, P. F., and Hertel, K. J. (2006). Genomic splice-site analysis reveals frequent alternative splicing close to the dominant splice site. *RNA* **12**, 2047–2056.
- Douris, V., Telford, M. J., and Averof, M. (2010). Evidence for multiple independent origins of trans-splicing in metazoa. *Mol. Biol. Evol.* **27**, 684–693.
- Evans, D., and Blumenthal, T. (2000). trans splicing of polycistronic *Caenorhabditis elegans* pre-mRNAs: analysis of the SL2 RNA. *Mol. Cell Biol.* **20**, 6659–6667.
- Evans, D., Perez, I., MacMorris, M., Leake, D., Wilusz, C. J., and Blumenthal, T. (2001). A complex containing CstF-64 and the SL2 snRNP connects mRNA 3' end formation and trans-splicing in *C. elegans* operons. *Genes Dev.* **15**, 2562–2571.
- Evans, D., Zorio, D., MacMorris, M., Winter, C. E., Lea, K., and Blumenthal, T. (1997). Operons and SL2 trans-splicing exist in nematodes outside the genus *Caenorhabditis*. *Proc. Natl Acad. Sci. U. S. A.* **94**, 9751–9756.

- Farrer, T., Roller, A. B., Kent, W. J., and Zahler, A. M. (2002). Analysis of the role of *Caenorhabditis elegans* GC-AG introns in regulated splicing. *Nucleic Acids Res.* **30**, 3360–3367.
- Ferguson, K. C., Heid, P. J., and Rothman, J. H. (1996). The SL1 trans-spliced leader RNA performs an essential embryonic function in *Caenorhabditis elegans* that can also be supplied by SL2 RNA. *Genes Dev.* **10**, 1543–1556.
- Ferguson, K. C., and Rothman, J. H. (1999). Alterations in the conserved SL1 trans-spliced leader of *Caenorhabditis elegans* demonstrate flexibility in length and sequence requirements *in vivo*. *Mol. Cell Biol.* **19**, 1892–1900.
- Fields, C. (1990). Information content of *Caenorhabditis elegans* splice site sequences varies with intron length. *Nucleic Acids Res.* **18**, 1509–1512.
- Fitzgerald, M., and Shenk, T. (1981). The sequence 5'-AAUAAA-3' forms parts of the recognition site for polyadenylation of late SV40 mRNAs. *Cell* **24**, 251–260.
- Ganot, P., Kalløsø, T., Reinhardt, R., Chourrout, D., and Tompson, E. M. (2004). Spliced-leader RNA trans splicing in a chordate, *Oikopleura dioica*, with a compact genome. *Mol. Cell Biol.* **24**, 7795–7805.
- Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D. L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.* **15**, 71–78.
- Goodall, G. J., and Filipowicz, W. (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* **58**, 473–483.
- Graber, J. H., Salisbury, J., Hutchins, L. N., and Blumenthal, T. (2007). *C. elegans* sequences that control trans-splicing and operon pre-mRNA processing. *RNA* **13**, 1409–1426.
- Guiliano, D. B., and Blaxter, M. L. (2006). Operon conservation and the evolution of trans-splicing in the phylum *Nematoda*. *PLoS Genet.* **2**, e198.
- Haenni, S., Sharpe, H. E., Gravato Nobre, M., Zechner, K., Browne, C., Hodgkin, J., and Furger, A. (2009). Regulation of transcription termination in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res.* **37**, 6723–6736.
- Hajarnavis, A., Korf, I., and Durbin, R. (2004). A probabilistic model of 3' end formation in *Caenorhabditis elegans*. *Nucleic Acids Res.* **32**, 3392–3399.
- Hannon, G. J., Maroney, P. A., Ayers, D. G., Shambaugh, J. D., and Nilsen, T. W. (1990). Transcription of a nematode trans-spliced leader RNA requires internal elements for both initiation and 3' end-formation. *EMBO J.* **9**, 1915–1921.
- Hannon, G. J., Maroney, P. A., and Nilsen, T. W. (1991). U small nuclear ribonucleoprotein requirements for nematode cis- and trans-splicing *in vitro*. *J. Biol. Chem.* **266**, 22792–22795.
- Hannon, G. J., Maroney, P. A., Yu, Y. T., Hannon, G. E., and Nilsen, T. W. (1992). Interaction of U6 snRNA with a sequence required for function of the nematode SL RNA in trans-splicing. *Science* **258**, 1775–1780.
- Hiller, M., Huse, K., Szafranski, K., Jahn, N., Hampe, J., Schreiber, S., Backofen, R., and Platzer, M. (2004). Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. *Nat. Genet.* **36**, 1255–1257.
- Hiller, M., Nikolajewa, S., Huse, K., Szafranski, K., Rosenstiel, P., Schuster, S., Backofen, R., and Platzer, M. (2007). TassDB: A database of alternative tandem splice sites. *Nucleic Acids Res.* **35**, D188–D192.
- Hillier, L. W., Reinke, V., Green, P., Hirst, M., Marra, M. A., and Waterston, R. H. (2009). Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res.* **19**, 657–666.
- Hollins, C., Zorio, D. A. R., MacMorris, M., and Blumenthal, T. (2005). U2AF binding selects for the high conservation of the *C. elegans* 3' splice site. *RNA* **11**, 248–253.
- Huang, P., Pleasance, E. D., Maydan, J. S., Hunt-Newbury, R., O'Neil, N. J., Mah, A., Baillie, D. L., Marra, M. A., Moerman, D. G., and Jones, S. J. M. (2007). Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. *Genome Res.* **17**, 1478–1485.
- Huang, T., Kuersten, S., Deshpande, A. M., Spieth, J., MacMorris, M., and Blumenthal, T. (2001). Intercistronic region required for polycistronic pre-mRNA processing in *Caenorhabditis elegans*. *Mol. Cell Biol.* **21**, 1111–1120.
- Huang, T., Vilardell, J., and Query, C. C. (2002). Pre-spliceosome formation in *S. pombe* requires a stable complex of SF1-U2AF(59)-U2AF(23). *EMBO J.* **21**, 5516–5526.

- Huang, X. Y., and Hirsh, D. (1989). A second trans-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. U. S. A.* **86**, 8640–8644.
- Jacob, M., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356.
- Jones, T. A., Otto, W., Marz, M., Eddy, S. R., and Stadler, P. F. (2009). A survey of nematode SmY RNAs. *RNA Biol.* **6**, 5–8.
- Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* **23**, 616–626.
- Kay, R. J., Russnak, R. H., Jones, D., Mathias, C., and Candido, E. P. (1987). Expression of intron-containing *C. elegans* heat shock genes in mouse cells demonstrates divergence of 3' splice site recognition sequences between nematodes and vertebrates, and an inhibitory effect of heat shock on the mammalian splicing apparatus. *Nucleic Acids Res.* **15**, 3723–3741.
- Keiper, B. D., Lamphear, B. J., Deshpande, A. M., Jankowska-Anyszka, M., Aamodt, E. J., Blumenthal, T., and Rhoads, R. E. (2000). Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. *J. Biol. Chem.* **275**, 10590–10596.
- Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedeá, E., Greenblatt, J. F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522.
- Kramer, J. M., French, R. P., Park, E. C., and Johnson, J. J. (1990). The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. *Mol. Cell Biol.* **10**, 2081–2089.
- Krause, M., and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell* **49**, 753–761.
- Kuersten, S., Lea, K., MacMorris, M., Spieth, J., and Blumenthal, T. (1997). Relationship between 3' end formation and SL2-specific trans-splicing in polycistronic *Caenorhabditis elegans* pre-mRNA processing. *RNA* **3**, 269–278.
- Kuhn, A. N., and Käufer, N. F. (2003). Pre-mRNA splicing in *Schizosaccharomyces pombe*: Regulatory role of a kinase conserved from fission yeast to mammals. *Curr. Genet.* **42**, 241–251.
- Kuwabara, P. E., Okkema, P. G., and Kimble, J. (1992). *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway. *Mol. Biol. Cell* **3**, 461–473.
- Lasda, E. L., Allen, M. A., and Blumenthal, T. (2010). Polycistronic pre-mRNA processing *in vitro*: snRNP and pre-mRNA role reversal in trans-splicing. *Genes Dev.* **24**, 1645–1658.
- Lee, K., and Sommer, R. J. (2003). Operon structure and trans-splicing in the nematode *Pristionchus pacificus*. *Mol. Biol. Evol.* **20**, 2097–2103.
- Legrand, P., Pinaud, N., Minvielle-Sébastien, L., and Fribourg, S. (2007). The structure of the CstF-77 homodimer provides insights into CstF assembly. *Nucleic Acids Res.* **35**, 4515–4522.
- Liou, R. F., and Blumenthal, T. (1990). trans-spliced *Caenorhabditis elegans* mRNAs retain trimethyl-guanosine caps. *Mol. Cell Biol.* **10**, 1764–1768.
- Liu, C., Oliveira, A., Chauhan, C., Ghedin, E., and Unnasch, T. R. (2010). Functional analysis of putative operons in *Brugia malayi*. *Int. J. Parasitol.* **40**, 63–71.
- Liu, Y., Kuersten, S., Huang, T., Larsen, A., MacMorris, M., and Blumenthal, T. (2003). An uncapped RNA suggests a model for *Caenorhabditis elegans* polycistronic pre-mRNA processing. *RNA* **9**, 677–687.
- Logan, J., Falck-Pedersen, E., Darnell, J. E., and Shenk, T. (1987). A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc. Natl Acad. Sci. U. S. A.* **84**, 8306–8310.
- Longman, D., Johnstone, I. L., and Cáceres, J. F. (2000). Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *EMBO J.* **19**, 1625–1637.
- Luo, W., Johnson, A. W., and Bentley, D. L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev.* **20**, 954–965.

- MacMorris, M., Kumar, M., Lasda, E., Larsen, A., Kraemer, B., and Blumenthal, T. (2007). A novel family of *C. elegans* snRNPs contains proteins associated with trans-splicing. *RNA* **13**, 511–520.
- Mandel, C. R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cell Mol. Life Sci.* **65**, 1099–1122.
- Mandel, C. R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J. L., and Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**, 953–956.
- Maroney, P. A., Hannon, G. J., Denker, J. A., and Nilsen, T. W. (1990). The nematode spliced leader RNA participates in trans-splicing as an Sm snRNP. *EMBO J.* **9**, 3667–3673.
- Maroney, P. A., Hannon, G. J., Shambaugh, J. D., and Nilsen, T. W. (1991). Intramolecular base pairing between the nematode spliced leader and its 5' splice site is not essential for trans-splicing *in vitro*. *EMBO J.* **10**, 3869–3875.
- Maroney, P. A., Romfo, C. M., and Nilsen, T. W. (2000). Functional recognition of 5' splice site by U4/U6. U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol. Cell* **6**, 317–328.
- Maroney, P. A., Yu, Y. T., Jankowska, M., and Nilsen, T. W. (1996). Direct analysis of nematode cis- and trans-spliceosomes: A functional role for U5 snRNA in spliced leader addition trans-splicing and the identification of novel Sm snRNPs. *RNA* **2**, 735–745.
- de la Mata, M., and Kornblihtt, A. R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat. Struct. Mol. Biol.* **13**, 973–980.
- Mazroui, R., Puoti, A., and Krämer, A. (1999). Splicing factor SF1 from *Drosophila* and *Caenorhabditis*: Presence of an N-terminal RS domain and requirement for viability. *RNA* **5**, 1615–1631.
- McCarthy, J. E., and Gualerzi, C. (1990). Translational control of prokaryotic gene expression. *Trends Genet.* **6**, 78–85.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**, 357–361.
- Merendino, L., Guth, S., Bilbao, D., Martínez, C., and Valcárcel, J. (1999). Inhibition of *msl-2* splicing by sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402**, 838–841.
- Miguel-Aliaga, I., Culetto, E., Walker, D. S., Baylis, H. A., Sattelle, D. B., and Davies, K. E. (1999). The *Caenorhabditis elegans* orthologue of the human gene responsible for spinal muscular atrophy is a maternal product critical for germline maturation and embryonic viability. *Hum. Mol. Genet.* **8**, 2133–2143.
- Morton, J. J., and Blumenthal, T. (2011). Identification of transcriptional start sites of trans-spliced genes: Uncovering unusual operon arrangements. *RNA* **17**, 327–337.
- Muhich, M. L., and Boothroyd, J. C. (1988). Polycistronic transcripts in trypanosomes and their accumulation during heat shock: Evidence for a precursor role in mRNA synthesis. *Mol. Cell Biol.* **8**, 3837–3846.
- Nilsen, T. W. (1994). Unusual strategies of gene expression and control in parasites. *Science* **264**, 1868–1869.
- Ogg, S. C., Anderson, P., and Wickens, M. P. (1990). Splicing of a *C. elegans* myosin pre-mRNA in a human nuclear extract. *Nucleic Acids Res.* **18**, 143–149.
- Osbourne, A. E., and Field, B. (2009). Operons. *Cell Mol. Life Sci.* **66**, 3755–3775.
- Park, Y. S., and Kramer, J. M. (1990). Tandemly duplicated *Caenorhabditis elegans* collagen genes differ in their modes of splicing. *J. Mol. Biol.* **211**, 395–406.
- Patel, S. B., and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* **36**, 6482–6493.
- Pettitt, J., Müller, B., Stansfield, I., and Connolly, B. (2008). Spliced leader trans-splicing in the nematode *Trichinella spiralis* uses highly polymorphic, noncanonical spliced leaders. *RNA* **14**, 760–770.
- Pouchkina-Stantcheva, N. N., and Tunnacliffe, A. (2005). Spliced leader RNA-mediated trans-splicing in phylum *Rotifera*. *Mol. Biol. Evol.* **22**, 1482–1489.

- Qian, W., and Zhang, J. (2008). Evolutionary dynamics of nematode operons: Easy come, slow go. *Genome Res.* **18**, 412–421.
- Reinke, V., and Cutter, A. D. (2009). Germline expression influences operon organization in the *Caenorhabditis elegans* genome. *Genetics* **181**, 1219–1228.
- Rino, J., and Carmo-Fonseca, M. (2009). The spliceosome: A self-organized macromolecular machine in the nucleus? *Trends Cell Biol.* **19**, 375–384.
- Rosonina, E., Kaneko, S., and Manley, J. L. (2006). Terminating the transcript: Breaking up is hard to do. *Genes Dev.* **20**, 1050–1056.
- Ross, L. H., Freedman, J. H., and Rubin, C. S. (1995). Structure and expression of novel spliced leader RNA genes in *Caenorhabditis elegans*. *J. Biol. Chem.* **270**, 22066–22075.
- Rushforth, A. M., and Anderson, P. (1996). Splicing removes the *Caenorhabditis elegans* transposon Tc1 from most mutant pre-mRNAs. *Mol. Cell Biol.* **16**, 422–429.
- Rushforth, A. M., Saari, B., and Anderson, P. (1993). Site-selected insertion of the transposon Tc1 into a *Caenorhabditis elegans* myosin light chain gene. *Mol. Cell Biol.* **13**, 902–910.
- Ruskin, B., Greene, J. M., and Green, M. R. (1985). Cryptic branch point activation allows accurate *in vitro* splicing of human beta-globin intron mutants. *Cell* **41**, 833–844.
- Rutz, B., and Séraphin, B. (1999). Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA* **5**, 819–831.
- Salisbury, J., Hutchison, K. W., and Graber, J. H. (2006). A multispecies comparison of the metazoan 3'-processing downstream elements and the CstF-64 RNA recognition motif. *BMC Genomics* **7**, 55.
- Séraphin, B., Kretzner, L., and Rosbash, M. (1988). A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J.* **7**, 2533–2538.
- Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A. R., and Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* **34**, 3955–3967.
- Soller, M. (2006). Pre-messenger RNA processing and its regulation: A genomic perspective. *Cell Mol. Life Sci.* **63**, 796–819.
- Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993). Operons in *C. elegans*: Polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell* **73**, 521–532.
- Spieth, J., and Lawson, D. (2006). Overview of gene structure (January 18, 2006). In (WormBook, ed.). The *C. elegans* Research Community (Wormbook, doi/10.1895/wormbook.1.65.1). Available at: <http://www.wormbook.org>.
- Stover, N. A., and Steele, R. E. (2001). Trans-spliced leader addition to mRNAs in a cnidarian. *Proc. Natl Acad. Sci. U. S. A.* **98**, 5693–5698.
- Sutton, R. E., and Boothroyd, J. C. (1986). Evidence for trans splicing in trypanosomes. *Cell* **47**, 527–535.
- The *C. elegans* Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012–2018.
- Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: A role for the *nrd1-nab3* pathway in genome surveillance. *Mol. Cell* **23**, 853–864.
- Thomas, J., Lea, K., Zucker-Aprison, E., and Blumenthal, T. (1990). The spliceosomal snRNAs of *Caenorhabditis elegans*. *Nucleic Acids Res.* **18**, 2633–2642.
- Thomas, J. D., Conrad, R. C., and Blumenthal, T. (1988). The *C. elegans* trans-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. *Cell* **54**, 533–539.
- Van Doren, K., and Hirsh, D. (1990). mRNAs that mature through trans-splicing in *Caenorhabditis elegans* have a trimethylguanosine cap at their 5' termini. *Mol. Cell Biol.* **10**, 1769–1772.
- Van Doren, K., and Hirsh, D. (1988). Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. *Nature* **335**, 556–559.
- Vandenbergh, A. E., Meedel, T. H., and Hastings, K. E. (2001). mRNA 5'-leader trans-splicing in the chordates. *Genes Dev.* **15**, 294–303.

- Wachtel, C., and Manley, J. L. (2009). Splicing of mRNA precursors: The role of RNAs and proteins in catalysis. *Mol. BioSyst.* **5**, 311–316.
- Wahl, M. C., Will, C. L., and Lührmann, R. (2009). The spliceosome: Design principles of a dynamic RNP machine. *Cell* **136**, 701–718.
- Wallace, A., Filbin, M. E., Veo, B., McFarland, C., Stepinski, J., Jankowska-Anyszka, M., Darzynkiewicz, E., and Davis, R. E. (2010). The nematode eIF4E/G complex works with a trans-spliced leader stem loop to enable efficient translation of trimethylguanosine-capped RNAs. *Mol. Cell Biol.* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20154140> [Accessed March 5, 2010].
- Wang, Q., Zhang, L., Lynn, B., and Rymond, B. C. (2008). A BBP–Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast. *Nucleic Acids Res.* **36**, 2787–2798.
- West, S., Gromak, N., and Proudfoot, N. J. (2004). Human 5' → 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522–525.
- Whittle, C. M., McClinic, K. N., Ercan, S., Zhang, X., Green, R. D., Kelly, W. G., and Lieb, J. D. (2008). The genomic distribution and function of histone variant HTZ-1 during *C. elegans* embryogenesis. *PLoS Genet.* **4**, e1000187.
- Wickens, M., and Stephenson, P. (1984). Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science* **226**, 1045–1051.
- Wieringa, B., Hofer, E., and Weissmann, C. (1984). A minimal intron length but no specific internal sequence is required for splicing the large rabbit beta-globin intron. *Cell* **37**, 915–925.
- Will, C. L., and Lührmann, R. (2005). Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol. Chem.* **386**, 713–724.
- Williams, C., Xu, L., and Blumenthal, T. (1999). SL1 trans splicing and 3'-end formation in a novel class of *Caenorhabditis elegans* operon. *Mol. Cell Biol.* **19**, 376–383.
- Wong, S., and Wolfe, K. H. (2005). Birth of a metabolic gene cluster in yeast by adaptive gene relocation. *Nat. Genet.* **37**, 777–782.
- WormBase and WS210 Wormbase, WormBase, WS210 Wormbase – Home Page. Available at: <http://www.wormbase.org/>.
- Wu, S., Romfo, C. M., Nilsen, T. W., and Green, M. R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **402**, 832–835.
- Xie, H., and Hirsh, D. (1998). *In vivo* function of mutated spliced leader RNAs in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. U. S. A.* **95**, 4235–4240.
- Zaslaver, A., Baugh, L. R., and Sternberg, P. W. (2011). Metazoan operons accelerate recovery from growth-arrested states. *Cell* **145**, 981–992.
- Zhang, H., and Blumenthal, T. (1996). Functional analysis of an intron 3' splice site in *Caenorhabditis elegans*. *RNA* **2**, 380–388.
- Zorio, D. A., and Blumenthal, T. (1999a). Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **402**, 835–838.
- Zorio, D. A., and Blumenthal, T. (1999b). U2AF35 is encoded by an essential gene clustered in an operon with RRM/cyclophilin in *Caenorhabditis elegans*. *RNA* **5**, 487–494.
- Zorio, D. A., Lea, K., and Blumenthal, T. (1997). Cloning of *Caenorhabditis* U2AF65: An alternatively spliced RNA containing a novel exon. *Mol. Cell Biol.* **17**, 946–953.

---

---

## CHAPTER 8

# Analysis of microRNA Expression and Function

**Priscilla M. Van Wynsberghe**<sup>\*,‡,1</sup>, **Shih-Peng Chan**<sup>†,¶,1</sup>,  
**Frank J. Slack**<sup>†,2</sup> and **Amy E. Pasquinelli**<sup>\*,2</sup>

<sup>\*</sup>Department of Biology, University of California San Diego, La Jolla, California, USA

<sup>†</sup>Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA

<sup>‡</sup>Current address: Department of Biology, Colgate University, Hamilton, NY, USA

<sup>¶</sup>Current address: Department of Microbiology, College of Medicine, National Taiwan University, Taipei, Taiwan

---

### Abstract

#### I. Introduction

#### II. Rationale, Methods, and Materials

A. Worm Collection and Preparation for RNA Isolation

B. RNA Isolation Method

C. Analysis of Primary miRNAs by Agarose Northern Blotting

D. Analysis of Precursor and Mature miRNAs by PAGE Northern Blotting

E. Analysis of Primary, Precursor, and Mature miRNAs by qRT-PCR

F. 5' and 3' RACE to Identify Ends of Primary Transcripts

G. Modified RACE to Map Cleavage Sites in miRNA Processing

H. Small RNA Purification

I. cDNA Library Preparation

J. Hybridization, Cluster Generation, and Sequencing

K. Computational Data Analysis

L. RIP

M. Gel Filtration Analysis of miRNPs

N. EMSA of miRNPs

O. Affinity Purification of miRNPs by Biotinylated 2'-O-methyl Oligonucleotides

P. miRNA Promoter-Reporter Fusion

Q. Reporter-miRNA Target 3'UTR Fusion

R. Transformation

#### III. Summary

Acknowledgments

References

<sup>1</sup> Equal joint contribution.

<sup>2</sup> Co-corresponding authors.

---

---

## Abstract

Originally discovered in *C. elegans*, microRNAs (miRNAs) are small RNAs that regulate fundamental cellular processes in diverse organisms. MiRNAs are encoded within the genome and are initially transcribed as primary transcripts that can be several kilobases in length. Primary transcripts are successively cleaved by two RNase III enzymes, Drosha in the nucleus and Dicer in the cytoplasm, to produce ~70 nucleotide (nt) long precursor miRNAs and 22 nt long mature miRNAs, respectively. Mature miRNAs regulate gene expression post-transcriptionally by imperfectly binding target mRNAs in association with the multiprotein RNA induced silencing complex (RISC). The conserved sequence, expression pattern, and function of some miRNAs across distinct species as well as the importance of specific miRNAs in many biological pathways have led to an explosion in the study of miRNA biogenesis, miRNA target identification, and miRNA target regulation. Many advances in our understanding of miRNA biology have come from studies in the powerful model organism *C. elegans*. This chapter reviews the current methods used in *C. elegans* to study miRNA biogenesis, small RNA populations, miRNA–protein complexes, and miRNA target regulation.

---

---

## I. Introduction

microRNAs (miRNAs) play a major role in regulating many important processes including cellular differentiation and proliferation, developmental timing, hematopoiesis, immune responses, apoptosis, and nervous system patterning (Fineberg *et al.*, 2009; Gangaraju and Lin, 2009; Latronico and Condorelli, 2009; Negrini *et al.*, 2009; O’Connell *et al.*, 2010; Subramanian and Steer, 2010). Consequently, alterations in miRNA levels, timing of expression, or target recognition can have devastating consequences including cancer (Kai and Pasquinelli, 2010; Medina and Slack, 2008).

Most miRNAs are transcribed by RNA polymerase II as independent transcripts or as RNAs embedded within introns of protein-coding messenger RNAs (Davis and Hata, 2009). These miRNA transcripts, called primary miRNAs, are capped, polyadenylated, and can be several thousand nucleotides (nt) long (Davis and Hata, 2009). Primary miRNAs are successively cleaved into ~70 nt hairpin precursor miRNAs and then to 22 nt mature miRNAs by two RNase III enzymes respectively called Drosha and Dicer (DRSH-1 and DCR-1 in *C. elegans*) (Davis and Hata, 2009). Mature miRNAs guide the argonaute protein ALG-1/2 to a specific target mRNA (Grishok *et al.*, 2001; Okamura *et al.*, 2004). In animals, the mature miRNA is imperfectly complementary to the target site and though most miRNA-ALG-1/2 complexes bind to the 3’UTR, target binding can also occur in the 5’UTR, exons, and introns (Davis and Hata, 2009; Zisoulis *et al.*, 2010). ALG-1/2 associates with the RISC to regulate gene expression by inhibiting translation or triggering degradation of the target mRNA (Chekulaeva and Filipowicz, 2009).

Originally discovered in *C. elegans* as genes that regulate developmental timing, miRNAs regulate fundamental processes in diverse organisms and are often conserved between species (Lee *et al.*, 1993; Pasquinelli *et al.*, 2000; Reinhart *et al.*, 2000). This conservation, as well as the availability of miRNA mutants makes *C. elegans* an important system to study the mechanisms of miRNA biogenesis, target identification, and function (Kato and Slack, 2008).

This chapter details current methods used to isolate miRNA species, analyze primary, precursor, and mature miRNA expression, map the 5' and 3' ends of miRNA primary and cleavage products, identify small RNA populations, analyze miRNA–protein complexes, and construct reporters and transgenic strains to analyze miRNA target regulation. Though specific for the analysis of small RNAs, these methods can also be adapted to study mRNA expression or mRNA–protein interactions.

---

---

---

## II. Rationale, Methods, and Materials

- I. Total RNA isolation
- II. Analysis of primary, precursor, and mature miRNA expression
- III. RACE mapping of miRNA primary and cleavage products
- IV. Deep sequencing of small RNA populations
- V. Analysis of miRNP complexes
- VI. Construction of transgenic strains to analyze miRNA expression and target regulation

### I. Total RNA Isolation

Though standard protocols can be used for RNA isolation in *C. elegans*, to ensure isolation of high-quality RNAs of all sizes, including small 22 nt mature miRNAs, the following double-RNA-extraction method is preferred. This method is based upon the Trizol RNA isolation protocol (Invitrogen) and can be used to isolate RNA from embryos or larval and adult stage worms. RNA generated from this procedure can be used for any application, including those described in Sections II to V.

#### A. Worm Collection and Preparation for RNA Isolation

For analysis of miRNA expression at specific developmental time points, plate synchronized larval stage 1 (L1) worms on the appropriate bacterial food source and wash worms off plates with M9 at the desired time point. Rock worms in M9 (22 mM KH<sub>2</sub>PO<sub>4</sub>, 42 mM Na<sub>2</sub>HPO<sub>4</sub>, 85.5 mM NaCl, 1 mM MgSO<sub>4</sub>) in 15 mL conical tubes for 20 min to allow digestion of bacterial food. Pellet worms with centrifugation and transfer to 1.5 mL tubes. Pellet again, remove excess liquid, and snap-freeze in a dry ice–ethanol bath. Worm pellets of 50–100  $\mu$ L will yield at least 50  $\mu$ g of total RNA.

## B. RNA Isolation Method

Add ~100  $\mu\text{L}$  of Trizol (Invitrogen) to the frozen worm pellet for a total volume of 200  $\mu\text{L}$ . Grind worms in Trizol with a hand-held electric homogenizer that fits 1.5 mL tubes until the solution is cloudy and no clumps are visible. Add 800  $\mu\text{L}$  of Trizol to lysed sample, mix, and let sit at room temperature (RT) for at least 5 min. After all samples have been homogenized as above, add 200  $\mu\text{L}$  of chloroform: isoamyl alcohol (24:1) to each sample, shake for 30 s, and let sit at RT for 3 min. Spin samples at 12,000g for 15 min at 4 °C. Transfer the RNA-containing supernatant to an equal volume of chloroform:isoamyl alcohol. Shake to mix, and spin at top speed for 5 min at RT. Transfer the supernatant to an equal volume of isopropanol and 1  $\mu\text{L}$  of 20 mg/mL glycogen. Mix samples briefly and incubate at  $-20$  °C for at least 1 h, which is important for efficiently precipitating small RNAs. Spin samples at 12,000 g for 10 min at 4 °C. Carefully remove the supernatant and add 1 mL of 75% ethanol to the RNA pellet. Spin samples at 12,000 g for 5 min at 4 °C. Carefully remove all of the supernatant. Briefly dry pellets in a 50–58 °C oven for ~5 min or at RT for ~10 min. Resuspend pellets in 200  $\mu\text{L}$  of DEPC water preheated to ~68 °C. It is very important to thoroughly resuspend the RNA at this step and ensure that the pellet has gone into solution entirely.

Next, a second RNA-extraction step is performed to further purify the RNA. If samples are to be analyzed by RT-PCR, a DNase treatment can be performed as described below in Section II before continuing with the second RNA extraction. Add 20  $\mu\text{L}$  of 3 M NaOAc pH 5.2 and 220  $\mu\text{L}$  of phenol:chloroform:isoamyl alcohol (25:24:1) to each RNA sample. Vortex briefly to mix, and spin at top speed for 5 min at RT. Transfer supernatant to 220  $\mu\text{L}$  of chloroform:isoamyl alcohol (24:1). Vortex briefly and spin at top speed for 5 min at RT. Transfer supernatant to 220  $\mu\text{L}$  of isopropanol and 1  $\mu\text{L}$  of 20 mg/mL glycogen. Vortex samples briefly and precipitate by incubating at  $-20$  °C for at least 1 h. Spin samples at 12,000 g for 10 min at 4 °C. Carefully remove the supernatant and add 1 mL of 75% ethanol to the pellet. Spin samples at 12,000 g for 5 min at 4 °C. Carefully remove all of the supernatant. Briefly, dry pellets in a 50–58 °C oven for ~5 min or at RT for ~10 min. Depending on the pellet size, resuspend pellets in 30–50  $\mu\text{L}$  of DEPC water preheated to ~68 °C. Determine RNA concentration by diluting RNA 100-fold in  $1 \times$  Tris EDTA (TE) buffer pH 7.4 and analyzing the A260 value by spectrophotometry. Pure RNA typically has an A260/A280 reading of 1.8–2. An A260/A230 ratio less than 2 or 1.5 signifies contamination by genomic DNA or extraction chemicals, respectively.

## II. Analysis of Primary, Precursor, and Mature miRNA Expression

Mature 22 nt miRNAs originate from genome-encoded primary miRNA (pri-miRNA) transcripts that are mostly transcribed by RNA polymerase II, capped, polyadenylated, and multiple kilobases in length (Davis and Hata, 2009). Cleavage of the pri-miRNA in the nucleus by the RNase III enzyme Drosha in association with Pasha produces the ~70 nt precursor miRNA (pre-miRNA) hairpin (Davis and Hata, 2009). Further cleavage by the RNase III enzyme Dicer in the

cytoplasm releases the 22 nt mature miRNA (Davis and Hata, 2009). Regulation of mature miRNA levels can be accomplished by controlling any step in the miRNA biogenesis pathway (Winter *et al.*, 2009). Though some control mechanisms ubiquitously affect miRNA expression, others are specific for particular miRNAs (Kai and Pasquinelli, 2010; Winter *et al.*, 2009). To gain a complete understanding of miRNA expression and determine the mechanisms controlling miRNA biogenesis, analysis of pri-, pre-, and mature miRNA levels is necessary.

### C. Analysis of Primary miRNAs by Agarose Northern Blotting

This method allows concurrent analysis of multiple pri-miRNA isoforms, depending on sufficient differences in their sizes. For example, the three pri-*let-7* transcripts or two pri-*lin-4* transcripts in *C. elegans* can be resolved by denaturing 1% agarose gels and detected by this northern blotting protocol (Bracht *et al.*, 2004; Bracht *et al.*, 2010; Van Wynsberghe *et al.*, 2011). The northern blotting protocol is comparable to that used to detect mRNAs except that the often low abundance of pri-miRNAs necessitates that at least 10  $\mu\text{g}$  of total RNA, prepared by the method described above in Section I, is used for each sample.

#### 1. Running and Transferring

Make a 1% agarose gel in MOPS/EDTA running buffer (40 mM MOPS pH 7.0, 2 mM EDTA). Prepare samples by adding 10  $\mu\text{g}$  of total RNA in a total volume of 10  $\mu\text{L}$  to 15  $\mu\text{L}$  of MOPS/EDTA running buffer and 5  $\mu\text{L}$  of 37% formaldehyde. Heat samples at 70°C for 10min and then transfer to ice for 1min. Add 6  $\mu\text{L}$  of 6 $\times$  glycerol loading buffer (30% glycerol, 0.25% bromphenol blue, 0.25% xylene cyanol) to each sample before loading. Run the gel at 80 to 125 V for at least 2 h. Visualization of the rRNA bands by ethidium bromide (EtBr) staining is used to confirm even loading and nondegraded RNA samples before proceeding with northern blotting. Rinse gel gently in sterile water with shaking. Equilibrate membrane (Bio-Rad Zeta Probe GT), cut to the same size as the gel, in ddH<sub>2</sub>O followed by 10 $\times$  SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0). Assemble the transferring setup at RT by layering: glass tray containing 10 $\times$  SSC, upside-down gel casting tray, two sheets of Whatman 3MM paper wetted in 10 $\times$  SSC with ends dipped in the 10 $\times$  SSC in the glass tray to act as wicks, gel with bottom side facing up, membrane, three pieces of Whatman 3MM paper wetted in 10 $\times$  SSC and cut to the size of the membrane, a stack of paper towels 5–10 cm high and cut to the size of the gel, and a light weight evenly stacked on top. Allow transfer to occur at least 6 h and up to overnight.

#### 2. Prehybridization, Probe Preparation, and Hybridization

Briefly rinse the membrane in 6 $\times$  SSC to remove any agarose particles. Place the side of the membrane that contacted the gel facing up on a piece of dry Whatman 3MM paper and cross-link at 1200  $\mu\text{J}$   $\times$  100 by using a Spectrolinker XL-1000 UV

Crosslinker. Sandwich the membrane between dry Whatman 3MM paper and bake in an oven at 80°C for 30 min. Incubate the membrane at 58–60°C with agitation for at least 2 h in a sealed bag containing 25–50 mL of prehybridization solution [ $5\times$  SSC, 7% SDS, 20 mM sodium phosphate, 0.1 mg/mL boiled, sheared salmon sperm DNA, and  $1\times$  Denhardt's solution (5 g of Ficoll, 5 g of polyvinylpyrrolidone, and 5 g of BSA in 500 mL of ddH<sub>2</sub>O)].

Multiple methods can be used to make a radioactive probe to target the pri-miRNA of interest. A simple, common method is based on the Prime-it II Random Primer Labeling Kit (Stratagene), which requires only a gel-purified, 500–1500 bp PCR product of the desired target and 80  $\mu$ Ci of alpha-<sup>32</sup>P dATP (6000 Ci/mmol). Add probe to 25–50 mL of fresh prehybridization solution. Remove the original prehybridization solution from the blot container before adding the fresh probe solution. Incubate blot in probe solution at 58–60 °C with agitation for at least 4 h or up to overnight.

### 3. Washing, Viewing, and Stripping

Probe solution can be stored at –20°C and reused. Remove the blot from the bag and transfer to a container with a fitted lid. Wash the blot twice at 58–60°C for 10 min in high salt wash solution ( $3\times$  SSC,  $10\times$  Denhardt's solution, 5% SDS, and 25 mM sodium phosphate), flipping the blot after 5 min for each wash. Wash the blot once at 58–60°C for 10 min in low salt wash solution ( $0.5\times$  SSC and 1.5% SDS). Repeat wash in low salt wash solution if the radioactive signal on the blot is too high or not specific. Wrap membrane in plastic wrap and expose to film or a phosphorimager screen. To strip the blot, incubate membrane in boiling 0.1% SDS with agitation. Repeat until radioactive signal is no longer detected.

## D. Analysis of Precursor and Mature miRNAs by PAGE Northern Blotting

To ensure detection of pre- and mature miRNAs at least 20  $\mu$ g of total RNA, prepared by the method described above in Section I, is recommended for each sample.

### 1. Running and Transferring

To visualize both pre- and mature miRNAs, make an 11% denaturing, urea PAGE gel (Sequagel, National Diagnostics). Pre-run PAGE gel in  $0.5\times$  Tris Borate EDTA (TBE) buffer at 150 V for 30 min. Prepare RNA samples by mixing 20  $\mu$ g of total RNA in a final volume of 15  $\mu$ L with an equal volume of  $1\times$  formamide loading dye (10 mM EDTA, 80% deionized formamide, and 1 mg/mL xylene cyanol and bromphenol blue dye mix). Heat RNA samples at 65°C for 10 min before loading. Run gel at 150 V until the bromphenol blue dye is at the bottom of the gel, ~1.5 h for 10 cm  $\times$  10 cm gel size. Stain PAGE gel with EtBr to visualize rRNA bands and confirm even

loading and nondegraded RNA samples. Equilibrate membrane (Zeta-Probe GT, Bio-Rad) cut to the same size as the gel in water followed by  $0.5\times$  TBE. Assemble transfer apparatus by layering: plastic holder, sponge pad wetted in  $0.5\times$  TBE, two sheets of Whatman 3MM paper cut to the size of the gel and wetted in  $0.5\times$  TBE, gel, membrane, two sheets of Whatman 3MM paper prepared as before, sponge pad wetted in  $0.5\times$  TBE, and the plastic holder. Transfer at 200 mA at  $4^{\circ}\text{C}$  for 2–4 h. In our experience, wet transfer methods are much more reliable than semidry dry methods for transfer of small RNAs for northern blotting analyses.

## 2. Prehybridization, Probe Preparation, and Hybridization

Place membrane RNA side up on a piece of dry Whatman 3MM paper and cross-link at  $1200\ \mu\text{J}\times 100$  by using a Spectrolinker SL-1000 UV Crosslinker. Sandwich membrane between dry Whatman 3MM paper and bake in an oven at  $80^{\circ}\text{C}$  for 30 min. Depending on the miRNA GC content, incubate blot at  $40\text{--}50^{\circ}\text{C}$  with agitation for at least 2 h in a sealed bag containing 25 mL prehybridization solution.

Multiple methods can be used to make a radioactive probe to target the mature miRNA of interest. A simple, common method is based on the microRNA Starfire Kit (IDT DNA) that requires an oligo complementary to the miRNA of interest and containing the proprietary starfire sequence (IDT DNA) and  $80\ \mu\text{Ci}$  of alpha- $^{32}\text{P}$  dATP (6000 Ci/mmol). For more abundant miRNAs, a probe can be made by labeling a DNA oligo complementary to the miRNA of interest. To do this incubate 25 pmol of the DNA oligo with  $\gamma\text{-}^{32}\text{P}$ ATP (6000 Ci/mmol), T4 polynucleotide kinase (PNK) buffer, and T4 PNK enzyme at  $37^{\circ}\text{C}$  for 1 h. Add the probe to 25 mL of fresh prehybridization solution. Remove the original prehybridization solution from the blot container before adding the fresh probe solution. Incubate the blot in probe solution at  $40\text{--}50^{\circ}\text{C}$  with agitation for at least 4 h or up to overnight.

## 3. Washing, Viewing, and Stripping

Probe solution can be stored at  $-20^{\circ}\text{C}$  and reused. Blots should be washed, viewed, and stripped as described above for agarose northern blotting.

## E. Analysis of Primary, Precursor, and Mature miRNAs by qRT-PCR

This analysis allows detection of miRNAs from a small amount of starting material, but does not distinguish between differently sized primary miRNA isoforms, or determine the origin of pre-miRNA since pre-miRNA is contained within pri-miRNA (Bracht *et al.*, 2010; Van Wynsberghe *et al.*, 2011). Inclusion of a control sample that lacks reverse transcriptase will test if any amplification is caused by genomic DNA. SYBR Green or Taqman qRT-PCR can be performed. Taqman qRT-PCR is more specific than SYBR Green qRT-PCR because of the additional probe, but it is also more costly. qRT-PCR can be performed in one step or in two steps of

cDNA synthesis and qPCR. Both methods are detailed below. Reactions should be done in duplicate or triplicate to reduce erroneous results due to pipetting error. A nontemplate control reaction should also be included and a dissociation curve should be calculated to determine if nonspecific primer interactions occur.

### 1. Primer Selection

Primary miRNAs can be amplified by selecting primers near to but not included in the pre-miRNA hairpin, while precursor miRNAs can be amplified by selecting primers within the hairpin. Since all pre-miRNAs are contained within pri-miRNAs, the origin of amplification occurring from pre-miRNA hairpin primers cannot be specifically assigned. However, comparison of signal from pre- and pri-miRNA primer sets can determine the amount of pre-miRNA signal derived from pri- or pre-miRNA (Van Wynsberghe *et al.*, 2011). Primers should have a  $T_m$  between 58 and 60°C, a GC content between 30% and 80%, and the 5 nt at the 3' end of the primer should have no more than two G or C nts. Amplicons should be less than 200 nt. Analysis of mature miRNAs by qPCR presents additional problems due to their small size, but can be accomplished by adding linkers to the mature miRNA (Benes and Castoldi, 2010). Below we detail methods for analyzing pri- and pre-miRNA levels by 1 and 2 step qPCR.

### 2. DNase Treatment

RNA prepared as described above in Section I is generally devoid of DNA contamination, but a DNase treatment of total RNA is recommended before RT-PCR. This step can be performed between the first Trizol RNA extraction and the second phenol extraction as described above in Section I. To the RNA resuspended in 200  $\mu\text{L}$  of DEPC water add 10  $\mu\text{L}$  of RQ1 DNase and 23  $\mu\text{L}$  of 10 $\times$  RQ1 DNase buffer (Promega). Incubate at 37°C for 1 h before adding 1  $\mu\text{L}$  of stop solution (Promega). Proceed with second RNA phenol extraction as described above in Section I.

### 3. One-Step qRT-PCR

The amount of total RNA and primers used in this reaction may need to be optimized depending on the abundance of the particular transcript being analyzed. In general 0.5  $\mu\text{g}$  of total RNA and 6.25 pmol of both forward and reverse primers are added to a minimum 20  $\mu\text{L}$  reaction volume containing 1 $\times$  RT Enzyme Mix, 1 $\times$  RT-PCR Mix, and the Taqman probe if performing a Taqman reaction (Applied Biosystems). Increasing the reaction volume will decrease pipetting error but increase the cost per reaction. Many companies sell kits or individual reagents for qRT-PCR including Applied Biosystems and Bio-Rad. Reactions are run in a qPCR thermocycler set to perform the following protocol: (1) 30 min at 48°C, (2) 10 min at 95°C, (3) 40 cycles of 15 s at 95°C and 1 min at 60°C, and (4) hold at 4°C.

#### 4. Two-Step qRT-PCR

Add 2.5 to 5  $\mu$ g of total, DNase-treated RNA to 30 pmol of oligo dT or random primers and 1  $\mu$ L of 10 mM dNTP in a 12  $\mu$ L reaction. Incubate at 65°C for 5 min followed by ice for 1 min. Add 4  $\mu$ L of 5 $\times$  buffer, 2  $\mu$ L of 0.1 M DTT, and 40 U of RNasin (Promega). Incubate at 42°C for 2 min before adding 200 U of Superscript III Reverse Transcriptase (Invitrogen). Incubate at 42°C for 50 min and 70°C for 15 min. Add 2 U of RNaseH (Invitrogen), and incubate at 37°C for 30 min. Dilute an aliquot of the cDNA reaction 1:5 for use in qPCR. As mentioned above the amount of cDNA and primers used in this reaction may need to be optimized depending on the abundance of the particular transcript being analyzed. In general, 1  $\mu$ L of cDNA diluted 1:5 and 6.25 pmol of both forward and reverse primers are added to a minimum 20  $\mu$ L reaction volume containing 1 $\times$  SYBR Green PCR Mastermix and the Taqman probe if performing a Taqman reaction (Applied Biosystems). Increasing the reaction volume will decrease pipetting error but increase the cost per reaction. Reactions are run in a qPCR thermocycler set to perform the following protocol: (1) 2 min at 48°C, (2) 10 min at 95°C, (3) 40 cycles of 15 s at 95°C and 1 min at 60°C, (4) hold at 4°C.

### III. RACE Mapping of miRNA Primary and Cleavage Products

Mature miRNAs originate from genome-encoded primary miRNA transcripts that are often multiple kilobases in length (Davis and Hata, 2009). Primary miRNA transcripts can be encoded within an intron in the same or opposite direction as its host gene, or transcribed independently of protein-coding genes (Davis and Hata, 2009). Some miRNAs are clustered together within the same primary miRNA transcript (Davis and Hata, 2009). Identification of primary miRNA transcripts defines miRNA clusters, allows further characterization of primary miRNA expression patterns both through proper promoter–fusion reporter analysis, as described below in Section VI, and primary miRNA analysis as described above in Section II. Identification of primary miRNA transcripts also enhances identification of *cis* and *trans* transcription elements. Here we describe RACE (Random Amplification of cDNA Ends) methods for the identification of 5' and 3' primary miRNA transcript ends (Bracht *et al.*, 2004; Bracht *et al.*, 2010).

Similar RACE methods can also be used to identify and characterize cleavage products of Drosha processing (Bracht *et al.*, 2010; Van Wynsberghe *et al.*, 2011). In doing so, this method provides a more sensitive method than northern blotting to determine if precursor miRNA is produced. This method identifies the site of Drosha processing, which can vary throughout development. Additionally, this method determines if 3' end modification, like uridylation, of precursor miRNA occurs. 5' or 3' Drosha cleavage fragments destined for degradation can also be identified.

#### F. 5' and 3' RACE to Identify Ends of Primary Transcripts

This method is adapted from the GeneRacer Kit (Invitrogen) and requires 1–5  $\mu$ g RNA isolated as described above in Section I. Dephosphorylate RNA by incubating

total RNA with 40 U of RNaseOut and 1 U of calf intestinal phosphatase (CIP) in 1× CIP buffer (Invitrogen) at 50°C for 1 h. Extract RNA as described above in the second RNA-extraction step of Section I. After precipitation, resuspend the RNA pellet in 7 μL of DEPC water. Decap the dephosphorylated RNA by adding 40 U of RNaseOut and 0.5 U of tobacco acid pyrophosphatase (TAP) to the RNA in 1× TAP buffer (Invitrogen). Incubate at 37°C for 1 h and extract RNA as described above. Resuspend the RNA in 7 μL of DEPC water after precipitation. Incubate the GeneRacer RNA Oligo (Invitrogen) with the dephosphorylated, decapped RNA at 65°C for 5 min and 4°C for 2 min. Add 40 U of RNaseOut, 5 U of T4 RNA ligase, 1 μL of 10 mM ATP, and 1 μL of 10× T4 ligase buffer. Incubate at 37°C for 1 h before extracting RNA as described above. Resuspend the ligated RNA in 10 μL of DEPC water after precipitation.

Next, the cDNA synthesis process is initiated by adding 1 μL of 50 μM GeneRacer oligo dT primer, 1 μL of 10 μM dNTP, and 1 μL of H<sub>2</sub>O to the ligated RNA and incubating at 65°C for 5 min and 4°C for 1 min. Add 4 μL of 5× First Strand Buffer, 1 μL of 0.1 M DTT, 1 μL of RNaseOut, and 1 μL of SuperScript III (Invitrogen). Incubate at 50°C for 1 h, 70°C for 15 min, and ice for 2 min. Add 2 U of RNase H and incubate at 37°C for 2 min.

To characterize the 5' end of the cDNA, amplify cDNA with the GeneRacer 5' primer (Invitrogen) and a gene-specific reverse primer 3' to the mature miRNA. Use PCR conditions appropriate for the primer melting temperatures and expected product sizes with 30 cycles of amplification in a 50 μL reaction using High Fidelity Taq (Invitrogen). Clean the PCR reaction using the Qiagen PCR column purification system and elute in 50 μL of ddH<sub>2</sub>O. Use 1 μL of the first PCR reaction to perform nested RACE, using a second set of primers corresponding to the GeneRacer 5' nested oligo (Invitrogen) and a gene-specific primer upstream of the original primer. To enable the identification of specific RACE products, also set up nested PCR reactions that contain the nested 5' and gene-specific primers individually. Use appropriate PCR conditions with 25–30 cycles of amplification in a 50 μL reaction using High Fidelity Taq (Invitrogen). Clean the reactions as described above and use half of the reaction (25 μL) to analyze the PCR products by separation in an agarose gel. Specific products should be present in the nested PCR using both primers but not in the control reactions using the single primers. It is not uncommon for the first PCR reaction to contain faint or multiple bands that do not correspond to the potentially specific bands in the nested RACE. Additional nested primers or diagnostic restriction digestions should be used to verify RACE products corresponding to the gene of interest before proceeding to cloning. If multiple products are present and only some appear specific, gel purification can be performed to eliminate nonspecific products before cloning. Otherwise, the PCR reaction can be used directly for cloning using the TOPO reagents (Invitrogen). Cloned DNA is transformed into bacteria, cultured, and the plasmids isolated by standard methods. Restriction digestion analyses should be used to verify inserts before subjecting clones for sequencing and identification of the primary miRNA 5' end(s).

Characterize the 3' end of the cDNA as described above for the 5' end cDNA analysis but use the GeneRacer 3' primer and 3' nested primer together with a forward gene-specific primer and a nested primer 5' to the mature miRNA.

### G. Modified RACE to Map Cleavage Sites in miRNA Processing

Drosha cleavage of miRNA primary transcripts releases the hairpin precursor miRNA, which contains 5' monophosphate and 3' hydroxyl groups, as well as upstream and downstream products that contain 3' hydroxyl and 5' monophosphate groups, respectively. Based on these chemical features, the products of Drosha processing can be analyzed by modification of the RACE methods (Bracht *et al.*, 2010; Van Wynsberghe *et al.*, 2011). Below we describe methods for analyzing the 3' end of precursor miRNA and the 5' end of 3' Drosha cleavage products. Similar methods can also be used to characterize the 3' end of the 5' Drosha cleavage product as well as the 5' end of precursor miRNA.

#### 1. Characterization of the pre-miRNA 3' End After Drosha Cleavage

Run 20  $\mu\text{g}$  of RNA, isolated as described above in Section I, on a 15% denaturing PAGE gel alongside 10 and 25 bp ladders. Stain the gel with EtBr and cut out a band of gel in the sample lane that corresponds to the appropriate size range. Shear gel by spinning through a 0.5 mL tube with multiple 21-gauge needle holes. Elute RNA by rocking sheared gel pieces in 500  $\mu\text{L}$  of 0.3 M NaCl at RT for 4 h. Transfer sample to a Spin-X Cellulose Acetate filter and spin at top speed for 2 min. Precipitate RNA with isopropanol, glycogen, and NaOAc, pH 5.2. Wash RNA pellet with 75% ethanol and resuspend the RNA pellet in 26  $\mu\text{L}$  of ddH<sub>2</sub>O. Dephosphorylate gel extracted RNA by incubating with 10 U of CIP in 1 $\times$  NEBuffer #3 (NEB) at 37°C for 1 h. Extract RNA as described above in the second RNA-extraction step of Section I. After precipitation, resuspend the RNA pellet in 3  $\mu\text{L}$  of DEPC water. Ligate dephosphorylated RNA to a 3' RNA linker with a 5' phosphate group and 3' puromycin tag by incubating RNA at 20°C overnight in a 10  $\mu\text{L}$  total reaction that contains: 1 $\times$  reaction buffer (Fermentas), 1 mM ATP, 0.1 mg/mL BSA, 20 U of RNasin (Promega), 10 U of T4 RNA ligase (Fermentas), and 5  $\mu\text{M}$  3' RNA linker. Extract RNA as described above in the second RNA-extraction step of Section I. After precipitation, resuspend RNA pellet in 20  $\mu\text{L}$  of ddH<sub>2</sub>O. Purify RNA by running on a 15% denaturing PAGE gel as described above. After gel extraction, precipitate RNA with isopropanol, glycogen, and NaOAc, pH 5.2. Resuspend RNA in 4  $\mu\text{L}$  of H<sub>2</sub>O. Reverse transcribe by adding 1.5  $\mu\text{L}$  of 10 mM oligo complementary to the 3' RNA linker and 0.5  $\mu\text{L}$  of 10 mM dNTP. Incubate at 65°C for 5 min and 4°C for 1 min. Add 2  $\mu\text{L}$  of 5 $\times$  buffer, 1  $\mu\text{L}$  of 0.1 M DTT, and 0.5  $\mu\text{L}$  of RNasin. Incubate at 42°C for 2 min. Add 0.5  $\mu\text{L}$  of Superscript III Reverse Transcriptase (Invitrogen) before incubating at 42°C for 50 min and 70°C for 15 min. Add 0.5  $\mu\text{L}$  of RNase H and incubate at 37°C for 30 min. To characterize the 3' end of the cleavage product,

amplify cDNA with a gene-specific forward primer and a reverse primer complementary to the 3' linker. Use PCR conditions appropriate for the primer melting temperatures and expected product sizes with 30 cycles of amplification in a 50  $\mu$ L reaction using High Fidelity Taq (Invitrogen). Clean the PCR reaction using the Qiagen PCR column purification system and elute in 50  $\mu$ L of ddH<sub>2</sub>O. Use 1  $\mu$ L of the first PCR reaction to perform a nested PCR, using a second set of primers corresponding to a gene-specific nested forward primer and a linker specific reverse nested primer. To identify specific PCR products, also set up nested PCR reactions that contain the gene-specific and 3' nested primers individually. Use appropriate PCR conditions with 25–30 cycles of amplification in a 50  $\mu$ L reaction using High Fidelity Taq (Invitrogen). Clean the reactions as described above and use half of the reaction (25  $\mu$ L) to analyze the PCR products by separation in an agarose gel. Specific products should be present in the nested PCR using both primers but not in the control reactions using the single primers. If multiple products are present, perform gel purification of the specific bands before cloning. Otherwise, the PCR reaction can be used directly for cloning using the TOPO reagents (Invitrogen). Cloned DNA is transformed into bacteria, cultured, and the plasmids isolated by standard methods. Restriction digestion analyses should be used to verify inserts before subjecting clones for sequencing and identification of the 3' cleavage site.

## 2. Characterize the 5' End of 3' Drosha Cleavage Products

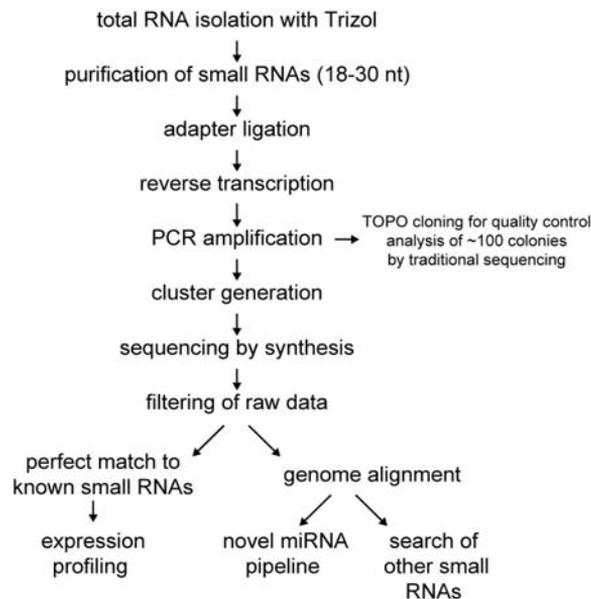
Incubate GeneRacer RNA Oligo with 1–5  $\mu$ g of total RNA, isolated as described above in Section I, at 65°C for 5 min and 4°C for 2 min. Add 40 U of RNaseOut, 5 U of T4 RNA ligase (Fermentas), 1  $\mu$ L of 10 mM ATP, and 1  $\mu$ L of 10 $\times$  T4 ligase buffer (Fermentas). Incubate at 37°C for 1 h before extracting RNA as described in the second RNA-extraction step of Section I. Resuspend ligated RNA in 10  $\mu$ L of DEPC water after precipitation. Reverse transcribe as described above with 1  $\mu$ L of 50  $\mu$ M gene-specific reverse primer. Amplify cDNA with GeneRacer 5' primer and a gene-specific reverse primer as described above. Clean the PCR reaction and perform a nested PCR with the GeneRacer 5' nested primer and a gene-specific reverse nested primer as described above. Analyze the nested PCR reaction, TOPO clone, and sequence as described above to identify the 5' cleavage site.

## IV. Deep Sequencing of Small RNA Populations

High-throughput “deep sequencing” technologies, such as 454 Life Science/Roche Genome Sequencer (Margulies *et al.*, 2005), Solexa/Illumina Genome Analyzer (Bentley *et al.*, 2008), and Applied Biosystems SOLiD System (for review see Mardis, 2008; Shendure and Ji, 2008), capable of simultaneously sequencing millions of molecules, have provided an excellent tool for small RNA profiling and discovery. Deep sequencing outperforms other existing profiling methods, such as array hybridization or qRT-PCR, in discriminating closely related RNAs, detecting 5'- or 3'-end variations and measuring absolute abundance with a better dynamic range. Furthermore, deep sequencing, unlike microarrays, allows the identification

of novel small RNAs, as it is not limited by *a priori* sequence information. For example, Ruby *et al.* (2006) applied 454 pyrosequencing to small RNAs expressed in mixed-staged *C. elegans* and identified the class of 21U-RNAs and additional previously unknown miRNAs (Ruby *et al.*, 2006). Our group also has used Illumina Genome Analyzer for profiling dynamic miRNA expression in *C. elegans* development and identifying novel miRNAs and 21U-RNAs (Kato *et al.*, 2009).

Protocols for creating cDNA libraries of small RNAs commonly include size fractionation of RNAs, 5'- and 3'-adaptor ligation, reverse transcription, and PCR amplification (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Lee and Ambros, 2001). These cDNA libraries are made amenable for deep sequencing by the inclusion of appropriate primer sequences. Commercially available kit systems for preparing cDNA libraries for deep sequencing also have been released. In the approaches described below, we follow the manufacturer's instruction in the Illumina's Small RNA Sample Prep Kit (catalog # FC-102-1009) for preparing cDNA libraries and discuss the principles of deep sequencing of *C. elegans* small RNA populations based on our approach using Illumina Genome Analyzer, summarized graphically in Fig. 1. These general principles are broadly applicable to other platforms.



**Fig. 1** An outline of deep sequencing of small RNA populations. See text for details.

## H. Small RNA Purification

Prepare 10  $\mu\text{g}$  of total RNA for each experimental condition (see Section I for total RNA isolation). The integrity of RNA can be evaluated by separating RNA on a 1% agarose gel and EtBr staining. A high-quality RNA sample should show two discrete bands at 3.5 kb and 1.8 kb that represent 26S rRNA and 18S rRNA, respectively. Separate 10  $\mu\text{g}$  of total RNA by electrophoresis on a 15% TBE–urea gel along with 20–100 bases ladder that should be loaded several lanes away from the total RNA sample. Stain the gel with EtBr and view the gel on a Dark Reader transilluminator (Clare Chemical Research) or a long wavelength UV transilluminator. Cut out a band of gel in the sample lane corresponding to the 18–30 nt bands in the marker lane. Elute and precipitate RNA from the gel slice (see Section III, Modified RACE to map cleavage sites). Resuspend the RNA pellet in ultrapure water.

## I. cDNA Library Preparation

In order to prepare the cDNA libraries, RNA adaptors for 5'- and 3'-ends are ligated to the purified small RNAs for use in reverse transcription and PCR amplification. The 3'-adaptor also possesses a sequence that is in the reverse complementary orientation to a surface-bound amplification primer on the Illumina Genome Analyzer flowcell. The sequences of another flowcell primer and the sequencing primer will be added to the other end of the template by a PCR primer. The adaptors must be designed to take advantage of the 5'-phosphate and 3'-hydroxyl termini of small RNAs, resulting from RNase III processing. In contrast, most RNase degradation products have 5'-hydroxyl and 3'-phosphate termini. The 5'-adaptor should carry a hydroxyl group at the 3'-terminus as an acceptor for 5'-phosphate containing small RNAs, while the 3'-adaptor should carry a 5'-phosphate terminus as a donor and a non-nucleotidic moiety at the 3'-terminus to prevent circularization. Use excess amounts of adaptors over 5'-phosphate containing small RNAs to avoid circularization. Alternatively, use a 3'-adaptor with a preadenylated 5'-end and a truncated T4 RNA ligase 2, RNL2(1-249), in the absence of ATP to reduce background ligation (Hafner *et al.*, 2008). Perform reverse transcription over the ligated RNA products and then amplify by PCR. Reagents, adaptors, primers, and detailed protocols are available from Illumina, Inc.

### 1. Adaptor Ligation

Dissolve purified small RNA in 5.7  $\mu\text{L}$  of ultrapure water, and mix with 1.3  $\mu\text{L}$  of 5'-RNA adaptor (Illumina part #1000595), 1  $\mu\text{L}$  of 10 $\times$  T4 RNA ligase buffer, 1  $\mu\text{L}$  of RNaseOUT and 1  $\mu\text{L}$  of T4 RNA ligase. Incubate the reaction mixture at 20°C for 6 h in a thermal cycler and hold at 4°C. Separate ligated RNAs by electrophoresis on a 15% TBE–urea gel along with 20–100 bases ladder and recover RNAs corresponding to 40–60 nt bands using the same elution and ethanol precipitation steps mentioned above. Dissolve purified small RNA in 6.4  $\mu\text{L}$  of ultrapure water, and mix

with 0.6  $\mu\text{L}$  of 3'-RNA adaptor (Illumina part #1000596), 1  $\mu\text{L}$  of 10 $\times$  T4 RNA ligase buffer, 1  $\mu\text{L}$  of RNaseOUT, and 1  $\mu\text{L}$  of T4 RNA ligase. Incubate the reaction mixture at 20°C for 6 h in a thermal cycler and hold overnight at 4°C. Separate ligated RNAs by electrophoresis on a 15% TBE–urea gel along with 20–100 bases ladder and recover RNAs corresponding to 70–90 nt bands using the same elution and ethanol precipitation steps mentioned above. Dissolve the RNA pellet in ultrapure water.

## 2. Adaptor Ligation (Alternative)

Dissolve purified small RNA in 5  $\mu\text{L}$  of ultrapure water and mix with 1  $\mu\text{L}$  of 1 $\times$  preadenylated 3'-sRNA adaptor (Illumina part #1000263). Incubate at 70°C for 2 min then immediately transfer onto ice. Add 1  $\mu\text{L}$  of 10 $\times$  T4 RNL2 truncated reaction buffer (NEB), 0.8  $\mu\text{L}$  of 100 mM  $\text{MgCl}_2$ , 1.5  $\mu\text{L}$  of T4 RNA ligase 2 truncated (NEB), and 0.5  $\mu\text{L}$  of RNaseOUT. Incubate the reaction mixture in a thermal cycler at 22°C for 1 h. Then add 1  $\mu\text{L}$  of 5'-RNA adaptor (Illumina part #1000595), 1  $\mu\text{L}$  of 10 mM ATP, and 1  $\mu\text{L}$  of T4 RNA ligase. Incubate the reaction mixture in a thermal cycler at 20°C for 1 h. Store the product at 4°C.

## 3. Reverse Transcription and PCR Amplification

Mix 4.5  $\mu\text{L}$  of adaptor-ligated RNAs and 0.5  $\mu\text{L}$  of RT primer (Illumina part #100597). Heat the mixture in a thermal cycler at 65°C for 10 min and transfer immediately onto ice. Add 2  $\mu\text{L}$  of 5 $\times$  first strand buffer (Invitrogen), 0.5  $\mu\text{L}$  of 12.5 mM dNTP mix, 1  $\mu\text{L}$  of 100 mM DTT and 0.5  $\mu\text{L}$  of RNaseOUT. Heat the sample to 48°C in a thermal cycler for 3 min. Add 1  $\mu\text{L}$  of SuperScript II Reverse Transcriptase (Invitrogen) and incubate the reaction mixture in a thermal cycler at 44°C for 1 h. The final volume is 10  $\mu\text{L}$ . Premix 40  $\mu\text{L}$  of PCR master mix with 28  $\mu\text{L}$  of ultrapure water, 8  $\mu\text{L}$  of 5 $\times$  Phusion HF buffer (Finnzymes), 0.5  $\mu\text{L}$  of Primer GX1 (Illumina part # 1000591), 0.5  $\mu\text{L}$  of Primer GX2 (Illumina part #1000592), 0.5  $\mu\text{L}$  of 25 mM dNTP mix, and 0.5  $\mu\text{L}$  of Phusion DNA Polymerase (Finnzymes). Mix 10  $\mu\text{L}$  of reverse-transcribed cDNAs and 40  $\mu\text{L}$  of PCR master mix and perform the PCR reaction using the following protocol: (1) 30 s at 98°C, (2) 15 cycles of 10 s at 98°C, 30 s at 60°C, and 15 s of 72°C, (3) 10 min at 72°C, (4) hold at 4°C. Purify the PCR products (a band of approximately 92 bp in length) by separating on a 6% TBE–urea gel followed by elution and ethanol precipitation. Resuspend purified cDNA libraries in 10  $\mu\text{L}$  of resuspension buffer (Illumina part # 1001388). Use 1  $\mu\text{L}$  of the product for TOPO cloning and validate  $\sim$ 100 colonies with traditional sequencing.

## J. Hybridization, Cluster Generation, and Sequencing

The following approaches require operators specifically trained in the use of Illumina Genome Analyzer. An average-size *C. elegans* laboratory may seek assistance or use the service of a core facility. Briefly, denature the cDNA templates by

NaOH and load the sample into an Illumina Genome Analyzer flowcell. Hybridize the complementary end of each template to a flowcell primer. Perform the first extension by *Taq* polymerase to generate a reverse complementary copy that is tethered to the flowcell surface. Remove the not-tethered original template by flushing with NaOH. Hybridize the free end of the complementary copy strand to another flow cell primer and perform extension by *Bst* polymerase, generating a double-stranded product. Denature the double-stranded product by formamide and then the free ends of two single strands can anneal to another set of two flowcell primers, respectively. Repeat cycles of denaturation, annealing, and extension to generate a cluster of  $\sim 1000$  double-stranded products. Cleave one of the flowcell primers and remove one strand selectively, resulting in clusters of single-stranded templates. This allows more efficient hybridization of the sequencing primer and ensures that the sequencing occurs only in one direction. Perform the sequencing by synthesis, 36 cycles of single-base extension, in the Illumina Genome Analyzer according to the manufacturer's instruction, using a modified DNA polymerase and a mixture of four dNTPs that are labeled by four different fluorophores and also 3' blocked. In every cycle, the fluorescence signal corresponding to the identity of incorporated nucleotide is imaged and then the fluorophore and the 3' blocking moiety are cleaved for the next cycle. Export the raw sequence data.

## K. Computational Data Analysis

Analysis of the raw sequencing data presents intense computational challenges and the methods often change based on newly proposed algorithms (Creighton *et al.*, 2009; Shendure and Ji, 2008). The first step is to filter out unusable reads from the raw data. For example, unique sequence reads of fewer than 10 copies may be considered as potential sequencing errors and be discarded. Sequence reads that match *E. coli* genome database are considered as contaminations and should also be removed. In searching the small RNAs with a size around  $\sim 17$ – $26$  nt in length, given that the average length of Illumina read is  $\sim 36$  nt, finding part of the 3'-adaptor in the 3'-end of the read sequence can also be used as a quality control step, while this may not apply to all noncoding RNAs.

For profiling expression of known miRNAs and 21U-RNAs, we trimmed the 3'-adaptor sequence from the reads and used an in-house alignment for perfectly matching known sequences from reference databases [miRBase (Griffiths-Jones *et al.*, 2008) and previously annotated 21U-RNAs (Batista *et al.*, 2008)]. The number of known small RNA reads is normalized to the total number of reads that matched the *C. elegans* genome (see below) and that can represent small RNA abundance. For alignment of large sets of short reads to genome databases, an increasing number of software tools have been developed, which also allow for mismatches and/or gaps (for review see Shendure and Ji, 2008). We loaded the adaptor-trimmed reads to the SOAP (short oligonucleotide alignment package) (Li *et al.*, 2008, 2009) for matching the *C. elegans* genome [WormBase (Harris *et al.*,

2010)], allowing a maximum of 2 bp mismatches in the reads. The resulting sequence reads from 17~26 nt in length that matched the *C. elegans* genome were considered as the total reads in the expression profiling. For examining the proportion of all noncoding RNAs in the deep sequencing database, we aligned genome-matched reads to the noncoding RNA list in WormBase using Blastn, with the following parameters: -e 0.001-G 5-E 2-q-3-r 1-W 7-v 10-b 10.

#### 1. Novel miRNA and 21U-RNA Discovery (Bracht *et al.*, 2010)

For discovering novel miRNAs from the deep sequencing database, several software tools based on different algorithms are publicly available, including miRDeep (Friedlander *et al.*, 2008), CID-miRNA (Tyagi *et al.*, 2008), miRank (Xu *et al.*, 2008), miRCat (Moxon *et al.*, 2008), and miRanalyzer (Hackenberg *et al.*, 2009). We have chosen miRDeep, which uses a probabilistic model scoring both the compatibility of the position and frequency of sequences with the model of miRNA biogenesis as well as the stability of the characteristic hairpin structures of predicted pre-miRNAs. We removed reads that match previously annotated noncoding RNAs and mRNAs using Blastn and also removed reads that perfectly match known miRNAs and 21U-RNAs. Then we loaded the remaining reads to miRDeep for searching for novel miRNAs and used RNAfold to predict secondary structures of putative pre-miRNA hairpins (Hofacker, 2009). For searching novel 21U-RNAs, we aligned the sequences of 21 nt in length with a 5' uracil in the remaining reads with the *C. elegans* genome and judged by the previously reported characteristics of 21U-RNAs, that is, localization on chromosomes and the upstream conserved motif (Ruby *et al.*, 2006).

#### V. Analysis of miRNP Complexes

miRNAs function as part of the RISC protein complex (Bartel, 2009). Multiple proteins also generally or specifically regulate miRNA biogenesis (Kai and Pasquinelli, 2010; Winter *et al.*, 2009). Protein immunoprecipitation followed by MASS spectrometry has been done to identify proteins associated with RISC proteins (Zhang *et al.*, 2007). RNA immunoprecipitation (RIP) experiments have been used to identify target mRNAs that associate with miRISC *in vivo* (Zhang *et al.*, 2007). The use of cross-linking to stabilize endogenous RNA-protein interactions followed by immunoprecipitation of the protein complex of interest and isolation of the RNA regions protected by the complex (cross-linking and immunoprecipitation, CLIP) has enabled the identification of sequences specifically bound by miRISC (Zisoulis *et al.*, 2010; Zisoulis *et al.*, 2011). Additionally, RIP allows the identification of the primary, precursor, or mature miRNA that a particular protein or protein complex binds *in vivo* (Van Wynsberghe *et al.*, 2011). This method requires an antibody against the protein or fusion-protein of interest. Parallel RIP analysis in a mutant worm strain automatically provides a negative control for nonspecific antibody binding. Additionally, an IgG antibody produced in the same species as the antibody of interest should be included as a negative control to determine fold-enrichment. UV-cross-linking of live worms allows endogenous RNA-protein interactions to be stabilized and avoids

potential nonphysiological interactions that can happen upon extract preparation (Mili and Steitz, 2004; Van Wynsberghe *et al.*, 2011).

In addition to CLIP and RIP experiments, which are aimed at the identification of RNA sequences that associate with proteins of interest, the characterization of the protein components of miRNA ribonucleoprotein complexes (miRNPs) contributes to the understanding of mechanistic aspects of miRNA pathways. Gel filtration and electrophoretic mobility shift assay (EMSA) allow for the estimation of the size of miRNPs both in wild-type organisms and in mutants or under designed experimental circumstances. UV-cross-linking of miRNPs assembled with radio-labeled miRNAs in EMSA experiments transfers the radioactive label to miRNA-binding proteins and can be used to visualize them on SDS-PAGE. Ultimately, affinity selection using biotinylated 2'-*O*-methyl oligonucleotides allows purification of miRNPs and further identification of miRNP component proteins by mass spectrometry.

## L. RIP

This protocol is adapted from the Dynabeads instruction protocol (Invitrogen) and published immunoprecipitation protocols (Van Wynsberghe *et al.*, 2011; Zhang *et al.*, 2007; Zisoulis *et al.*, 2010, 2011). Cross-linking of the antibody to the beads and blocking of the beads with salmon-sperm DNA is recommended to decrease background, and enhance reproducibility and specificity.

### 1. Worm Collection and Cross-Linking

At least 200  $\mu\text{L}$  of worms collected at the appropriate stage are required for this protocol. After collection, rock worms in M9 for  $\sim 10$  min to allow residual bacteria to be digested. Resuspend worms in  $\sim 200$   $\mu\text{L}$  of M9 and plate several drops onto 2–3 100-mm worm plates. Spread worms equally around the plate. UV cross-link worms in a Spectrolinker XL-1000 UV Crosslinker with an energy output of  $3 \text{ kJ/m}^2$  at a distance of  $\sim 10$  cm from the light source. Collect cross-linked worms with M9. Flash freeze worm pellet in a dry ice–ethanol bath.

### 2. Lysate Preparation

Crush worms in liquid nitrogen with mortar and pestle. Transfer worms to a 1.5 mL eppendorf tube containing three times the pellet volume of lysis buffer [150 mM NaCl, 25 mM HEPES pH 7.5, 0.2 mM DTT, 10% glycerol, 0.025 U/ $\mu\text{L}$  RNasin, 1% Triton X-100, and protease inhibitors (Roche)]. Spin sample at top speed at  $4^\circ\text{C}$  for 15 min. Transfer supernatant to a new eppendorf tube and determine the protein concentration by the Bradford protein assay (Bio-Rad).

### 3. Preclear Lysate

At least 500  $\mu\text{g}$  of lysate protein is required per sample. Equal amounts of sample should be used for each reaction (typically: antibody of interest and IgG control).

Add lysate and lysis buffer in a total of 1 mL to 50  $\mu\text{L}$  of Dynabeads (Invitrogen) that have been washed twice in 100  $\mu\text{L}$  of WB buffer (0.1 M  $\text{NaPO}_4$ , 0.1% Tween 20, pH 8.2). Rotate at 4°C for 1 h.

#### 4. Immunoprecipitation

Wash 50  $\mu\text{L}$  of Dynabeads per sample with 100  $\mu\text{L}$  WB buffer. Block beads by resuspending in 180  $\mu\text{L}$  of WB buffer and 20  $\mu\text{L}$  of 10  $\mu\text{g}/\mu\text{L}$  sheared, salmon-sperm DNA (ssDNA). Incubate on rotator at RT for 30 min. Add 5  $\mu\text{g}$  of antibody and incubate on rotator at RT for 10 min. Wash antibody–bead complex twice with 100  $\mu\text{L}$  of conjugation buffer (20 mM  $\text{NaPO}_4$ , 150 mM NaCl). Incubate beads in 250  $\mu\text{L}$  bis(sulfosuccinimidyl)suberate ( $\text{BS}^3$ ) conjugation buffer (5 mM  $\text{BS}^3$  in conjugation buffer containing 0.5  $\mu\text{g}/\mu\text{L}$  ssDNA) at RT for 30 min with rotation. Quench cross-linking reaction by adding 12.5  $\mu\text{L}$  of 1 M Tris-Cl pH 7.4 and incubating at RT for 15 min with rotation. Wash beads twice with 100  $\mu\text{L}$  of lysis buffer. Add antibody-cross-linked beads to the precleared lysate from the previous step, and rotate at 4°C overnight.

Save the supernatant from the immunoprecipitation reaction for IP efficiency analysis by western blotting. Wash beads twice for 1 min in a Thermomixer R (Eppendorf) set to 4°C with: (1) wash buffer [ $1\times$  PBS pH 7.4, 0.1% SDS, 0.5% sodium deoxycholate, and 0.5% NP-40], (2) high salt wash buffer [ $5\times$  PBS pH 7.4, 0.1% SDS, 0.5% sodium deoxycholate, and 0.5% NP-40], and (3) PK buffer [100 mM Tris-Cl pH 7.4, 50 mM NaCl, and 10 mM EDTA]. Resuspend beads in 100  $\mu\text{L}$  of PK buffer. Remove a 15  $\mu\text{L}$  aliquot of bead solution for IP efficiency analysis by western blotting. Remove excess PK buffer and resuspend beads in 100  $\mu\text{L}$  of proteinase K solution (5 mg of proteinase K in 1 mL of PK buffer) at 37°C for 20 min in a Thermomixer R (Eppendorf) set to 1200 rpm. Quench reaction by adding 100  $\mu\text{L}$  of PK buffer containing 7 M urea, and incubating the reaction at 37°C for 20 min at 1000rpm.

#### 5. RNA Extraction

Extract total and immunoprecipitated RNA by adding 800  $\mu\text{L}$  of Trizol (Invitrogen) directly to an aliquot of input protein lysate or the proteinase K bead solution and extracting as described above in Section I. Resuspend extracted RNA in 120  $\mu\text{L}$  of DEPC  $\text{H}_2\text{O}$ . Treat with DNase by adding 15  $\mu\text{L}$  of  $10\times$  RQ1 DNase Buffer and 15  $\mu\text{L}$  of RQ1 DNase to each sample and incubating at 37°C for 1 h. Extract RNA a second time as described above in the second RNA-extraction step of Section I. Resuspend RNA samples in 10  $\mu\text{L}$  of  $\text{ddH}_2\text{O}$ .

#### 6. RT-PCR

Add 9  $\mu\text{L}$  of immunoprecipitated RNA or 2.5  $\mu\text{g}$  of total RNA in a 9  $\mu\text{L}$  volume to 1  $\mu\text{L}$  of random primers (250 ng/ $\mu\text{L}$ ), 1  $\mu\text{L}$  of 10 mM dNTPs, and 1.2  $\mu\text{L}$  of  $\text{ddH}_2\text{O}$ . Incubate at 65°C for 5 min and 4°C for 1 min. Add 4  $\mu\text{L}$  of  $5\times$  buffer (Invitrogen),

2  $\mu\text{L}$  of 0.1 M DTT, and 1  $\mu\text{L}$  of RNasin (Promega). Incubate at 42°C for 2 min before adding 1  $\mu\text{L}$  of Superscript III Reverse Transcriptase enzyme (Invitrogen). Incubate at 42°C for 50 min and 70°C for 15 min. Add 1  $\mu\text{L}$  of RNaseH (Invitrogen), and incubate at 37°C for 30 min. Dilute an aliquot of the cDNA reaction 1:5 for use in qPCR. Analyze samples for the primary or processed miRNA or mRNA of interest by PCR or qPCR as described above in Section II.

## M. Gel Filtration Analysis of miRNPs

Gel filtration has been used to fractionate cell extracts from *C. elegans*, *Drosophila*, and human cells to determine the sizes of complexes corresponding to RNAi activity, siRNAs, miRNAs, or miRNP/RISC (Caudy *et al.*, 2002, 2003; Chan *et al.*, 2008; Gregory *et al.*, 2005; Pham *et al.*, 2004). Gel filtration can also be used to detect cofractionation of proteins of interest with miRNAs or to purify fractions that possess miRNA-related activities, that is miRNA binding or pre-miRNA processing.

### 1. Cell Extract Preparation

Harvest mixed-staged (or staged) worms grown in S medium supplied with *E. coli* HB101. Clean the worms by sucrose floatation and then wash three times with 0.1 M NaCl. Store the worms at  $-80^{\circ}\text{C}$ . Thaw 1 mL of packed worms on ice and wash twice with wash buffer (50 mM Tris-HCl pH 7.5 and 10 mM potassium acetate). Resuspend the worms in 4 mL of homogenization buffer [50 mM Tris-HCl pH 7.5, 10 mM potassium acetate, 5 mM DTT, 10 U/mL SuperaseIn RNase inhibitor (Ambion), 10% glycerol, and 1 $\times$  Roche Complete protease inhibitor, EDTA-free] and incubate on ice for 20 min with intermittent agitation. Transfer the worms to a 7 mL glass dounce homogenizer tube (Kontes Glass Co., Vineland, N.J., article no. 885303-0007) and homogenize by 30 strokes on ice using the tight-fitting pestle B (Kontes Glass Co., Vineland, N.J., article no. 885302-0007). Transfer homogenate to a 15-mL tube and add magnesium acetate to 2 mM and adjust the concentration of potassium acetate to 100 mM. Incubate the homogenate on ice for 20 min with intermittent agitation and then aliquot the homogenate to 1.5 mL eppendorf tubes. Centrifuge the tubes at 16,100  $\times$  g at 4°C for 30 min. Store the supernatant at  $-80^{\circ}\text{C}$ .

### 2. Gel Filtration

Equilibrate a Superdex-200 HR 10/30 column with cold homogenization buffer. Load 400  $\mu\text{L}$  of cell extract into the column. Run the chromatography at 4°C at the rate of 0.35 mL per minute using AKTA FPLC (Pharmacia) and collect 60 fractions (0.5 L each). Use the following size markers: thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), and albumin (67 kDa). For monitoring miRNAs in the fractions, mix 250  $\mu\text{L}$  of fraction sample with 750  $\mu\text{L}$

of Trizol LS and follow the RNA-extraction method described in Section I. Refer to Section II for a description of miRNAs detection by northern blotting. Use 10–20  $\mu\text{L}$  of fraction sample for western analysis of known RISC components with appropriate antibodies.

## N. EMSA of miRNPs

EMSA can be used to qualitatively identify large ribonucleoprotein complexes, for example, holo-RISC (Pham *et al.*, 2004), or to detect miRNA-binding proteins that may only form small complexes with their substrates (Chan *et al.*, 2008). Below, we show a basic EMSA setup using radiolabeled miRNAs and crude cell extracts. However, the design of EMSA should depend on the purpose of the experiment and the parameters, such as the gel, buffer, running temperature, and/or the presence of competitor or heparin, have to be optimized.

### 1. Cell Extract Preparation

See above.

### 2. Radiolabeled Synthetic miRNA Preparation

Obtain synthetic miRNAs from Dharmacon Research, Inc. or Integrated DNA Technologies, Inc. Mix 1  $\mu\text{L}$  of 50  $\mu\text{M}$  synthetic miRNA (50 pmol), 5  $\mu\text{L}$  of gamma- $^{32}\text{P}$  ATP (6000 Ci/mmol, NEN), 2.5  $\mu\text{L}$  of  $10\times$  T4 PNK buffer, 2.5  $\mu\text{L}$  of T4 polynucleotide kinase, and 14  $\mu\text{L}$  of RNase-free water. Incubate the reaction mixture at  $37^\circ\text{C}$  for 30 min. Remove free gamma- $^{32}\text{P}$  ATP by passing the mixture through a Sephadex G-25 column (Roche) according to the manufacturer's instruction.

### 3. Native Gel Preparation

Prepare a 16 cm  $\times$  16 cm gel of 0.8 mm thickness with  $1\times$  TBE and 5% acrylamide. The ratio of acrylamide/bis-acrylamide depends on the experiment. Use 40:1 for large complexes and 19:1 for small complexes.

### 4. miRNA Binding Reaction and Gel Running

Prepare miRNA binding reaction mixtures (10  $\mu\text{L}$ ) that contain 20 mM Tris pH 7.5, 1 mM magnesium acetate, 1 mM calcium chloride, 0.01% Nonidet P-40, 2 mM ATP, 250 ng/ $\mu\text{L}$  *E. coli* tRNA,  $5\times 10^5$  c.p.m radiolabeled miRNA and cell extract that can be used in different amounts. Incubate the reaction at RT for 45 min. Mix with 2  $\mu\text{L}$  of  $6\times$  native gel loading dye (30% glycerol, 0.25% BPB, and 0.25% xylene cyanol). Run the gel at 100 V at  $4^\circ\text{C}$  until the BPB dye migrates to 3 cm from

the bottom of the gel. Lay the gel onto Whatman 3MM paper and dry it using a gel dryer. Detect the signals by autoradiography or phosphoimaging.

#### 5. UV-Cross-Linking of miRNPs

Prepare 150  $\mu\text{L}$  of miRNA binding reaction mixture and separate the sample on a nondenaturing gel (see above). Do not dry the gel. Instead, expose the gel wrapped in a transparent plastic film to a BioMax MS film (Kodak) for 1 h. After autoradiography, excise gel bands that contain radioactive signals corresponding to miRNPs of interest. Place the gel slices on a clean glass plate that is sitting on ice and expose them to 254 nm UV light in Stratalink UV Crosslinker (Stratagene) with an energy output of 2  $\text{J}/\text{cm}^2$  at a distance of approximately 10 cm from the light source. Mince the gel slice to small pieces and incubate with 150  $\mu\text{L}$  of RNase A (250 U/mL in 50 mM Tris-HCl pH 7.5 buffer) at RT for 30 min. Add 50  $\mu\text{L}$  of XT sample buffer 4X (Bio-Rad) and agitate vigorously on a vortex mixer for 3 h. Analyze the eluted proteins by denaturing gel electrophoresis using a Criterion XT Bis-Tris 4–12% gradient gel (Bio-Rad) and autoradiography. To detect all UV-cross-linking signals not specific to any individual complex, irradiate 20  $\mu\text{L}$  of miRNA binding reaction mixture placed in a 96-well plate on ice following RNase treatment and denaturing gel electrophoresis as mentioned above.

### O. Affinity Purification of miRNPs by Biotinylated 2'-O-methyl Oligonucleotides

Antisense 2'-O-methyl oligonucleotides block siRNA and miRNA function *in vitro* and *in vivo* (Hutvagner *et al.*, 2004). In addition, immobilized 2'-O-methyl oligonucleotides complementary to siRNAs or miRNAs capture Argonaute proteins in *C. elegans* extract (Hutvagner *et al.*, 2004; Steiner *et al.*, 2007; Yigit *et al.*, 2006). Below, we describe an adapted procedure to purify miRNPs from *C. elegans* crude extract using biotinylated 2'-O-methyl oligonucleotides. The RNA or protein contents in purified miRNPs are subjected to further investigation.

#### 1. Cell Extract Preparation

Depending on the particular miRNA of interest, harvest synchronized worms of the appropriate developmental stage. For example, to capture *let-7* binding ribonucleoprotein complexes, prepare cell extract from L4 stage worms. For one volume of packed worms, use five volumes of homogenization buffer [10 mM Tris-HCl pH 7.5, 100 mM KCl, 2 mM  $\text{MgCl}_2$ , 0.5% (v/v) Nonidet P-40, 15% (v/v) glycerol, 5 mM DTT, 50 U/mL SuperaseIn RNase inhibitor (Ambion) and 1 $\times$  Roche Complete protease inhibitor, EDTA-free].

#### 2. Biotinylated 2'-O-Methyl Oligonucleotides

Obtain 5'-biotinylated 2'-O-methyl oligonucleotides from Integrated DNA Technologies, Inc. In previous studies, five nucleotides complementary to the

sequences flanking each side of the siRNA/miRNA target site were added to enhance the efficiency of blocking siRNA/miRNA activity (Hutvagner *et al.*, 2004). For affinity purification of miRNPs, we found that the flanking sequences are not necessary (our unpublished observation). For example, we used a 5'-biotinylated 22-nt 2'-*O*-methyl oligonucleotide complementary to *let-7* and a 22-nt oligonucleotide complementary to luciferase mRNA sequence as negative control.

### 3. Affinity Purification

Incubate 300  $\mu$ L of cell extract with 15 pmol of 5'-biotinylated 2'-*O*-methyl oligonucleotide at RT for 1 h. Subsequently, add 20  $\mu$ L of Streptavidin Sepharose (GE Healthcare), pre-equilibrated with the homogenization buffer, and incubate at RT for 1 h with gentle agitation on a nutator. Centrifuge at 8000 rpm for 30 s and remove the unbound fraction. Wash five times by resuspending with 1 mL of homogenization buffer and centrifuging at 8000 rpm for 30 s each time. For analysis of purified proteins, resuspend the beads in a final volume of about 28  $\mu$ L of water and add 10  $\mu$ L XT sample buffer (4 $\times$ ) and 2  $\mu$ L XT reducing agent (20 $\times$ ) (Bio-Rad) or 2-mercaptoethanol. Heat at 95°C for 5 min. Remove the beads by passing the sample through a Spin X cellulose acetate filter (Corning Costar). Analyze the eluted proteins by denaturing gel electrophoresis using a Criterion XT Bis-Tris 4–12% gradient gel (Bio-Rad) and western blotting with antibodies against proteins of interest. RNAs in the beads or unbound fraction can be extracted by Trizol and analyzed by northern blotting.

## VI. Construction of Transgenic Strains to Analyze miRNA Expression and Target Regulation

Northern blotting, real-time quantitative PCR, and deep sequencing are powerful tools for analysis of miRNAs in *C. elegans*, but they lack the ability to detect spatiotemporal patterns of expression. Reporter genes, such as *gfp* (encoding green fluorescent protein) (Chalfie *et al.*, 1994) or *lacZ* (encoding beta-galactosidase) (Fire *et al.*, 1990), driven by a miRNA promoter (*Pmir*) are more amenable to analyze spatiotemporal miRNA expression *in vivo*. Moreover, a series of deletions in the promoter region of the construct can be used to determine the transcriptional regulatory *cis*-elements of miRNAs. For example, our group has examined the expression pattern of *lin-4* and *let-7* family members using *Pmir::gfp* constructs (Esquela-Kerscher *et al.*, 2005; Johnson *et al.*, 2005). In addition, by serially deleting the *let-7* promoter region upstream of the *gfp* reporter, we have succeeded in pinpointing a short temporal regulatory *cis*-element (TRE) of *let-7* (Johnson *et al.*, 2003). We also have demonstrated that *hbl-1*, one target of *let-7*, inhibits the transcription of *let-7* using a *Plet-7::gfp* construct in which the putative HBL-1 binding sites were deleted (Roush and Slack, 2009). Recently, Martinez *et al.* (2008) have used *Pmir::gfp* reporter constructs to analyze spatiotemporal promoter activity of miRNA on a genomic scale (Martinez *et al.*, 2008). With all its promises, one obvious limit of the miRNA promoter-reporter fusion approach is that it can only

reflect the transcriptional control of primary miRNAs production, and it cannot provide information on post-transcriptional regulation or turnover.

Similarly, reporter gene fusion can also be used to analyze miRNA-mediated gene regulation at the level of 3'UTRs (see below), which contain regulatory elements such as miRNA-binding sites.

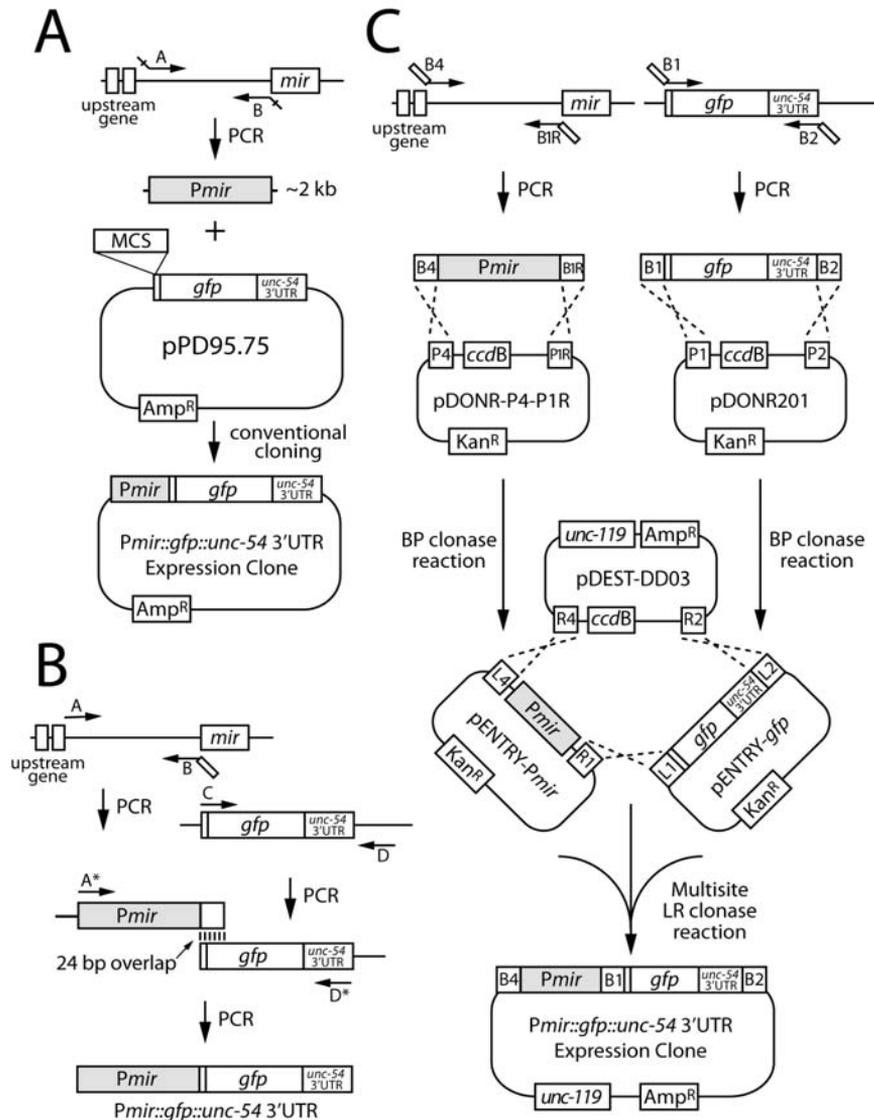
## P. miRNA Promoter–Reporter Fusion

The first step in designing a miRNA promoter–reporter fusion construct is to decide on the putative promoter sequence. Several factors should be taken into consideration, including the intergenic, intragenic, or operonic location of miRNA loci and the distance to the next upstream transcript, as well as the phylogenetic conservation of promoter sequences and the presence of known transcription factor-binding sites. Determination of the 5' end of the primary miRNA by 5' RACE (see Section III) would help design the promoter–reporter fusion construct and it is debatable whether the primary miRNA sequence upstream of the mature miRNA should be kept in the construct. Previous studies have shown that fragments between a few hundred basepairs to 2 kb in length upstream of the mature miRNA or known pre-miRNA hairpin [miRBase (Griffiths-Jones *et al.*, 2008)] contain sufficient information to faithfully recapitulate the pattern of expression of endogenous miRNAs. For example, we have used a 0.5 kb fragment upstream of *lin-4* and a 2.2 kb fragment upstream of *mir-84* as the promoters (Esquela-Kerscher *et al.*, 2005; Johnson *et al.*, 2005). Martinez *et al.* (2008) have used the intergenic genomic sequences upstream of miRNA genes from ~300 bp to 2 kb as the promoters in their genome-scale analysis (Martinez *et al.*, 2008). Putative promoter regions located in vast intergenic regions that lack homology across species or known transcription factor binding sites may thus be safely limited to approximately 2 kb in most cases.

The second step is to choose the reporter gene and the vector backbone. Andrew Fire's laboratory has developed a large array of vectors and reporter gene derivatives, using *gfp* or *lacZ* as the reporter, on the backbone of a pUC19 plasmid for fusion gene expression in *C. elegans* (see Links). The cloning strategies described below are adapted from several common cloning techniques for reporter gene fusions in *C. elegans* (Boulin *et al.*, 2006). Due to the relatively short length of miRNA promoters we have chosen, we only discuss three strategies more suitable for short DNA fragment cloning, including conventional cloning, PCR fusion (Horton *et al.*, 1989), and Multisite Gateway cloning (Hope *et al.*, 2004).

### 1. Conventional Cloning

The most common strategy to construct a miRNA promoter–reporter fusion is conventional restriction enzyme-based cloning, graphically summarized in Fig. 2A. Perform PCR to amplify the promoter region (~300 bp to 2 kb upstream in length



**Fig. 2** Generating a *Pmir::gfp::unc-54* 3' UTR reporter construct. (A) By conventional cloning. Use primer A and B to amplify a 2 kb fragment upstream of mature miRNA or pre-miRNA hairpin and insert the fragment into MCS of pPD95.75. (B) By PCR fusion. Use Primer A and B to amplify the promoter region and primer C and D to amplify *gfp::unc-54* 3' UTR from pPD95.75. Primer B adds a 24-bp region from MCS upstream of *gfp* gene. Anneal these two amplicons and use primer A\* and D\* to amplify the fusion product. (C) By Multisite Gateway cloning. Amplify the promoter region and the reporter gene with primers that introduce appropriate *attB* recombination sites and recombine PCR products with donor plasmids pDONR-P4-P1R and pDONR201, respectively. Recombine the products pENTRY-*Pmir* and pENTRY-*gfp* with pDEST-DD03. See text for details.

from the start of the mature miRNA) from genomic DNA and introduce appropriate restriction enzyme sites at both ends (restriction enzyme sites are introduced via the primers generated for amplification). Generate compatible ends on the purified PCR product and a reporter gene vector by restriction enzyme digestion. Ligate the promoter fragment and the linearized vector and amplify the ligated product in *E. coli*. For promoter-*gfp* fusion, two vectors, pPD95.70 and pPD95.75 from the Fire Vector Kit, have been broadly used in the *C. elegans* community. Both of them carry a S65C variant of the *gfp* gene and pPD95.70 has a SV40 nuclear localization sequence (NLS) for nuclear localization of GFP. The NLS may help with observations of weak GFP signals but should not be used to represent the subcellular localization of the miRNA. The *gfp* gene is linked to a 3'UTR derived from the muscle myosin heavy chain gene *unc-54*, permissive for expression in all cell types. Although conventional cloning is relatively time-consuming and depends on restriction enzyme sites, it has the irreplaceable advantage of generation of a reusable reporter gene construct. This construct can be verified by sequencing, repeatedly amplified, stored for a long time, and easily subjected to further subcloning.

## 2. PCR Fusion

To prepare the promoter-reporter fusion construct more quickly than conventional cloning and escape the limitation of restriction enzyme sites, a two-step approach of PCR can be used to fuse a miRNA promoter and a reporter gene, generating a linear-formed DNA ready for transformation. Below, we describe the procedures for fusing a miRNA promoter to the *gfp* reporter with an *unc-54* 3'UTR, derived from pPD95.75, by PCR fusion (Fig. 2B).

Prepare the following primers:

Primer A: promoter-specific outside forward primer

Primer A\*: promoter-specific nested forward primer

Primer B: promoter-specific reverse primer with a 24-bp sequence complementary to the MCS of the expression vector. (5'-AGTCGACCTGCAGGCATGCAAGCT-promoter specific sequence-3')

Primer C: *gfp* forward primer (5'-AGCTTGCATGCCTGCAGGTCG-3')

Primer D: *gfp* reverse primer (5'-AAGGGCCCGTACGGCCGACTA-3')

Primer D\*: *gfp* nested reverse primer (5'-GGAAACAGTTATGTTTTGGTATA-3')

Perform PCR using primers A and B to amplify a DNA fragment from genomic DNA that contains the promoter region and a 24-bp fragment overlapping with the MCS of pPD95.75. Perform PCR using primer C and D to amplify the *gfp::unc-54* fragment from pPD95.75, which includes the MCS and an artificial intron between MCS and the *gfp* gene. Inclusion of the artificial intron will enhance reporter gene expression. Purify the two PCR products by electrophoresis on an agarose gel and recover the DNA from the corresponding bands. Mix 1–50 ng of each purified DNA fragment and use nested primers A\* and D\* for PCR fusion. Purify the fused

fragment by agarose gel electrophoresis and the recovered DNA, if sufficient, can be used directly for microinjection (Mello and Fire, 1995; Mello *et al.*, 1991).

The primary advantages of PCR fusion are speed, ability to carry out multiple constructions in parallel, and the lack of a requirement for restriction enzyme sites. However, the construct cannot be amplified in bacteria and additive PCR rounds may increase point mutations. These shortcomings can be circumvented by cloning the PCR fusion product, verifying its sequence, and propagating it in bacteria.

### 3. Multisite Gateway Cloning

Invitrogen's Gateway cloning system has provided a versatile way to shuttle cloned DNA fragments among a variety of vector backbones for different applications (Walhout *et al.*, 2000). This technology utilizes the site-specific recombination properties of bacteriophage lambda. Compatible recombination sites (*attB* and *attP* sites; B for bacteria and P for phage) on the DNA fragment of interest and the cloning vector are recognized by bacteriophage lambda integrase proteins, and recombination between *attB* and *attP* sites is carried out. The result is an entry clone carrying the DNA fragment of interest flanked with *attL* (L for left) and *attR* (R for right) sites, which consist of DNA sequences from *attB* and *attP*. The reverse reaction between compatible *attL* and *attR* of the entry clone and the destination vector moves the cloned DNA from the entry clone to the destination vector. The result is an expression clone carrying the DNA fragment of interest flanked with original *attB* and *attP* sites. Different core sequences in the *att* sites provide specificity when recombination occurs, that is, *attB1* will only react with *attP1* and *attB2* will only react with *attP2*. This specificity ensures the correct orientation of inserted DNA fragment. Moreover, by using multiple designed *att* sites, one can perform recombination between up to four DNA fragments. Below, we describe a two-fragment recombination application of multisite Gateway cloning (Hope *et al.*, 2004) (Fig. 2C) that has been used to build a *C. elegans* Promoterome library (Dupuy *et al.*, 2004) and also has been used to systematically generate miRNA promoter-*gfp* reporter constructs (Martinez *et al.*, 2008).

Prepare the following primers: [adapted from (Hope *et al.*, 2004)]

For miRNA promoter:

5' primer (*attB4*): 5'-GGGGACAACCTTTGTATAGAAAAGTTGTG-promoter specific sequence (18–25 bp)

3'-primer (*attB1.1R*): 5'-GGGGACAACCTTTTTGTACAAAGTTGC-promoter specific sequence (18–25 bp)

For *gfp* gene:

5' primer (*attB1.1*): 5'-GGGGACAACCTTTGTACAAAAAAGTTGTG-*gfp* gene-specific sequence (18–25 bp)

3' primer (*attB2.1*): 5'-GGGGACAACCTTTGTACAAGAAAGTTGC-*gfp* gene-specific sequence (18–25 bp)

As shown in Fig. 2C, amplify *attB4*-miRNA promoter (*Pmir*)-*attB1.1R* and *attB1.1-gfp::unc-54* 3'UTR-*attB2.1* fragments by PCR from source DNA (genomic DNA and pPD95.75). Purify the PCR products by agarose gel electrophoresis. Recombine *attB4-Pmir-attB1.1R* with pDONR-P4-P1R (Invitrogen) and *attB1.1-gfp::unc-54* 3'UTR-*attB2.1* with pDONR201 or pDONR221 (Invitrogen). This will generate two entry vectors, pENTRY-*Pmir* and pENTRY-*gfp*. Recombine pENTRY-*Pmir* and pENTRY-*gfp* with pDEST R4-R2 or pDEST-DD03 (Dupuy *et al.*, 2004) (available from Marc Vidal, Harvard University). pDEST-DD03 contains *attR4* and *attR2* sites along with an *unc-119* gene which is used in ballistic transformation protocols (Praitis *et al.*, 2001). All recombination reactions are according to manufacturer's instruction.

### Q. Reporter-miRNA Target 3'UTR Fusion

In addition to determining the expression pattern of miRNAs during development, reporter fusion constructs can be used to study the miRNA-mediated regulation of target genes. Usually, target genes of miRNAs are identified by genetic screening or predicted by computational analysis, which is mostly based on the assumption that miRNAs associate with regulatory elements in the 3'UTR of target genes. An experimental approach to verify a newly found or predicted miRNA/target pair is to examine the expression of a reporter that is fused with the 3'UTR of a target gene. For example, our group has used a *lacZ* reporter construct, *Pcol-10::lacZ::lin-41* 3'UTR, to observe the regulation of heterochronic gene *lin-41* by *let-7* (Reinhart *et al.*, 2000). The actual *let-7* binding sites on *lin-41* 3'UTR were also verified by a series of mutagenesis experiments on such a reporter (Vella *et al.*, 2004). Moreover, this *Pcol-10::lacZ::lin-41* 3'UTR reporter construct can be used as a tool to evaluate the efficacy of *let-7* in wild-type or *let-7* mutant animals (Chan and Slack, 2009; Reinhart *et al.*, 2000). In these studies, we use the promoter of collagen gene *col-10* to specifically express *lacZ* in hypodermal seam cells where *let-7* is expressed.

A reporter-3'UTR fusion can be constructed by cloning strategies similar to that for building a promoter-reporter fusion (see above and Fig. 2). A fixed and cell-specific promoter can be chosen to drive the reporter expression for the purpose of observation. Usually, a promoter of a housekeeping gene or a gene that is not linked to miRNA regulation will be chosen, like the *col-10* promoter. In other cases, researchers will choose the promoter from the gene of interest along with its 3'UTR and examine the reporter expression in the specific spatiotemporal pattern of the endogenous gene. The 3'UTR region of most genes is annotated in WormBase (Harris *et al.*, 2010) or can be inferred by putative polyadenylation signal (PAS) sites (Hajarnavis *et al.*, 2004) or by RNA-seq data available as part of the track for ALG-1 binding at the UCSC Genome Browser (Hillier *et al.*, 2009; Zisoulis *et al.*, 2010). Putative miRNA binding sites can be predicted by algorithms derived from PicTar (Lall *et al.*, 2006), TargetScan (Lewis *et al.*, 2005), or miRanda (Enright *et al.*, 2003). The presence of ALG-1 binding sites can also be used to prioritize 3'UTR sequences to test for miRNA regulation by reporter analyses (Zisoulis *et al.*, 2010).

## R. Transformation

Transgenic *C. elegans* strains can be obtained by either microinjection (Mello and Fire, 1995; Mello *et al.*, 1991) or ballistic transformation (Wilm *et al.*, 1999). For microinjection, a coinjection marker, like *rol-6(su-1006)* (Kramer *et al.*, 1990), can be used to select the transformed animals. For ballistic transformation, a copy of the wild-type *unc-119* gene is included in the DNA construct as a selection marker when the *unc-119(ed3)* mutant strain is used for transformation (Praitis *et al.*, 2001). The recent development of the Mos1 mediated insertion technique for integrating single copy transgenes into a specific chromosomal location can also be used for generating 3'UTR reporter strains (Frokjaer-Jensen *et al.*, 2008).

### Links

For EST, mRNA and miRNA registries

WormBase: <http://www.wormbase.org>

miRBase: <http://microrna.sanger.ac.uk>

For vectors

Fire Lab *C. elegans* Vector Kit: [http://www.addgene.org/Andrew\\_Fire](http://www.addgene.org/Andrew_Fire) or [ftp://ftp.wormbase.org/pub/wormbase/datasets/fire\\_vectors](ftp://ftp.wormbase.org/pub/wormbase/datasets/fire_vectors)

For miRNA target prediction algorithms and databases

PicTar: <http://pictar.mdc-berlin.de/>

TargetScan: [http://www.targetscan.org/worm\\_12/](http://www.targetscan.org/worm_12/)

miRanda: [http://cbio.mskcc.org/research/sander/data/miRNA2003/miranda\\_new.html](http://cbio.mskcc.org/research/sander/data/miRNA2003/miranda_new.html) and <http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/>

For mRNA 5' and 3' end annotations based on RNA-seq and ALG-1 binding sites based on CLIP-seq

<http://genome.ucsc.edu>

---

---

---

## III. Summary

mRNAs are essential regulators of many basic cellular processes. As such, slight deviations in the amount, timing, or location of miRNA expression can have large effects on cell and organismal growth. The analysis of primary, precursor, and mature miRNA levels as well as the identification and characterization of miRNA targets is thus crucial for determining the step in miRNA biogenesis or function affected in a particular mutant or disease. The methods described in this chapter provide a foundation for analyzing these steps in the powerful model organism *C. elegans*.

## Acknowledgments

The authors thank members of the Pasquinelli and Slack labs for their comments, critical reading of this manuscript, and sharing of protocols. This work was supported by funding to A. E. Pasquinelli from NIH (GM071654), and the Keck, Searle, V. Emerald, and Peter Gruber Foundations, and funding to F. J. Slack from NIH (GM064701, AG033921) and the Ellison Medical Foundation. P. M. Van Wynsberghe was supported by a Ruth L. Kirschstein National Research Service Award Number F32GM087004 from the National Institute of General Medical Sciences.

## References

- Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell* **136**, 215–233.
- Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., Chaves, D. A., Gu, W., Vasale, J. J., Duan, S., Conte Jr., D., Luo, S., Schroth, G. P., Carrington, J. C., Bartel, D. P., and Mello, C. C. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* **31**, 67–78.
- Benes, V., and Castoldi, M. (2010). Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods* **50**, 244–249.
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheatham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Irving, L. J., Karbelashvili, M. S., Kirk, S. M., Li, H., Liu, X., Maisinger, K. S., Murray, L. J., Obradovic, B., Ost, T., Parkinson, M. L., Pratt, M. R., Rasolonjatovo, I. M., Reed, M. T., Rigatti, R., Rodighiero, C., Ross, M. T., Sabot, A., Sankar, S. V., Scally, A., Schroth, G. P., Smith, M. E., Smith, V. P., Spiridou, A., Torrance, P. E., Tzonev, S. S., Vermaas, E. H., Walter, K., Wu, X., Zhang, L., Alam, M. D., Anastasi, C., Aniebo, I. C., Bailey, D. M., Bancarz, I. R., Banerjee, S., Barbour, S. G., Baybayan, P. A., Benoit, V. A., Benson, K. F., Bevis, C., Black, P. J., Boodhun, A., Brennan, J. S., Bridgman, J. A., Brown, R. C., Brown, A. A., Buermann, D. H., Bundu, A. A., Burrows, J. C., Carter, N. P., Castillo, N., Chiara, E. C. M., Chang, S., Neil Cooley, R., Crane, N. R., Dada, O. O., Diakoumakos, K. D., Dominguez-Fernandez, B., Earnshaw, D. J., Egbujor, U. C., Elmore, D. W., Etchin, S. S., Ewan, M. R., Fedurco, M., Fraser, L. J., Fuentes Fajardo, K. V., Scott Furey, W., George, D., Gietzen, K. J., Goddard, C. P., Golda, G. S., Granieri, P. A., Green, D. E., Gustafson, D. L., Hansen, N. F., Harnish, K., Haudenschild, C. D., Heyer, N. I., Hims, M. M., Ho, J. T., and Horgan, A. M. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59.
- Boulin, T., Etchberger, J. F., and Hobert, O. (2006). Reporter gene fusions. WormBook: The Online Review of *C. elegans* Biology, The Online *C. elegans* Research Community, pp. 1–23.
- Bracht, J., Hunter, S., Eachus, R., Weeks, P., and Pasquinelli, A. E. (2004). Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA* **10**, 1586–1594.
- Bracht, J. R., Van Wynsberghe, P. M., Mondol, V., and Pasquinelli, A. E. (2010). Regulation of lin-4 miRNA expression, organismal growth and development by a conserved RNA binding protein in *C. elegans*. *Dev. Biol.* **348**, 210–221.
- Caudy, A. A., Ketting, R. F., Hammond, S. M., Denli, A. M., Bathoorn, A. M., Tops, B. B., Silva, J. M., Myers, M. M., Hannon, G. J., and Plasterk, R. H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. *Nature* **425**, 411–414.
- Caudy, A. A., Myers, M., Hannon, G. J., and Hammond, S. M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev.* **16**, 2491–2496.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Chan, S. P., Ramaswamy, G., Choi, E. Y., and Slack, F. J. (2008). Identification of specific let-7 microRNA binding complexes in *Caenorhabditis elegans*. *RNA* **14**, 2104–2114.
- Chan, S. P., and Slack, F. J. (2009). Ribosomal protein RPS-14 modulates let-7 microRNA function in *Caenorhabditis elegans*. *Dev. Biol.* **334**, 152–160.
- Chekulaeva, M., and Filipowicz, W. (2009). Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.* **21**, 452–460.
- Creighton, C. J., Reid, J. G., and Gunaratne, P. H. (2009). Expression profiling of microRNAs by deep sequencing. *Brief Bioinform.* **10**, 490–497.
- Davis, B. N., and Hata, A. (2009). Regulation of MicroRNA biogenesis: A miRiad of mechanisms. *Cell Commun. Signal* **7**, 18.
- Dupuy, D., Li, Q. R., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stamm, L., Hope, I. A., Hill, D. E., Walhout, A. J., and Vidal, M. (2004). A first version of the *Caenorhabditis elegans* promoterome. *Genome Res.* **14**, 2169–2175.

- Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in *Drosophila*. *Genome Biol.* **5**, R1.
- Esquela-Kerscher, A., Johnson, S. M., Bai, L., Saito, K., Partridge, J., Reinert, K. L., and Slack, F. J. (2005). Post-embryonic expression of *C. elegans* microRNAs belonging to the lin-4 and let-7 families in the hypodermis and the reproductive system. *Dev. Dyn.* **234**, 868–877.
- Fineberg, S. K., Kosik, K. S., and Davidson, B. L. (2009). MicroRNAs potentiate neural development. *Neuron* **64**, 303–309.
- Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**, 189–198.
- Friedlander, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., and Rajewsky, N. (2008). Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **26**, 407–415.
- Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Gen.* **40**, 1375–1383.
- Gangaraju, V. K., and Lin, H. (2009). MicroRNAs: key regulators of stem cells. *Nat. Rev. Mol. Cell Biol.* **10**, 116–125.
- Gregory, R. I., Chendrimada, T. P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640.
- Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008). miRBase: Tools for microRNA genomics. *Nucleic Acids Res.* **36**, D154–D158.
- Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Bailie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34.
- Hackenberg, M., Sturm, M., Langenberger, D., Falcon-Perez, J. M., and Aransay, A. M. (2009). miRanalyzer: A microRNA detection and analysis tool for next-generation sequencing experiments. *Nucleic Acids Res.* **37**, W68–W76.
- Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., Holoch, D., Lim, C., and Tuschl, T. (2008). Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12.
- Hajarnavis, A., Korf, I., and Durbin, R. (2004). A probabilistic model of 3' end formation in *Caenorhabditis elegans*. *Nucleic Acids Res.* **32**, 3392–3399.
- Harris, T. W., Antoshechkin, I., Bieri, T., Blasiar, D., Chan, J., Chen, W. J., De La Cruz, N., Davis, P., Duesbury, M., Fang, R., Fernandes, J., Han, M., Kishore, R., Lee, R., Muller, H. M., Nakamura, C., Ozersky, P., Petcherski, A., Rangarajan, A., Rogers, A., Schindelman, G., Schwarz, E. M., Tuli, M. A., Van Auken, K., Wang, D., Wang, X., Williams, G., Yook, K., Durbin, R., Stein, L. D., Spieth, J., and Sternberg, P. W. (2010). WormBase: A comprehensive resource for nematode research. *Nucleic Acids Res.* **38**, D463–D467.
- Hillier, L. W., Reinke, V., Green, P., Hirst, M., Marra, M. A., and Waterston, R. H. (2009). Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res.* **19**, 657–666.
- Hofacker, I. L. (2009). RNA secondary structure analysis using the Vienna RNA package. *Curr Protoc Bioinformatics (Chapter 12, Unit 12. 2)* **26**, 12.2.1–12.2.16.
- Hope, I. A., Stevens, J., Garner, A., Hayes, J., Cheo, D. L., Brasch, M. A., and Vidal, M. (2004). Feasibility of genome-scale construction of promoter::reporter gene fusions for expression in *Caenorhabditis elegans* using a multisite gateway recombination system. *Genome Res.* **14**, 2070–2075.
- Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. *Gene* **77**, 61–68.
- Hutvagner, G., Simard, M. J., Mello, C. C., and Zamore, P. D. (2004). Sequence-specific inhibition of small RNA function. *PLoS Biol.* **2**, E98.

- Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647.
- Johnson, S. M., Lin, S. Y., and Slack, F. J. (2003). The time of appearance of the *C. elegans* let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259**, 364–379.
- Kai, Z. S., and Pasquinelli, A. E. (2010). MicroRNA assassins: Factors that regulate the disappearance of miRNAs. *Nat. Struct. Mol. Biol.* **17**, 5–10.
- Kato, M., de Lencastre, A., Pincus, Z., and Slack, F. J. (2009). Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. *Genome Biol.* **10**, R54.
- Kato, M., and Slack, F. J. (2008). microRNAs: Small molecules with big roles – *C. elegans* to human cancer. *Biol. Cell* **100**, 71–81.
- Kramer, J. M., French, R. P., Park, E. C., and Johnson, J. J. (1990). The *Caenorhabditis elegans* rol-6 gene, which interacts with the sqt-1 collagen gene to determine organismal morphology, encodes a collagen. *Mol. Cell Biol.* **10**, 2081–2089.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858.
- Lall, S., Grun, D., Krek, A., Chen, K., Wang, Y. L., Dewey, C. N., Sood, P., Colombo, T., Bray, N., Macmenamin, P., Kao, H. L., Gunsalus, K. C., Pachter, L., Piano, F., and Rajewsky, N. (2006). A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr. Biol.* **16**, 460–471.
- Latronico, M. V., and Condorelli, G. (2009). MicroRNAs and cardiac pathology. *Nat. Rev. Cardiol.* **6**, 419–429.
- Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862.
- Lee, R. C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864.
- Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.
- Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20.
- Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: Short oligonucleotide alignment program. *Bioinformatics* **24**, 713–714.
- Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., and Wang, J. (2009). SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967.
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F., and Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380.
- Martinez, N. J., Ow, M. C., Reece-Hoyes, J. S., Barrasa, M. I., Ambros, V. R., and Walhout, A. J. (2008). Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity. *Genome Res.* **18**, 2005–2015.
- Medina, P. P., and Slack, F. J. (2008). microRNAs and cancer: An overview. *Cell Cycle* **7**, 2485–2492.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.

- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Mili, S., and Steitz, J. A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: Implications for the interpretation of immunoprecipitation analyses. *RNA* **10**, 1692–1694.
- Moxon, S., Schwach, F., Dalmay, T., Maclean, D., Studholme, D. J., and Moulton, V. (2008). A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics* **24**, 2252–2253.
- Negrini, M., Nicoloso, M. S., and Calin, G. A. (2009). MicroRNAs and cancer – new paradigms in molecular oncology. *Curr. Opin. Cell Biol.* **21**, 470–479.
- O’Connell, R. M., Rao, D. S., Chaudhuri, A. A., and Baltimore, D. (2010). Physiological and pathological roles for microRNAs in the immune system. *Nat. Rev. Immunol.* **10**, 111–122.
- Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M. C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* **18**, 1655–1666.
- Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Muller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., and Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86–89.
- Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W., and Sontheimer, E. J. (2004). A Dicer-2-dependent 80 s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* **117**, 83–94.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906.
- Roush, S. F., and Slack, F. J. (2009). Transcription of the *C. elegans* let-7 microRNA is temporally regulated by one of its targets, hbl-1. *Dev. Biol.* **334**, 523–534.
- Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**, 1193–1207.
- Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145.
- Steiner, F. A., Hoogstrate, S. W., Okihara, K. L., Thijssen, K. L., Ketting, R. F., Plasterk, R. H., and Sijen, T. (2007). Structural features of small RNA precursors determine Argonaute loading in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **14**, 927–933.
- Subramanian, S., and Steer, C. J. (2010). MicroRNAs as gatekeepers of apoptosis. *J. Cell Physiol.* **223**, 289–298.
- Tyagi, S., Vaz, C., Gupta, V., Bhatia, R., Maheshwari, S., Srinivasan, A., and Bhattacharya, A. (2008). CID-miRNA: A web server for prediction of novel miRNA precursors in human genome. *Biochem. Biophys. Res. Commun.* **372**, 831–834.
- Van Wynsberghe, P. M., Kai, Z. S., Massire, K. B., Burton, V. H., Yeo, G. W., and Pasquinelli, A. E. (2011). LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **18**, 302–308.
- Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K., and Slack, F. J. (2004). The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3’UTR. *Genes Dev.* **18**, 132–137.
- Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000). GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* **328**, 575–592.
- Wilm, T., Demel, P., Koop, H. U., Schnabel, H., and Schnabel, R. (1999). Ballistic transformation of *Caenorhabditis elegans*. *Gene* **229**, 31–35.
- Winter, J., Jung, S., Keller, S., Gregory, R. I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **11**, 228–234.
- Xu, Y., Zhou, X., and Zhang, W. (2008). MicroRNA prediction with a novel ranking algorithm based on random walks. *Bioinformatics* **24**, i50–i58.

- Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J., and Mello, C. C. (2006). Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**, 747–757.
- Zhang, L., Ding, L., Cheung, T. H., Dong, M. Q., Chen, J., Sewell, A. K., Liu, X., Yates 3rd, J. R., and Han, M. (2007). Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol. Cell* **28**, 598–613.
- Zisoulis, D. G., Lovci, M. T., Wilbert, M. L., Hutt, K. R., Liang, T. Y., Pasquinelli, A. E., and Yeo, G. W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **17**, 173–179.
- Zisoulis, D. G., Yeo, G. W., and Pasquinelli, A. E. (2011). Comprehensive identification of miRNA target sites in live animals. *Methods Mol. Biol.* **732**, 169–185.

---

---

## CHAPTER 9

# *In situ* Hybridization of Embryos with Antisense RNA Probes

**Gina Broitman-Maduro and Morris F. Maduro**

Department of Biology, University of California at Riverside, Riverside, California, USA

---

- Abstract
- I. Introduction
- II. Rationale
- III. Methods
  - A. Probe Design and Synthesis
  - B. PCR to Generate Probe Synthesis Template
  - C. DIG-Labeled Probe Synthesis
- IV. Animal Preparation and Fixation
  - A. Synchronization to Produce Gravid Hermaphrodites
  - B. Fixation of Gravid Hermaphrodites and Embryos
  - C. Freeze-Cracking and Fixation
  - D. Hydration Series
- V. Hybridization and Signal Development
  - A. Prehybridization and Probe Hybridization
  - B. Rinsing and Antibody Incubation
  - C. Rinsing and Signal Development
- VI. Materials
- VII. Notes
  - A. Suggested Positive Controls
  - B. Hypaque meglumine
  - C. Nontoxic Fixative
  - D. Purchase of Ready-Made Reagents
  - E. Labeling of Nuclei
  - F. Staining Small Quantities of Embryos or Other Stages
  - G. Staining of Other Species
- VIII. Troubleshooting Guide
- IX. Imaging of Stained Embryos
- X. Summary
- Acknowledgments
- References

---

---

## Abstract

Detection of transcripts *in situ* is a rapid means by which gene expression can be characterized in many systems. In the nematode, *Caenorhabditis elegans*, the ease with which transgenics can be made and the general reliability of reporter fusion expression patterns, have made this technique comparatively less popular than in other systems. There are, however, still applications in which *in situ* hybridization is desired, such as for maternally expressed genes, or in related species without established transgene methods. The most frequently used method of *in situ* hybridization uses DNA probes and formaldehyde fixation. A newer approach that permits single-transcript detection has been reported and will not be described here (Raj and Tyagi, 2010). Rather, we describe an alternative protocol that uses RNA probes with a different fixative. This approach has been applied to *C. elegans* and related nematodes, providing reliable, sensitive detection of endogenous transcripts.

---

---

## I. Introduction

Localization of transcripts by *in situ* hybridization is a desirable way to determine expression patterns, because it can detect endogenous mRNA in its natural context, and because it is a method that, once established, can be repeated on any number of genes by changing only the antisense probe used. In addition, regulatory mechanisms might also be identified, such as subcellular localization of the transcripts. Molecular approaches, such as quantitative PCR (qPCR), Northern blots, or genome-wide approaches such as microarrays or RNA-Seq require isolation of tissues, and in most cases, amplification of the endogenous material. In the nematode, *Caenorhabditis elegans*, *in situ* hybridization has historically not been the method of choice for assessing endogenous gene expression, due largely to the ease of construction of transgenic reporter strains and the general reliability of the expression patterns produced (see Chapter on Transgenesis in this volume). Nonetheless, there remain instances in which *in situ* detection of endogenous mRNA may be desired. Reporters may be difficult to construct for particular genes, or transgenes may not express, in particular those activated in the *C. elegans* germ line and very early embryo. Other nematode species remain refractory to transgene expression techniques, either because the DNA becomes silenced in the soma as well as germ line, or because the necessary reagents (e.g., specific mutant backgrounds in which to make transgenics) do not yet exist. Newer transgene protocols may overcome some of these limitations (Giordano-Santini *et al.*, 2010; Praitis *et al.*, 2001; Schlager *et al.*, 2009; Semple *et al.*, 2010); however, it remains to be seen whether transgenes in other nematode species will in general be as reliable as those seen in *C. elegans*.

Historically, *in situ* detection of mRNA has relied on detection of colorimetric or fluorescent signals from localized antisense DNA probes in whole-mount embryos (Seydoux and Fire, 1995; Tabara *et al.*, 1996). For low-abundance transcripts, signal amplification can be used (Bobrow and Moen, 2001). Recently, a new approach

using multiple short, nonoverlapping fluorescent probes has been described (Raj and Tyagi, 2010; Raj *et al.*, 2008). This approach permits detection of mRNAs as diffraction-limited spots, with the advantage that it is highly sensitive, yet specific, and permits quantification of individual transcripts. As this protocol has very recently been described in detail (Raj and Tyagi, 2010; Raj *et al.*, 2008), we shall instead focus here on an alternative protocol that uses a nontoxic fixative (NTF) and less-expensive antisense RNA probes, which offers a qualitative assessment of gene expression. For the classic *C. elegans in situ* hybridization protocol, readers are referred to the original description (Seydoux and Fire, 1995), updated versions of which can be found on the WormMethods section of WormBook ([http://www.wormbook.org/toc\\_wormmethods.html](http://www.wormbook.org/toc_wormmethods.html)). The procedure described here can be performed in less than 3 days and offers good sensitivity with preservation of fine structure. It has been used to successfully detect embryonic transcripts in *C. elegans* and other nematode species (Broitman-Maduro *et al.*, 2009; Coroian *et al.*, 2005). Others have adapted the procedure, for example, for detection of transcripts in extruded *C. elegans* gonads (Sheth *et al.*, 2010).

---

---

---

## II. Rationale

In the method described here, whole embryos are mounted on coated glass microscope slides, fixed, and permeabilized. Antisense RNA probes are synthesized by *in vitro* transcription of PCR products, and include the use of Digoxigenin-tagged UTP. The use of RNA probes may improve sensitivity because RNA:RNA hybrids are more stable than DNA:RNA hybrids (Sugimoto *et al.*, 1995). Following hybridization of the probe, the slides are rinsed and processed for detection of the DIG moiety using an anti-DIG antibody conjugated to alkaline phosphatase (AP). The bound antibodies are detected by the use of the standard AP substrates Nitro blue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP). Staining is observed with a microscope equipped with differential interference contrast (DIC), and color images are acquired with a digital camera. Positive controls are performed with known probes, or on transgene strains with known expression. Use of a sense RNA probe, or a mutant background known to result in the absence of the endogenous transcript, can each serve as negative controls.

---

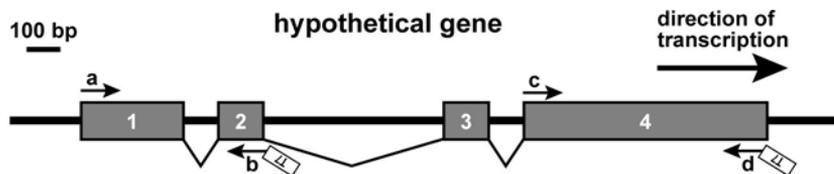
---

---

## III. Methods

### A. Probe Design and Synthesis

Antisense RNA probes are created from *in vitro* transcription of short (200 bp–1.5 kbp) PCR products carrying the T7 RNA polymerase recognition sequence at one end. The most convenient template is genomic DNA, using primers that will amplify as high a proportion of exon-containing sequence as possible (Fig. 1). Alternatively, a cloned cDNA fragment can be used. Some may choose to also synthesize the



**Fig. 1** Typical primer design considerations for a hypothetical four-exon gene. A good set of primers might be the pairs indicated by a/b or c/d. The 3' primer carries the recognition sequence for T7 RNA polymerase, 5'-TAATACGACTCACTATAGGG-3', followed by 25–30 bases of target homology; the forward primer has 25–30 bases of identity but lacks the T7 tag.

complementary (sense) strand as a control probe, or use one of the suggested probes (Section 4) for specificity.

## B. PCR to Generate Probe Synthesis Template

Assemble the following in a 0.6-mL (PCR) tube:

Forward primer (25 pmol/ $\mu$ L)	1 $\mu$ L
Reverse primer (25 pmol/ $\mu$ L)	1 $\mu$ L
dNTPs (10 mM)	0.5 $\mu$ L
PCR buffer (10 $\times$ stock)	2.5 $\mu$ L
Genomic DNA (200 ng/ $\mu$ L) <sup>a</sup>	1 $\mu$ L
Taq polymerase <sup>b</sup>	0.5 $\mu$ L
ddH <sub>2</sub> O	18.5 $\mu$ L (total volume: 25 $\mu$ L)

<sup>a</sup> Alternatively, use 1  $\mu$ L of a solution carrying 10 ng of plasmid template.

<sup>b</sup> We routinely use a crude Taq preparation with good results (Engelke *et al.*, 1990).

Perform a standard PCR reaction, for example 95°C for 3 min and then repeat [95°C for 30 s, 72°C for 30 s, 55°C for 30 s] for 30 cycles, 72°C for 10 min and ending at 4°C. Check an aliquot (5  $\mu$ L) on an agarose gel to make sure the PCR product is of the expected size.

## C. DIG-Labeled Probe Synthesis

Use DIG-RNA labeling kit (Roche, #1175025)

PCR product generated above (not purified)	3 $\mu$ L
10 $\times$ NTP mixture with DIG-11-UTP	1 $\mu$ L
10 $\times$ Transcription Buffer	1 $\mu$ L
RNase Inhibitor	0.5 $\mu$ L

T7 RNA Polymerase	1 $\mu\text{L}$
DEPC-ddH <sub>2</sub> O	3.5 $\mu\text{L}$ (total volume: 10 $\mu\text{L}$ )

Incubate in a thermocycler with a heated lid at 37°C, overnight. In the morning add 30  $\mu\text{L}$  DEPC-ddH<sub>2</sub>O and 0.5  $\mu\text{L}$  RNase Inhibitor to the incubations and store at -20°C.

---

---

---

## IV. Animal Preparation and Fixation

### A. Synchronization to Produce Gravid Hermaphrodites

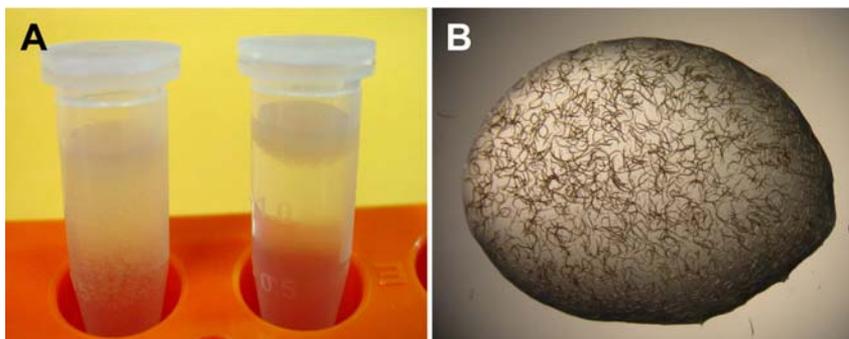
Hermaphrodites must be synchronized for *in situ* experiments because the fixation and freeze-crack techniques work best on animals that are similarly sized, rather than a mixture of adults and larvae. In order to synchronize worms, two to three 10-cm plates containing gravid hermaphrodites are bleached and the embryos are hatched overnight in M9 + cholesterol [10  $\mu\text{g}/\text{mL}$ ] at 20°C. Synchronized L1 larvae are collected by centrifugation and plated on 10-cm NGM plates for 3 days at 20°C. After approximately 3 days the plates should contain gravid hermaphrodites with early embryos. If later-staged embryos are required incubation may be extended a further 5 h, but will depend on incubation temperature, abundance of food, general health of the strain, and tendency of the worms to retain embryos.

### B. Fixation of Gravid Hermaphrodites and Embryos

Harvest the worms by washing each 10-cm plate with 3 mL of M9 into several Eppendorf tubes and centrifuging at 2000 rpm. This procedure often carries bacterial contamination along with gravid hermaphrodites, depending on how much bacteria remains on the plates. Excess bacteria will result in slides with high background after development. We routinely resuspend the worm pellet in 1 mL of Hypaque meglumine (60%, available from <http://www.nanric.com> as Reno-60). Invert the tube several times and centrifuge for 30 s at 2000 rpm. After centrifugation, the worms will float to the top of the Eppendorf tube (Fig. 2A). Remove the worm layer and rinse 2–3 $\times$  in M9. Examine an  $\sim 80$   $\mu\text{L}$  droplet of worms on a microscope slide under a dissecting microscope. If there are no bacterial clumps, then the worms are clean and ready for freeze-cracking and fixation (Fig. 2B). If many bacterial clumps are still present, repeat the meglumine flotation until worms appear clean. Resuspend worms in 750  $\mu\text{L}$  of M9.

### C. Freeze-Cracking and Fixation

The following steps are carried out in RNase-free glassware, which should be designated for RNA use and be rinsed with DEPC-ddH<sub>2</sub>O and baked for 4 h at 180°C. Alternatively, glassware may be cleaned thoroughly with RNase away (MBP, #7003, 1L). Where indicated, disposable plastics may be used for convenience.



**Fig. 2** (A) Appearance of worms in Hypaque meglumine in 1.5-mL Eppendorf tubes. The tube on the left has been inverted and worms are suspended throughout. After 5 min at room temperature, or after a brief 2000 rpm centrifugation, the worms will float to the top (right tube) and can be removed by pipetting the top layer. (B) 80  $\mu$ L droplet of worms on a glass slide. (For color version of this figure, the reader is referred to the web version of this book.)

Prepare a hydration series of methanol:DEPC-ddH<sub>2</sub>O dilutions in five clean 50-mL Coplin jars as follows:

- #1 – methanol pre-chilled at  $-20^{\circ}\text{C}$
- #2 – 90% methanol at room temperature
- #3 – 70% methanol at room temperature
- #4 – 50% methanol at room temperature
- #5 – DEPC-ddH<sub>2</sub>O at room temperature

Prepare 50 mL of NTF in a Coplin jar and warm it to  $37^{\circ}\text{C}$  in an incubator.

Place an aluminum disc (Fig. 3) onto crushed dry ice for at least 5 min and bring the disc, still on dry ice, near a dissecting microscope. Make sure the block is smooth and free from surface irregularities, to allow microscope slides to make good contact with it.



**Fig. 3** Aluminum disc, approximately 125 mm in diameter  $\times$  15 mm, on dry ice inside a Styrofoam shipping carton.

Resuspend worms in the M9 by inversion. Pipette 80  $\mu\text{L}$  of worm suspension onto the center of a poly-L-lysine coated slide (Fisher, #12-550-19) and spread the droplet with a clean razor blade so that the worms come in contact with the surface of the slide. Begin cutting worms with the razor in an up-down motion while observing the worms through the dissecting microscope. Stop when the majority of worms are cut in half and the embryos are liberated. Starting at one side, cover the worms with a 22  $\times$  40 mm cover glass (Gold Seal #3316), taking care not to introduce air bubbles. While looking at the worms through the dissecting microscope, wick away excess liquid by placing the edge of a Kimwipe or paper towel in contact with the edge of the coverslip. This should be stopped just as adult worm carcasses cease to move but before embryos burst open (Fig. 4). As this step is critical for proper permeabilization, it may need to be practiced until the researcher is confident that worms and embryos have adhered properly to the slide. When a slide is ready, place it coverslip side up on the aluminum disc and press down on the side that has no coverslip to ensure complete contact. Incubate slides at least 5 min before proceeding to the hydration series.

#### D. Hydration Series

Wedge a clean razor blade under one corner of the coverslip, and with a twisting motion quickly pop off the cover glass in one motion. There should be an audible “cracking” sound as the coverslip is popped off. Incubate the slide as follows:

- 1) 5 min – 100% methanol (at  $-20^{\circ}\text{C}$ )
- 2) 5 min – 90% methanol (room temperature)
- 3) 5 min – 70% methanol (room temperature)
- 4) 5 min – 50% methanol (room temperature)
- 5) 5 min – DEPC-ddH<sub>2</sub>O (room temperature)
- 6) 1 h – NTF (prewarmed to  $37^{\circ}\text{C}$ )



**Fig. 4** Appearance of desired density of worms and embryos under coverslip, prior to freezing, as viewed through a dissecting microscope.

Place Coplin jar, with NTF and slides, into a 37°C incubator during the 1 h incubation.

During incubation, prepare 10 mL prehybridization buffer (PB) and incubate at 65°C with occasional vortexing to resuspend components. PB may require up to 30 min at 65°C to become fully resuspended.

## ===== V. Hybridization and Signal Development

### A. Prehybridization and Probe Hybridization

Rinse slides in the following buffers:

- 1) 5 min – DEPC-ddH<sub>2</sub>O
- 2) 5 min – DEPC-ddH<sub>2</sub>O
- 3) 5 min – 2× SSC
- 4) 5 min – 2× SSC

Have a humid chamber ready. It is convenient to use empty pipette tip boxes, with the rack in place, with water in the bottom to a depth of a few centimeters. Place slides face up in the chamber and add 300 μL PB onto fixed worms. Be sure that samples are completely covered with PB. Place humid chambers into a hybridization oven prewarmed to 42°C and incubate 1 h. There is no need to cover the slides with a coverslip as the hybridization buffer will not evaporate under these conditions.

Towards the end of the incubation, thaw frozen probe(s) and dilute 1:500 to 1:1000 in PB (typically, 1–2 μL probe in 1 mL PB) and heat to 65°C for 5 min. Add 100 μL of the diluted probe to the PB already on the top of the slides. Alternatively, tilt the slides so that the PB runs off, and add 300 μL of the PB+Probe to the slides to cover the worms. Hybridize at 42°C overnight. It is convenient to prepare two 50 mL aliquots each of 2× SSC and formamide buffer (FB) and incubate them at 42°C overnight. These will be used for rinses on the following day.

### B. Rinsing and Antibody Incubation

Place slides back into Coplin jars with the following buffers at 42°C:

- 1) 5 min – 2× SSC (42°C)
- 2) 5 min – 2× SSC (42°C)
- 3) 5 min – FB (42°C)
- 4) 5 min – FB (42°C)

Rinse slides in the following buffers at room temperature:

- 1) 5 min – 2× SSC (room temperature)
- 2) 5 min – 2× SSC (room temperature)
- 3) 5 min – Tris-NaCl (TN) (room temperature)
- 4) 5 min – TN (room temperature)

Following the TN rise, block slides prior to antibody incubation at 37°C in 30 mL blocking buffer (BB) for 30 min.

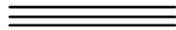
Add 10 µL of anti-Digoxigenin-AP Fab fragments (Roche, #11093274910) to 30 mL fresh BB and incubate for 3 h at 37°C.

### C. Rinsing and Signal Development

Rinse slides in the following buffers:

- 1) 10 min – TN (room temperature)
- 2) 10 min – TN (room temperature)
- 3) 10 min – TNM pH 9.5 (room temperature)
- 4) Developer – incubate with slides overnight in the dark (e.g., by covering entire Coplin jar in aluminum foil) at room temperature.

The next day, rinse twice in TN-EDTA for 10 min each, and add mounting medium (e.g., Vectashield, Vector Labs). Cover with 22×40 mm coverslip, wipe off excess mounting medium, and seal the coverslip with clear nail polish. Slides can be kept up to 6 months at 4°C, but for best image quality observe within a week. After several weeks, a colored precipitate will begin to form.



## VI. Materials

1. DEPC-ddH<sub>2</sub>O (0.1% v/v)

Millipore water or equivalent	500 mL
DEPC (Sigma)	0.5 mL

Suspend the DEPC by shaking vigorously, leave in a fume hood over night, and autoclave (121°C at 15 lb/in<sup>2</sup> for 15 min).

2. NTF

2-bromo-2-nitro-1,3-propanediol (Fisher, # AC15882-1000)	30 g
Diazolidinyl Urea (MP Biomedical, # ICN19019183)	30 g
Zinc Sulfate Heptahydrate, ZnSO <sub>4</sub> ·7H <sub>2</sub> O	12 g
Sodium Citrate (anhydrous), Na <sub>3</sub> C <sub>6</sub> H <sub>5</sub> O <sub>7</sub>	2.9 g

Add DEPC-ddH<sub>2</sub>O up to 1 L. Make and store in an autoclaved glass container at room temperature. Heat 30–50 mL to 37°C in a Coplin jar prior to use (Of this heated aliquot, do not add unused fixative back to stock, and do not reuse.).

## 3. 50× Denhardt's solution

Ficoll (Type 400)	0.1 g
Polyvinylpyrrolidone (PVP)	0.1 g
Bovine Serum Albumin (BSA)	0.1 g
DEPC-ddH <sub>2</sub> O	to 10 mL

Dissolve components, pass through a 0.45 μm filter, and store in 200 μL aliquots at −20°C.

## 4. Salmon sperm DNA

Lyophilized DNA	1 g
Tris-EDTA pH 8.0 (TE, Fisher #BP2473-1)	100 mL

Add TE to DNA in an autoclavable glass Wheaton bottle. Autoclave (121°C at 15 lb/in<sup>2</sup> for 15 min), cool to room temperature, dispense into Eppendorf tubes in 1 mL aliquots, and freeze at −20°C. As a result of autoclaving, the DNA will have been sheared to the 200 bp–5 kbp range, which can be checked by running an aliquot on an agarose gel.

## 5. Prehybridization buffer

20× SSC (Fisher, #BP1325-1)	2 mL
Dextran Sulfate	1 g
Formamide	5 mL
0.5 M EDTA	40 μL
50× Denhardt's Solution	200 μL
Sheared Salmon Sperm DNA	1 mL

Add DEPC-ddH<sub>2</sub>O up to 10 mL. Vortex the suspension and heat to 65°C to get into solution. This may take 30 min or more.

## 6. Formamide buffer (FB)

Formamide	30 mL
20× SSC	0.5 mL

Add up to 50 mL with DEPC-ddH<sub>2</sub>O in sterile plastic Falcon tube or equivalent.

## 7. Tris pH 7.5, NaCl (TN), and TN-EDTA

1M Tris, pH 7.5 (Fisher, #1757-500)	50 mL
2M NaCl	37.5 mL

Add up to 500 mL DEPC-ddH<sub>2</sub>O and autoclave.

For TN-EDTA, add 1 mL of 0.5M EDTA to 50 mL of TN.

## 8. Blocking buffer

Milk Blocker (Bio-Rad # 170-6404)	2.5 g
TN	to 50 mL

Make in sterile plastic 50 mL Falcon tubes. Suspend by vortexing.

## 9. 0.93M Tris, pH 9.5

1M Tris, pH 7.5 (Fisher, BP1757500)	500 mL
10M NaOH	43.5 mL

Dispense aliquots into sterile 50-mL Falcon tubes and freeze at  $-20^{\circ}\text{C}$ . May be thawed in  $37^{\circ}\text{C}$  incubator prior to use. (We have found pH adjustment of the RNase-free Fisher stock to be a more convenient way to make this buffer than to make 1M Tris pH 9.5 by other means.)

10. Tris, NaCl,  $\text{MgCl}_2$ , (pH 9.5) ( $2\times$  TNM),

0.93M Tris (pH 9.5) (see above)	100 mL
2M NaCl	50 mL
1M $\text{MgCl}_2$	50 mL
DEPC-dd $\text{H}_2\text{O}$	to 500 mL

Autoclave for 15 min at  $121^{\circ}\text{C}$ , 15 lb/in<sup>2</sup>. Dispense aliquots into 50 mL Falcon tubes. Store at  $-20^{\circ}\text{C}$ . Thaw the morning that it is needed. If a white precipitate occurs, heat in a  $65^{\circ}\text{C}$  waterbath and vortex, then cool before use.

To 25 mL of  $2\times$  TNM, add 25 mL DEPC-dd $\text{H}_2\text{O}$ . This is used as the last wash prior to development. The other 25 mL aliquots are used to make the Developer solution.

## 11. Nitro Blue Tetrazolium (NBT)

NBT is resuspended to 100 mg/mL in 70% dimethyl formamide and stored in aliquots at  $-20^{\circ}\text{C}$ .

## 12. 5-Bromo-4-Chloro-3-Indolyl Phosphate (BCIP)

BCIP is resuspended to 50 mg/mL in anhydrous dimethyl formamide and stored in aliquots at  $-20^{\circ}\text{C}$ .

13.  $2\times$  Polyvinyl alcohol (PVA)

PVA (Fluka 40-88, Sigma #81386)	50 g
DEPC-dd $\text{H}_2\text{O}$	to 500 mL

This solution takes very long to go into solution. It is best to start with freshly autoclaved (warm) DEPC-dd $\text{H}_2\text{O}$  and add PVA to a sterile container. Incubate at  $65^{\circ}\text{C}$  for several days, stirring often with a 25 mL plastic serological pipette; A stir bar is not recommended. When it is fully in suspension, the solution will have the consistency of glycerol. Store at room temperature.

In our hands, PVA in the developer greatly improves signal quality, and is worth the trouble to make.

#### 14. Developer Solution

2× TNM	25 mL
2× PVA	25 mL
NBT	200 μL
BCIP	200 μL
Levamisole	7.2 mg

Mix reagents in a Falcon tube and mix by inversion. Prepare immediately before use.

## ==== VII. Notes

### A. Suggested Positive Controls

Gene	Expression Pattern	Primers (forward primer, T7-tagged reverse primer)
<i>pgl-1</i>	Germ line and embryos	5'-gtt caaggaatcaactcgaag act c-3', 5'-TAA TAC GAC TCA CTA TAG GGA CTtggcagagct act gat ttcgtt gga-3'
<i>end-1</i>	Early E lineage (E, Ea/Ep)	5'-ttc aatcgtacgatc cag cacaacaat cg-3', 5'-TAA TAC GAC TCA CTA TAG GGA CTT GCA CT caatagctcctgaatcagt t-3'
<i>hll-1</i>	Mid to late embryogenesis, muscle lineages	5'-aaaccagccagcttactacctcccgctcta-3', 5'-TAATACGACTCACTATAGGGACTcgttcccga gcttatgatgatctctatc-3'
<i>opt-2</i>	Mid to late embryogenesis, intestine	5'-gtaatggcgattggactctcacatagacc-3', 5'-TAATACGACTCACTATAGGGACTtctgctgccg tgatgacaacgattgtg-3'
<i>myo-2</i>	Late stage embryos, pharynx muscle	5'-gga tttgtccaaga gat gaatccacc a-3', 5'-TAA TAC GAC TCA CTA TAG GGA CTtgttcaa tat cgcaagaagcgacac gtc-3'
<i>myo-3</i>	Late stage embryos, body muscle	5'-tct cgtgtcctcctcag gcaccagga gag-3, 5'-TAA TAC GAC TCA CTA TAG GGA CTcctggtgatg atccactgaacata cgg-3'
<i>cup-4</i>	4 coelomocytes in embryos, 6 in adults	5'-gtagtagcatctctaataatcatgacgttc-3', 5'-TAATACGACTCACTATAGGGACTtccttgaac gtattaggaatgtattctt-3'

### B. Hypaque meglumine

Bacterial contamination (usually *E. coli* OP50 from plates) often forms clumps around worms and deposits onto slides, attracting probe and contributing to high background. In our hands, cleaning worms with meglumine gives the most consistent results. It may be possible to bypass this step if the source plates are clean enough (David Fitch, personal communication). Alternatively, sucrose flotation may be tried (suspend animals in 30% sucrose at 4°C and centrifuge at 4000 rpm).

### C. Nontoxic Fixative

Although formaldehyde is a widely used fixative, NTF is an alternative to formaldehyde that is easily made, stable, and nontoxic. More importantly, we have found it to greatly improve sensitivity and preservation of fine structure, and it appears to be less prone to overfixation. This protocol was originally developed using STF (Streck Tissue Fixative; Streck Laboratories; [Montgomery et al., 1998](#)), but this reagent has been discontinued. NTF is based on the composition of STF (documented in United States Patent 5460797) and in our hands appears to be equally effective.

### D. Purchase of Ready-Made Reagents

Many of the reagents in this protocol, such as 2× SSC and 1 M Tris-HCl, can be easily made. The ready-made forms are purchased only to make it less likely that RNase contamination might be introduced.

### E. Labeling of Nuclei

We have found that use of fluorescence (e.g., to visualize DAPI) causes the purple color to develop very rapidly in a nonspecific manner. If the slides are dehydrated and mounted in a permanent mounting medium (e.g., Permount with DAPI) the nuclei can be visualized, although the morphology is adversely affected.

### F. Staining Small Quantities of Embryos or Other Stages

A larger number of animals is necessary for consistent freeze-cracking. When the number of specimens is limiting, sterile adults of recognizable body morphology could be added (e.g., Dpy; Glp). If it is desired to stain larvae or adults, they should be synchronized so that they are all the same size prior to mounting on the slides.

### G. Staining of Other Species

We and others have been successful in staining other nematode species, including *C. remanei* and *C. briggsae* ([Coroian et al., 2005](#); [Lin et al., 2009](#)) and

*Pristionchus pacificus* (George Hsu, Heather Roberson, G.B.-M. and M.M., unpublished observations) with species-specific probes. One consideration for embryonic staging is the degree to which embryos are retained by adults. In our hands, *C. remanei* (strain PB4641) and *C. briggsae* (AF16) retain fewer embryos than *C. elegans*. Possible solutions are to limit the amount of food to promote egg retention; use of a mutant background (e.g., Egl) that retains eggs (we have used *ir12*, an uncharacterized recessive *dpy* mutant of *C. briggsae* with good results); to bleach gravid adults and let the isolated embryos develop for several hours; or to bleach to isolate embryos from a plate. In our experience, bleaching solution negatively affects staining quality, so it is recommended to use a diluted bleach solution and rinse animals thoroughly. For consistent freeze-cracking, adults can be added back as suggested above.

## VIII. Troubleshooting Guide

Problem	Cause	Remedy
Lack of signal	Lack of template for transcription	Confirm amplification of PCR product with gel electrophoresis.
	Error in T7 sequence	Confirm that primer sequence was correct as ordered; subclone PCR product into T/A vector and sequence it; be certain correct (reverse) primer had the T7 sequence, not the sense primer
	Wrong synthesis kit or components	Confirm use of all components from a DIG-UTP kit, not DIG-dUTP (used for DNA probes)
	Poor transcription	Run side-by-side aliquots of (+)T7 and (-)T7 transcription reactions to confirm synthesis of RNA
	Developer problem	Spot the probe onto a blotting membrane (e.g., nitrocellulose or Hybond-N+) and test for color turnover. Try a previously successful probe. Remake developer components and solutions. (It is recommended to include a positive control in all experiments to rule out problems with the reagents or solutions.)

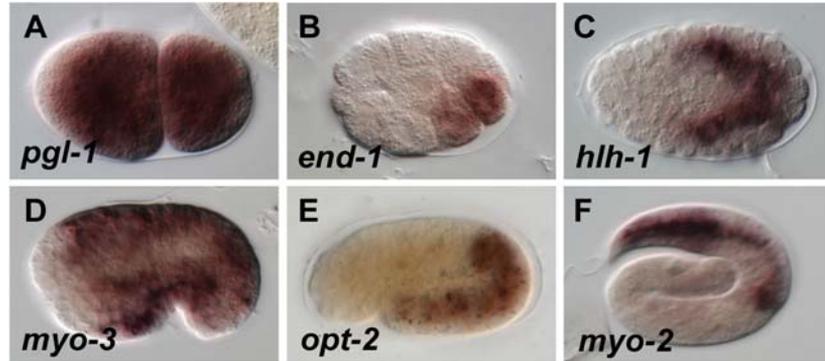
(Continued)

(Continued)

Problem	Cause	Remedy
Slides have very high background or are all purple	Probe excess (signal throughout embryos and adults; developer may turn purple overnight)	Dilute probe further. Typically 1 $\mu$ L of probe, diluted in 1 mL of hybridization buffer, is sufficient
	Inconsistent freeze-cracking technique (nonspecific signal on “top” of animals)	<ol style="list-style-type: none"> <li>1. When wicking liquid away from the animals, avoid having floating worm carcasses (excess liquid left) or burst embryos (too little liquid). If slight pressure is applied to the coverslip and the specimens do not float around but swell slightly in response to the pressure, then it is likely a good mount.</li> <li>2. When “cracking” the cover slip off, make sure you feel pressure as the cover slip is removed. If it comes off too easily, then the outer cuticle or egg shell of the animals will likely not be permeabilized properly</li> </ol>
	Choice of probe	Try a different or longer probe to your gene of interest
Uneven signal (some staining that looks real, but some areas with nothing)	Overcrowding of animals	This is usually the result of inconsistent freeze-cracking. Make sure animals are evenly dispersed on the slide prior to adding the coverslip. Try using fewer animals. Make sure there is no debris or worm clumps that would hold up the coverslip in parts

## IX. Imaging of Stained Embryos

For imaging of stained embryos, we use an Olympus BX51 equipped with DIC optics and a 60 $\times$ /1.4 oil immersion lens. It is highly recommended that a color camera be used to document *in situ* staining. DIC images of embryos can contain regions that are darker simply due to the nature of DIC itself. As the *in situ* staining appears brown or purple, faint signal can more readily be seen and distinguished from the surrounding tissue if a color camera is used. We have had good results with a consumer grade Canon APS-C-size sensor Digital SLR camera with a C-mount converter from LMscope (<http://www.lmscope.com/>)



**Fig. 5** Antisense RNA probes detect endogenous mRNAs consistent with published reports. (A) Maternal transcripts of *pgl-1* at the two-cell stage (Kawasaki *et al.*, 1998). (B) Expression of *end-1* in the E daughter cells (Zhu *et al.*, 1997). (C) Expression of *hlh-1* in muscle precursors (Krause *et al.*, 1990). (D) Activation of *myo-3* in body muscle cells (Okkema *et al.*, 1993). (E) Expression of *opt-2* in intestine cells (Nehrke, 2003). (F) Expression of *myo-2* in pharynx muscle cells (Okkema *et al.*, 1993). Anterior is left, and dorsal is up. Embryos are approximately 50  $\mu\text{m}$  long. (For color version of this figure, the reader is referred to the web version of this book.)

[index\\_e.html](#)), a platform that together costs approximately \$1500, and which is also suitable for fluorescence microscopy. It is also recommended that many animals (>25) of specific stages be examined for staining, and that the number of animals with good staining be quantified. Under optimal conditions, we routinely observe that at least 75%, and usually greater than 80%, of embryos will show detectable signal (Lin *et al.*, 2009). The sensitivity of the approach has been confirmed by quantification of transcripts by single molecule detection (Raj *et al.*, 2010). Quantification of endogenous transcripts of zygotic genes expressed in endoderm specification has shown that there are at most some 400 transcripts of *end-3* in the early E lineage (Raj *et al.*, 2010). We have observed very strong expression of *end-3* in the E cell of early embryos (Fig. 5B) (Maduro *et al.*, 2007), suggesting that this procedure is sensitive enough to detect several hundred transcripts. Given that the signals observed for *end-3* are fairly strong, it is likely that smaller numbers of transcripts (e.g., around 100) could be detected with this approach.

## ===== X. Summary

The detection of mRNA *in situ* provides a rapid means by which to determine the expression pattern of endogenous genes. A timetable is provided to assist in planning (Fig. 6). We have described a protocol that, in our hands, results in reproducible staining of endogenous mRNAs with a lower limit of at most several hundred

**in situ hybridization timetable**

<b><u>Advance preparation</u></b>	<b><u>Day 1</u></b>	<b><u>Day 2</u></b>	<b><u>Day 3</u></b>
<p><u>Prepare and autoclave:</u></p> <p>DEPC-ddH<sub>2</sub>O (5 L in 500 mL bottles)<sup>1</sup>            1M Tris-HCl, pH 7.5 (2 L)<sup>1</sup>            1M MgCl<sub>2</sub> (1 L)<sup>1</sup>            1M NaCl (1 L)<sup>1</sup>            20xSSC (1 L)<sup>1</sup>            0.93M Tris-HCl, pH 9.5 (1 L)<sup>2</sup>            10mg/mL Salmon Sperm DNA in TE<sup>2</sup>            0.5 M EDTA<sup>1</sup></p> <p><u>Prepare:</u></p> <p>10% PVA<sup>1</sup>            NBT<sup>2</sup>            BCIP<sup>2</sup>            Probe(s) (order oligos, PCR, synthesis)<sup>2</sup>            Denhardt's<sup>2</sup>            2xDeveloper<sup>2</sup></p> <p><u>Worm Culture</u></p> <p>Grow two 10 cm plates of worms until gravid and bleach. Allow to hatch in two falcon tubes with 5 mL M9 + cholesterol in each. Spin down 1 mL aliquots and plate 1mL large plate. Allow 3 days for worms to become gravid (e.g. Friday-Monday at 20°C).</p>	<p><u>Clean Glassware</u></p> <p>Rinse Coplin jars with RNase-away and DEPC-ddH<sub>2</sub>O</p> <p><u>Make working solutions</u></p> <p>NTF<sup>1</sup>            Hydration series<sup>3</sup>            2xSSC<sup>3</sup>            Prehybe buffer<sup>2</sup>            FB<sup>3</sup></p> <p><u>Experimental Technique</u></p> <p>Rinse, clean worms            Freeze crack worms            Hydration Series            Fixation            Rinse            Pre-hybridization            Hybridization</p>	<p><u>Make working solutions</u></p> <p>2xSSC<sup>3</sup>            TN<sup>3</sup>            Blocker<sup>3</sup>            TNM (thaw 2x Developer)<sup>3</sup>            1xDeveloper<sup>3</sup></p> <p><u>Experimental Technique</u></p> <p>Rinse            Antibody incubation            Rinse            Development (overnight)</p>	<p><u>Experimental Technique</u></p> <p>Rinse;            Mount with Vectashield or mounting medium of choice;            Add cover slip and seal with nail polish;            Analyze signal by light microscopy.</p> <p>Slides will keep at 4°C for several weeks, but are best within first week.</p>
<p><small>1-stable at room temperature; 2-store at -20°C in aliquots, 3-make fresh from stock components</small></p>			

**Fig. 6** Suggested timetable of steps in the procedure.

transcripts per cell. This approach should be useful for laboratories that wish to make a semiquantitative determination of gene expression at a modest cost in *C. elegans* or other related species.

## Acknowledgments

The authors thank others who have helped them test and develop the protocol and provide advice about reagents, in particular George Hsu, Andre Pires da Silva, Lenore Price, David Fitch, and Andrew Fire. The protocol was developed through funding from the NSF (grants IBN#0416922, IOS#0643325) to Morris F. Maduro.

## References

- Bobrow, M. N., and Moen, P. T., Jr. (2001). Tyramide signal amplification (TSA) systems for the enhancement of ISH signals in cytogenetics. *Curr Protoc Cytom* Chapter 8, Unit 8.9. Available at <http://onlinelibrary.wiley.com/doi/10.1002/0471142956.cy0809s11/pdf>.
- Broitman-Maduro, G., Owrighi, M., Hung, W. W., Kuntz, S., Sternberg, P. W., and Maduro, M. F. (2009). The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in *C. elegans* mesoderm development. *Development* **136**, 2735–2746.
- Coroian, C., Broitman-Maduro, G., and Maduro, M. F. (2005). Med-type GATA factors and the evolution of mesendoderm specification in nematodes. *Dev. Biol.* **289**, 444–455.

- Engelke, D. R., Krikos, A., Bruck, M. E., and Ginsburg, D. (1990). Purification of *Thermus aquaticus* DNA polymerase expressed in *Escherichia coli*. *Anal. Biochem.* **191**, 396–400.
- Giordano-Santini, R., Milstein, S., Svrzikapa, N., Tu, D., Johnsen, R., Baillie, D., Vidal, M., and Dupuy, D. (2010). An antibiotic selection marker for nematode transgenesis. *Nat. Methods* **7**, 721–723.
- Kawasaki, I., Shim, Y. H., Kirchner, J., Kaminker, J., Wood, W. B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. *Cell* **94**, 635–645.
- Krause, M., Fire, A., Harrison, S. W., Priess, J., and Weintraub, H. (1990). CeMyoD accumulation defines the body wall muscle cell fate during *C. elegans* embryogenesis. *Cell* **63**, 907–919.
- Lin, K. T., Broitman-Maduro, G., Hung, W. W., Cervantes, S., and Maduro, M. F. (2009). Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in *C. briggsae* as compared with *C. elegans*. *Dev. Biol.* **325**, 296–306.
- Maduro, M. F., Broitman-Maduro, G., Mengarelli, I., and Rothman, J. H. (2007). Maternal deployment of the embryonic SKN-1→MED-1,2 cell specification pathway in *C. elegans*. *Dev. Biol.* **301**, 590–601.
- Montgomery, M. K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *PNAS* **95**, 15502–15507.
- Nehrke, K. (2003). A reduction in intestinal cell pH<sub>i</sub> due to loss of the *Caenorhabditis elegans* Na<sup>+</sup>/H<sup>+</sup> exchanger NHX-2 increases life span. *J. Biol. Chem.* **278**, 44657–44666.
- Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A., and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* **135**, 385–404.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Raj, A., Rifkin, S. A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913–918.
- Raj, A., and Tyagi, S. (2010). Detection of individual endogenous RNA transcripts *in situ* using multiple singly labeled probes. *Methods Enzymol.* **472**, 365–386.
- Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879.
- Schlager, B., Wang, X., Braach, G., and Sommer, R. J. (2009). Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode *Pristionchus pacificus*. *Genesis* **47**, 300–304.
- Semple, J. I., Garcia-Verdugo, R., and Lehner, B. (2010). Rapid selection of transgenic *C. elegans* using antibiotic resistance. *Nat. Methods* **7**, 725–727.
- Seydoux, G., and Fire, A. (1995). Whole-mount *in situ* hybridization for the detection of RNA in *Caenorhabditis elegans* embryos. *Methods Cell Biol.* **48**, 323–337.
- Sheth, U., Pitt, J., Dennis, S., and Priess, J. R. (2010). Perinuclear P granules are the principal sites of mRNA export in adult *C. elegans* germ cells. *Development* **137**, 1305–1314.
- Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995). Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **34**, 11211–11216.
- Tabara, H., Motohashi, T., and Kohara, Y. (1996). A multi-well version of *in situ* hybridization on whole mount embryos of *Caenorhabditis elegans*. *Nucleic Acids Res.* **24**, 2119–2124.
- Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997). *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. *Genes Dev.* **11**, 2883–2896.

---

---

## CHAPTER 10

# Gene-Centered Regulatory Network Mapping

**Albertha J.M. Walhout**

Program in Gene Function and Expression and Program in Molecular Medicine, University of Massachusetts Medical School, USA

---

Abstract

- I. Introduction
  - II. Gene Regulatory Networks
  - III. Identifying Gene Regulatory Network Nodes
    - A. Regulatory Regions
    - B. Regulators
    - C. Delineating Gene Regulatory Network Edges
  - IV. Gene Regulatory Network Visualization and Analysis
    - A. Gene Regulatory Network Validation
  - V. Future Challenges
- Acknowledgments  
References

---

---

---

## Abstract

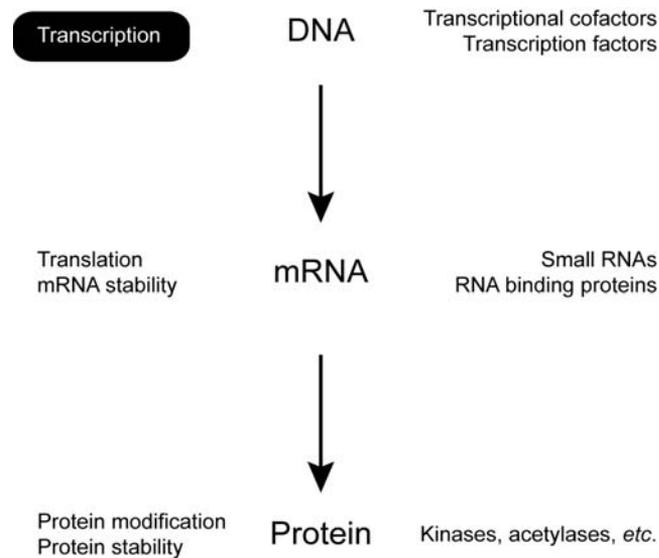
The *Caenorhabditis elegans* hermaphrodite is a complex multicellular animal that is composed of 959 somatic cells. The *C. elegans* genome contains ~20,000 protein-coding genes, 940 of which encode regulatory transcription factors (TFs). In addition, the worm genome encodes more than 100 microRNAs and many other regulatory RNA and protein molecules. Most *C. elegans* genes are subject to regulatory control, most likely by multiple regulators, and combined, this dictates the activation or repression of the gene and corresponding protein in the relevant cells and under the appropriate conditions. A major goal in *C. elegans* research is to determine the spatiotemporal expression pattern of each gene throughout development and in response to different signals, and to determine how this expression pattern is accomplished. Gene regulatory networks describe physical and/or functional interactions between genes and their regulators that result in specific spatiotemporal gene expression. Such regulators can act at transcriptional or post-transcriptional levels. Here, I

will discuss the methods that can be used to delineate gene regulatory networks in *C. elegans*. I will mostly focus on gene-centered yeast one-hybrid (Y1H) assays that are used to map interactions between non-coding genic regions, such as promoters, and regulatory TFs. The approaches discussed here are not only relevant to *C. elegans* biology, but can also be applied to other model organisms and humans.

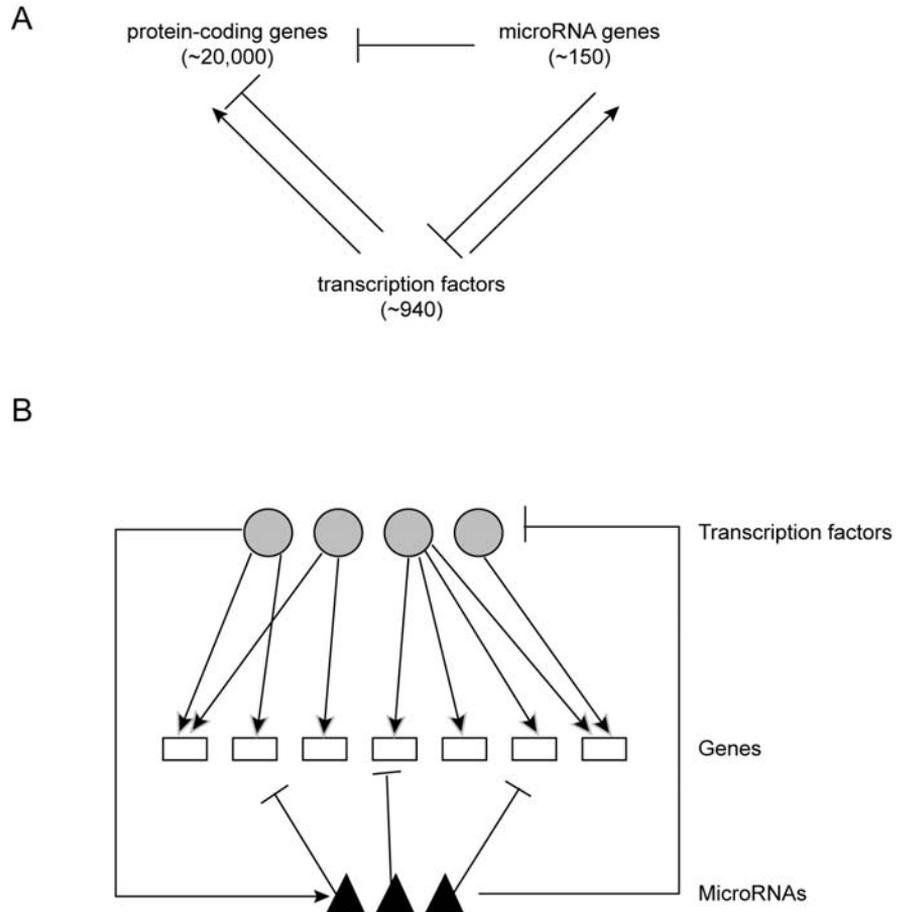
## I. Introduction

Complex multicellular model organisms such as *C. elegans* need to faithfully develop from a fertilized oocyte into a complete and fully functioning animal that is composed of different cell and tissue types. After development is completed, metazoan organisms also need mechanisms for homeostasis and to adequately respond to physiological and environmental cues, in order to find mating partners, to detect food, and to avoid pathogens. For correct functionality, cells and tissues need to compute an appropriate biological output based on the input they receive. Such an output can, for instance, be to differentiate, to move, or to enter the dauer stage. Biological outputs result from interactions between the different biomolecules cells and tissues contain, including the genome, proteins and RNA molecules as well as small molecules such as metabolites.

Developmental and post-developmental processes are controlled, at least in part, by the specific spatiotemporal expression of each of the ~20,000 protein-coding genes in the *C. elegans* genome. Each gene/protein is likely controlled by multiple regulators and at multiple levels (Figs. 1 and 2). First, genes are transcribed into



**Fig. 1** The different levels of differential gene expression. This review focuses mainly on the transcription, and to a lesser extent, on microRNAs.



**Fig. 2** A) The first levels of gene regulation of the ~20,000 protein-coding genes are controlled transcriptionally by ~940 TFs that can either activate or repress transcription and post-transcriptionally by ~150 microRNAs that repress mRNA translation and/or stability. B) Cartoon of a gene regulatory network involving genes, TFs, and microRNAs.

mRNAs, and this is controlled by the action of regulatory transcription factors (TFs) that can repress or activate gene expression by directly interacting with the genome. Second, mRNA stability and translation are controlled by small RNAs such as microRNAs, and by RNA binding proteins that frequently interact with the 3'UTR of their target mRNAs. Third, after translation, proteins can be stabilized or destabilized due to post-translational modifications by, for example, kinases or acetylases. Finally, sub-cellular mRNA and protein localization can be subject to control mechanisms as well.

The delineation of the complex networks that comprehensively describe the physical and regulatory interactions at each of these levels and between all

biomolecules is a daunting task. Here, I will focus specifically on *C. elegans* gene regulatory networks that control gene expression at the transcriptional and post-transcriptional levels. I will briefly discuss the methods that can be used to identify the players in gene regulatory networks, as well as approaches to identify interactions between them, with a primary focus on gene-centered yeast one-hybrid (Y1H) assays that are used to identify interactions between non-coding regulatory DNA regions and TFs.

---

---

---

## II. Gene Regulatory Networks

Gene regulatory networks are composed of two main components: nodes and edges. The network nodes are the players involved, that is, the genes and their regulators. The edges are the physical and/or regulatory relationships between the nodes (Fig. 2B). Gene regulatory networks are different from better-known protein–protein interaction networks, because gene regulatory networks are both bipartite and directional. They are bipartite because there are two types of nodes: genes and regulators, although of course some genes are themselves regulators of other genes or proteins. Gene regulatory networks are directional because regulators control genes and usually not the other way around. In order to map and characterize gene regulatory networks, one needs to first identify the nodes. For the genes this means to identify the non-coding genomic DNA sequences that participate in the control of gene expression, and for the regulators this means to identify which protein-coding genes encode TFs, RNA binding proteins, and other regulators, as well as to determine the complete collection of regulatory RNA molecules. Here, I will mostly focus on TFs and microRNAs, and the types of genic regions they interact with.

---

---

---

## III. Identifying Gene Regulatory Network Nodes

### A. Regulatory Regions

Different parts of a gene can contribute to its regulation. The more complex an organism, the more complex its gene regulation is. In *C. elegans* there are two main regulatory regions: gene promoters in the genome and 3'UTRs in mRNAs.

#### 1. Promoters

A gene promoter is the genomic DNA sequence immediately upstream of the transcription start site. Generally, promoters are composed of a basal element where the general transcriptional machinery binds (e.g., RNA polymerase II and general TFs), and the proximal gene promoter that serves as a landing site for regulatory TFs. Since the majority of *C. elegans* genes are subject to trans-splicing, precise transcription start sites have not been determined for most genes. However, 5'UTRs are

short compared to more complex organisms such as humans, and for practical purposes, promoters can therefore be defined as the region immediately upstream of the translational start site. It is difficult to determine the 5' start point of gene promoters. However, since most intergenic regions are shorter than 2 kb, most studies have limited their analyses to this length (Deplancke *et al.*, 2004; Dupuy *et al.*, 2004; Hunt-Newbury *et al.*, 2007). Importantly, it has been shown that this region, when fused to a reporter gene such as that encoding the green fluorescent protein (GFP) often drives gene expression in a manner that recapitulates the expression of the endogenous gene (Dupuy *et al.*, 2004; Grove *et al.*, 2009; Hunt-Newbury *et al.*, 2007; Martinez *et al.*, 2008b; Reece-Hoyes *et al.*, 2007).

To facilitate the system-level analysis of gene expression, a clone resource comprised of ~6000 *C. elegans* promoters, referred to as the Promoterome, has been generated (Dupuy *et al.*, 2004). This resource is based on the Gateway cloning system and consists of promoter Entry clones that can be easily transferred to various Destination vectors by a simple recombination reaction (Hartley *et al.*, 2000; Walhout *et al.*, 2000b). Destination vectors that can be used to analyze gene regulatory networks include a GFP vector for the creation of transgenic animals to study promoter activity *in vivo*, and Y1H vectors for the identification of TFs that can interact with the promoter (see below). So far, systematic efforts have determined the *in vivo* activity of ~350 TF-encoding gene promoters (Grove *et al.*, 2009; Reece-Hoyes *et al.*, 2007), ~1800 additional gene promoters (Hunt-Newbury *et al.*, 2007), and 73 microRNA gene promoters (Martinez *et al.*, 2008b). Many of the corresponding transgenic lines are available to the community through the *C. elegans* genetics center (CGC).

## 2. 3' UTRs

The 3'UTR is the untranslated region in the mRNA, immediately downstream of the stop codon. This region is subject to post-transcriptional control by microRNAs and RNA binding proteins. Recently, a comprehensive collection of 3'UTRs has been delineated for most *C. elegans* genes (Mangone *et al.*, 2010). Cloning these 3'UTRs into Gateway-compatible vectors will provide a resource for experimental gene regulatory network mapping that is similar to the ORFeome (see below) and Promoterome resources.

## 3. Other Genic Regulatory Regions

It is not clear to what extent other regulatory regions function in gene regulatory networks in *C. elegans*. So far, transcriptional studies have mostly focused on promoters. However, it is clear that other regions, such as introns and sequences downstream of the gene, can also play a role. Similarly, microRNAs and RNA binding proteins could target regions outside 3'UTRs within their mRNA targets. Systematic studies are required to elucidate the relative role different genic regions play in complex gene regulatory networks.

## B. Regulators

### 1. Transcription Factors

TFs provide the first level of gene control. They bind directly to DNA through their sequence-specific DNA binding domain and can be grouped into families based on the type of DNA binding domain they possess (Reece-Hoyes *et al.*, 2005). Well-known DNA binding domains include the homeodomain, the basic helix-loop-helix (bHLH) domain, C2H2 zinc fingers, the ETS domain, the bZIP domain, and C4-type zinc fingers found in nuclear hormone receptors (NHRs). TFs can be predicted in a genome of interest by searching the complete collection of proteins for the presence of a known DNA binding domain. This is usually done by computational methods, for instance using Interpro (Mulder *et al.*, 2003) or SMART (Letunic *et al.*, 2004) databases. However, we have found that visual inspection of predicted DNA binding domains using knowledge of their sequence and structure is highly useful as well. Indeed, by doing so we increased the predicted set of *C. elegans* TFs from ~600 (Ruvkun and Hobert, 1998) to 940, or ~5% of all protein-coding genes (Reece-Hoyes *et al.*, 2005; Vermeirssen *et al.*, 2007b). Most *C. elegans* TF-encoding genes encode a single splice variant; however in some cases multiple variants are present, and some of these may encode proteins with different DNA binding domains (Reece-Hoyes *et al.*, 2005). Interestingly, different TF variants can have different biological functions. For instance, different variants of the forkhead protein DAF-16 were recently found to be expressed in distinct patterns and to confer different functions related to metabolism and aging (Kwon *et al.*, 2010). Several proteins have been identified that can bind *C. elegans* gene promoters but that do not possess a known DNA binding domain (Deplancke *et al.*, 2006a; Vermeirssen *et al.*, 2007a). Thus, the total collection of *C. elegans* TFs may be slightly larger, but is likely not to exceed 1000 (unpublished data).

More than 12,000 *C. elegans* full-length open reading frames (ORFs) have been cloned into a Gateway-compatible resource called the ORFeome (Lamesch *et al.*, 2004; Reboul *et al.*, 2003). We obtained the TF-encoding ORFs from this resource and supplemented that with TF-encoding ORFs that we cloned *ab initio* (Deplancke *et al.*, 2004; Vermeirssen *et al.*, 2007b). The resulting clone collection currently contains ~90% of all full-length TFs and can be directly used in assays for the delineation of gene regulatory networks such as Y1H assays (unpublished data, see below).

### 2. MicroRNAs

MicroRNAs regulate gene expression post-transcriptionally by sequence-specific but imperfect basepairing with the 3'UTR of their target mRNAs. It has been estimated that the *C. elegans* genome encodes more than 110 microRNAs (Lehrbach and Miska, 2008). Some of these have been identified genetically (e.g., *lin-4*, *let-7*), some have been predicted computationally (Lim *et al.*, 2003), and others were more recently found by deep sequencing small RNA populations purified from

worms (Friedlander *et al.*, 2008; Kato *et al.*, 2009). As with TFs, microRNAs can also be grouped into families, based on their seed sequence, the part with which they basepair with their target genes. It is not yet clear whether all *C. elegans* microRNAs have been identified. Indeed, it may be that additional microRNAs will be uncovered when the animal is exposed to particular conditions, in males or dauers, or when sequencing techniques further improve to detect microRNAs of very low abundance.

### 3. Other Regulators

In addition to TFs and microRNAs, other RNA and protein molecules contribute to differential gene regulation. These include RNA binding proteins, transcriptional co-factors, and signaling molecules such as kinases and phosphatases, as well as endogenous siRNAs and, perhaps, long non-coding RNAs. Systematic computational and experimental analyses will shed light on the number of molecules in each class of regulators.

## C. Delineating Gene Regulatory Network Edges

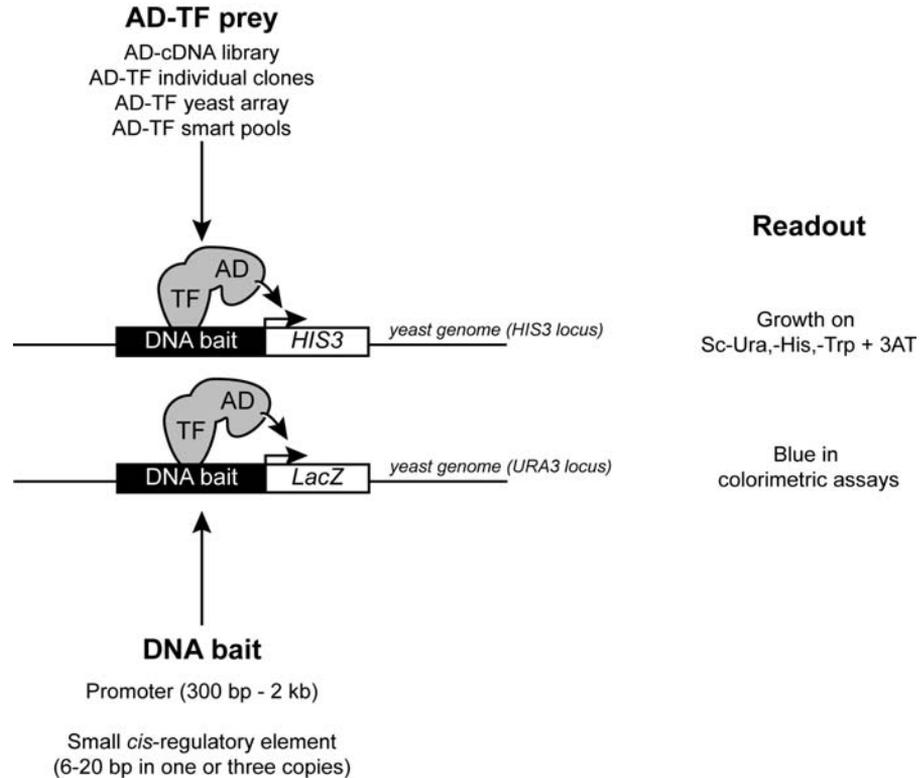
### 1. TF-Target Gene Interactions

Interactions between TFs and their target genes can be identified using two conceptually different and highly complementary strategies. The first are TF-centered (protein-to-DNA); they start with a TF of interest and identify the genes with which this factor interacts. The second are gene-centered (DNA-to-protein); they start with a gene of interest and identify the TFs with which it interacts (Fig. 3).

### 2. Transcription Factor-Centered Methods: ChIP

The most widely used TF-centered method is chromatin immunoprecipitation (ChIP). In ChIP assays, an anti-TF antibody is used to precipitate TFs *in vivo*. Briefly, worm extracts are first treated with formaldehyde to crosslink proteins to proteins and proteins to DNA. After precipitation of the TF, associated DNA molecules can be identified 1) by PCR using primer sets of interest (Deplancke *et al.*, 2006a); 2) by cloning and sequencing (Oh *et al.*, 2006); 3) using microarrays that tile the entire *C. elegans* genome (Tabuchi *et al.*, 2011; Whittle *et al.*, 2009); or more recently 4) by deep sequencing (e.g., 454 or Solexa). Controls include a non-relevant antibody and, if possible, mutant animals that do not express the TF of interest (Walhout, 2011).

ChIP is a powerful method to identify TF-target gene interactions that occur *in vivo*. However, it is mostly limited to TFs that are highly and/or broadly expressed throughout the lifetime of the animal, and to TFs for which ChIP-grade antibodies are available. It is, however, also feasible to use ChIP in transgenic animals that overexpress an epitope-tagged TF. Although ChIP is usually the method of choice when one is interested in one or a few TFs, it is less suitable when one is interested in



**Fig. 3** Both TF-centered and gene-centered methods can be used for the identification of TF-target gene interactions.

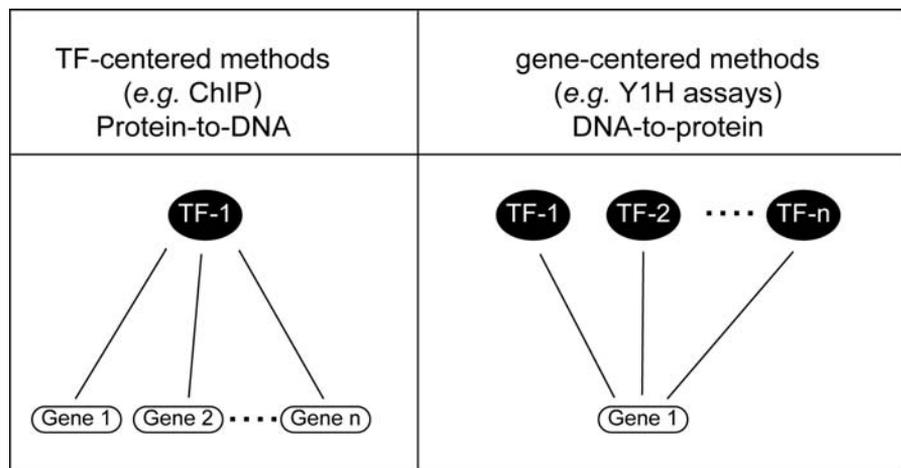
a single gene (or a set of genes) and wants to identify the TFs that contribute to its (their) regulation. This is because all 940 TFs would have to be tested and under all relevant developmental and physiological conditions. Detailed discussion and protocols for ChIP in worms are provided elsewhere (Mukhopadhyay *et al.*, 2008).

### 3. Gene-Centered Methods: Y1H Assays

Y1H assays provide a genetic method for the gene-centered identification of TF-target gene interactions. The Y1H system is conceptually similar to yeast two-hybrid (Y2H) assays that have been used extensively to map *C. elegans* protein-protein interaction networks (Li *et al.*, 2004; Walhout *et al.*, 2000a, 2002). Here, I will discuss the principles of the Y1H system. Detailed Y1H protocols are available elsewhere (Deplancke *et al.*, 2006b).

The Y1H system uses a reporter gene readout in yeast to detect interactions between a “DNA bait” and a “protein prey” (e.g., TF) (Fig. 4). The first step in Y1H assays involves the selection of the DNA bait. In most of the cases, this will be a gene promoter or a small *cis*-regulatory element. Next, the DNA bait is cloned upstream of two reporter genes, *HIS3* and *LacZ* (Fig. 4). Traditionally, this was done by restriction enzyme/ligation-based methods (Li and Herskowitz, 1993). However, this is difficult to standardize and thus not amenable to the high-throughput settings that are required for regulatory network studies. To enable high-throughput cloning of DNA baits, we have combined the Y1H system with Gateway cloning, a recombination-based method that is compatible with the Promoterome resource (Deplancke *et al.*, 2004). With this method, multiple DNA baits can be transferred to the Y1H reporter Destination vectors simultaneously (e.g., in 96-well plates).

After cloning, the two DNA bait::reporter constructs are linearized and integrated into the genome of a suitable yeast strain. DNA bait::*HIS3* constructs are integrated into a mutant *HIS3* locus and plated on media lacking histidine. There is enough background His3 expression conferred by the basal yeast promoter present in the DNA bait::reporter constructs to enable growth on media lacking histidine. When the same construct is used in a protein–DNA interaction assay, however, the media are supplemented with 3-aminotriazole (3AT), a competitive inhibitor of the His3 enzyme. That way, the growth of the yeast depends on an increase in expression of His3, conferred by an interacting AD-TF hybrid protein (see below and Fig. 4). DNA bait::*LacZ* constructs contain a wild-type *URA3* gene and are integrated into a mutant *URA3* locus, thereby rescuing the Ura3 deficiency when plated on media lacking uracil. The DNA bait::reporter constructs do not carry a yeast origin of replication and, therefore, the formation of colonies is strictly dependent on their



**Fig. 4** Cartoon of yeast one-hybrid assays. AD – Gal4 transcription activation domain; TF – transcription factor.

integration into the yeast genome. Integrations are generally done sequentially, either by first integrating the DNA bait::*HIS3* or the DNA bait::*LacZ* construct, and following with the other. However, it is possible to integrate both constructs simultaneously, but the efficiency will be much lower and only a handful of colonies is usually obtained (unpublished data).

After picking integrant colonies, they need to be tested for background reporter gene expression (auto-activation). Levels of auto-activation can differ between integrants from the same DNA bait::reporter construct, most likely because of differences in copy number (Deplancke *et al.*, 2004). The degree of auto-activation of DNA bait::*HIS3* strains is determined by plating the colonies on media lacking histidine, and with increasing concentrations of 3AT (5, 10, 20, 40, 60, and 80 mM). Preferably colonies are selected that do not confer growth on low concentrations (5–40 mM) of 3AT. The degree of auto-activation of DNA bait::*LacZ* strains is determined by a colorimetric assay where white indicates no expression and darker shades of blue indicate increasing induction of  $\beta$ Galactosidase. Colonies with little or no blue should be selected where possible. In our hands 10–20% of all DNA baits exhibit high levels of auto-activation. These baits are difficult to use in Y1H assays although interacting TFs can sometimes be detected, particularly in directed Y1H assays (Vermeirssen *et al.*, 2007b).

After obtaining double integrant DNA bait strains that exhibit the lowest possible levels of auto-activation, the actual Y1H experiment can be performed to detect interacting TFs. In Y1H assays, TFs are fused to the transcription activation domain (AD) of the yeast Gal4 protein. This ensures that both activators and repressors of transcription can be detected. In other words, only physical protein–DNA interactions are examined in Y1H assays. AD-TF clones can be obtained from different sources and can be introduced into the DNA bait strain in different ways (Fig. 4). In our Y1H system, AD-TF clones carry wild-type yeast *TRP1* gene and, therefore, colonies containing the plasmid are selected on media lacking tryptophan.

The most commonly used method is by transforming an AD-cDNA library into haploid DNA bait strains (Arda *et al.*, 2010; Deplancke *et al.*, 2004, 2006a, 2006b; Martinez *et al.*, 2008a; Vermeirssen *et al.*, 2007a). Another source for such haploid transformations was created by cherry-picking relevant clones from the ORFeome, transferring them to the AD Y1H Destination vector by Gateway cloning, and combining them into a single AD-TF mini library (Deplancke *et al.*, 2004). This library consists of ~650 full-length TFs. Screening such a mini library enables the detection of TFs that are underrepresented in non-normalized cDNA libraries. Since TFs are often of low abundance, this can be very useful. In library screens, interacting TFs are identified by yeast colony PCR and sequencing. We have also developed mini pools of individual AD-TF clones that can be introduced into DNA bait strains by transformation (Vermeirssen *et al.*, 2007b). These pools are designed using a “Smart pool” strategy, based on a Steiner Triple System that is used in combinatorial mathematics. We have generated these pools as well as the scripts to deconvolute the resulting interactions. This method is useful for higher throughput, cost-effective Y1H experiments because it does not rely on extensive prey sequencing. Single AD-TF

clones can of course also be transformed individually when particular pre-defined interactions are to be examined (Reece-Hoyes *et al.*, 2009).

In addition to transformation into haploid DNA bait strains, AD-TF clones can also be introduced by mating. For this, we have transformed the AD-TF clones (~755 in the first iteration) into yeast of mating type  $\alpha$ , which is compatible for mating with the DNA bait strains that have the “a” mating type (Vermeirssen *et al.*, 2007b). DNA bait strains are mated with the AD-TF clone array and positives are examined in diploids. Each of these different methods for introducing AD-TFs into DNA bait strains has advantages and disadvantages (Vermeirssen *et al.*, 2007b). Generally, transformation detects more interacting TFs than mating. However, mating is fast, less labor-intensive, and much less costly. Further, interactions detected by mating are highly reproducible. When comparing library screens to more directed experiments with smart pools or individual clones, it is clear that many more protein–DNA interactions are found by the latter methods. However, with directed experiments only cloned TFs can by definition be found, which in our current collection is about 850 (~90%) (Vermeirssen *et al.*, 2007b) (unpublished data). Proteins that do not have a recognizable DNA binding domain can only be retrieved in unbiased cDNA library screens (Deplancke *et al.*, 2006a; Vermeirssen *et al.*, 2007a). However, we do include these in TF resources after confirming their capability of interacting with *C. elegans* promoters and obtaining a suitable clone.

#### 4. MicroRNA–mRNA Interactions

Putative interactions between the 3'UTRs of mRNAs and microRNAs are mostly identified genetically or computationally predicted using one or more algorithms that are publicly available. These include PicTar (Lall *et al.*, 2006), MiRanda (Griffiths-Jones *et al.*, 2006), TargetScan (Lewis *et al.*, 2005), RNA hybrid (Rehmsmeier *et al.*, 2004), and mirWIP (Hammell *et al.*, 2008). These algorithms are challenging to use because they are often too greedy (high rate of false positive predictions), or too stringent (high rate of false negative predictions). In order to alleviate this, at least to some extent, we have previously used predictions that were found by at least two of the four algorithms used (Martinez *et al.*, 2008a). Future experimental approaches will shed light onto physical and functional microRNA–mRNA interactions that occur *in vivo* (Lall *et al.*, 2006; Zisoulis *et al.*, 2010).

#### 5. Other Regulatory Interactions

In addition to protein–DNA and microRNA–mRNA interactions, other relationships are involved in gene control. An important class involves sequence-specific RNA binding proteins that interact with the 3'UTR of mRNAs. It is not yet clear how many sequence-specific RNA binding proteins are encoded by the *C. elegans* genome, and only few have been studied genetically or biochemically. For instance, detailed binding sites have been determined *in vitro* for MEX-3, MEX-5, and a handful of other RNA binding proteins (Farley *et al.*, 2008; Pagano *et al.*, 2007,

2009). However, the functionality of most RNA binding proteins and their mRNA targets remains largely unexplored.

---

---

---

## IV. Gene Regulatory Network Visualization and Analysis

The identification of physical and functional relationships between genes and their regulators is only the first step in the characterization of gene regulatory networks. Lists of interactions are usually difficult to navigate through. Network models, however, provide a visually attractive method for gene regulatory network analysis. We usually use the publicly available Cytoscape tool (Shannon *et al.*, 2003) for network visualization and analysis (Arda *et al.*, 2010; Deplancke *et al.*, 2006a; Grove *et al.*, 2009; Martinez *et al.*, 2008a; Vermeirssen *et al.*, 2007a). Subsequently, we use a variety of tools for network analysis. Most notably we use topological overlap coefficient analysis to compare gene expression patterns and to identify TF or gene network modules. These methods are discussed elsewhere (Arda *et al.*, 2010; Arda and Walhout, 2009; Ravasz *et al.*, 2002; Vermeirssen *et al.*, 2007a).

### A. Gene Regulatory Network Validation

As with any method, the identification of physical and functional interactions between genes and their regulators is subject to issues related to both assay sensitivity and assay specificity (Walhout, 2011). Sensitivity refers to the proportion of real interaction that can be identified by the assay; interactions that cannot be detected are referred to as false negatives. Specificity refers to the proportion of interactions detected that are real, that is, that do occur *in vivo* and/or that have a biological consequence. Interactions that are detected but that are not “biologically meaningful” are referred to as false positives.

#### 1. False Negatives

Previously, we estimated the coverage of our Y1H screens to be ~35% (Deplancke *et al.*, 2006a). This number is based on a very small number of available published interactions, but is very similar to the coverage obtained with Y2H (Braun *et al.*, 2009). There are several reasons that not all possible TF-promoter interactions can be detected by Y1H assays: 1) Several TFs bind DNA as obligatory dimers. Although homodimers can be detected, the Y1H assay currently is not configured to detect heterodimers. In the future, we hope to develop approaches that enable the detection of heterodimeric TF-DNA interactions in directed Y1H assays. 2) We will not find TFs that depend on specific post-translational modification or co-factor interactions with *C. elegans* proteins for DNA binding. 3) We can obviously only find TFs that are available in the TF resource used. However, it is highly encouraging to note that we have already detected interactions for about 25% of all predicted *C. elegans* TFs with only ~1% of all gene promoters.

## 2. False Positives

As with Y2H assays, there are two types of false positives with Y1H assays: *technical false positives* that cannot be reproduced in the same assay and *biological false positives* that represent genuine Y1H/Y2H interactions that nonetheless do not occur *in vivo*. To keep the rate of technical false positives low several issues need to be taken into consideration. First, it is best to only consider interactions that score positively for both Y1H reporters, that is, that induce growth on media lacking histidine and containing 3AT and that are bluer than an “AD only” control. Second, it is important to make sure that the TF retrieved is in frame (only relevant to cDNA library screens). Third, all Y1H interactions need to be retested in fresh DNA bait cells (i.e., from a frozen stock that has not been used in the screen itself), either by gap-repair (Walhout and Vidal, 2001) or by directly transforming an AD-TF clone. This is necessary because baits can mutate in yeast and give rise to a colony with an apparent interaction phenotype that is not reproducible (Walhout and Vidal, 1999). Fourth, it is absolutely critical to integrate DNA bait::reporter constructs into the yeast genome. We have tried to perform the assay with replicating plasmids, but the background expression was highly variable, probably due to different plasmid copy numbers. Finally, it is important to note that interactions obtained with highly auto-active DNA baits are more difficult to assess and may be less specific. We have developed an interaction scoring scheme to assess the results obtained from Y1H library screens (Vermeirssen *et al.*, 2007a).

Biological false positives are more challenging to assess. First, the genome itself is the same in every cell and thus, when a TF is expressed in any given cell one may expect the interaction to occur. However, the nucleosome occupancy likely varies in different cell types and this may prevent interactions from occurring *in vivo*. The integration of the DNA baits into the yeast genome ensures that they are incorporated into chromatin and, thus, Y1H assays are not based on interactions with naked DNA. However, it could be that the integration of the DNA baits in yeast only partially recapitulates the chromatin state in any *C. elegans* cell *in vivo*. In Y1H assays, we can find multiple members of a TF family binding to a particular DNA bait. This could be because these members have very similar DNA binding specificities and that this does not reflect *in vivo* functionality. However, we, and others, have found that multiple members of a TF family can bind the same DNA targets *in vivo* and can function redundantly (Hollenhorst *et al.*, 2007; Ow *et al.*, 2008). For instance, multiple TFs with a FLYWCH DNA binding domain were found to interact with microRNA promoters in Y1H assays and to redundantly repress microRNA expression in the early *C. elegans* embryo (Ow *et al.*, 2008). It is also important to note that not all TF-DNA interactions lead to a regulatory consequence. For instance, ChIP has identified numerous interactions that do not have an apparent biological function (Li *et al.*, 2008). This should be taken into account when physical interactions are being assessed by regulatory assays such as target gene expression in TF mutants or by TF knockdown with RNAi. Finally, different validation assays each have their own rate of false negatives, that is, they cannot detect every single genuine

interaction. For instance, assays that are performed with mixed populations of animals can easily miss interactions that occur only in a few cells or only during a short developmental time. Indeed, in our study of the B0507.1 promoter, we found a reduction in expression upon loss of the TF CES-1 only in the spermatheca, rectal gland, and pharyngeal-intestinal valve and, since these are not large tissues, this would be extremely difficult to detect in mixed population whole animal assays such as qPCR (Reece-Hoyes *et al.*, 2009).

---

---

---

## V. Future Challenges

The comprehensive mapping of gene regulatory networks in *C. elegans* has only just started. Future studies are needed to complete transcriptional networks by high-throughput Y1H assays, and by other complementary assays such as ChIP. In addition, it will be highly useful to systematically generate promoter::GFP constructs and corresponding transgenic *C. elegans* lines for all worm genes. Such lines can then be used to examine promoter activity under different experimental or physiological conditions and to validate transcriptional networks, for instance using TF mutants or TF knockdown. Further, the continued experimental analysis of microRNAs and other small RNAs will be of extremely high value. Experimental methods also need to be developed and applied to assess other regulatory networks, such as those involving RNA binding proteins, signaling molecules, and metabolites. Finally, it will be exciting to go beyond static network models that represent a compilation of the interactions that can occur in the animal and to incorporate the dynamics and levels of gene and regulator expression and activation throughout the lifetime of the nematode.

## Acknowledgments

I thank members of my lab for their hard work and especially John Reece-Hoyes and Lesley MacNeil for critical reading of the manuscript. Work in my lab is supported by the National Institutes of Health (DK068429 and GM082971) and by the Ellison Medical Research Foundation.

## References

- Arda, H. E., Taubert, S., Conine, C., Tsuda, B., Van Gilst, M. R., Sequerra, R., Doucette-Stam, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a *C. elegans* gene regulatory network. *Mol. Syst. Biol.* **6**, 367.
- Arda, H. E., and Walhout, A. J. M. (2009). Gene-centered regulatory networks. *Briefings Funct. Genomic. Proteomic* doi:10.1093/elp049 **9**, 4–12.
- Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J. M., Murray, R. R., Roncari, L., de Smet, A. S., Venkatesan, K., Rual, J. F., Vandenhaute, J., Cusick, M. E., Pawson, T., Hill,

- D. E., Tavernier, J., Wrana, J. L., Roth, F. P., and Vidal, M. (2009). An experimentally derived confidence score for binary protein-protein interactions. *Nat. Methods* **6**, 91–97.
- Deplancke, B., Dupuy, D., Vidal, M., and Walhout, A. J. M. (2004). A gateway-compatible yeast one-hybrid system. *Genome Res.* **14**, 2093–2101.
- Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A. M., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stam, L., Reece-Hoyes, J. S., Hope, I. A., Tissenbaum, H. A., Mango, S. E., and Walhout, A. J. M. (2006a). A gene-centered *C. elegans* protein-DNA interaction network. *Cell* **125**, 1193–1205.
- Deplancke, B., Vermeirssen, V., Arda, H. E., Martinez, N. J., and Walhout, A. J. M. (2006b). Gateway-compatible yeast one-hybrid screens. *CSH Protocols* doi:10.1101/pdb.prot4590 .
- Dupuy, D., Li, Q., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stam, L., Hope, I. A., Hill, D., Walhout, A. J. M., and Vidal, M. (2004). A first version of the *Caenorhabditis elegans* promoterome. *Genome Res.* **14**, 2169–2175.
- Farley, B. M., Pagano, J. M., and Ryder, S. P. (2008). RNA target specificity of the embryonic cell fate determinant POS-1. *RNA* **14**, 2685–2697.
- Friedlander, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., and Rajewski, N. (2008). Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **26**, 407–415.
- Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34**, D140–D144.
- Grove, C. A., deMasi, F., Barrasa, M. I., Newburger, D., Alkema, M. J., Bulyk, M. L., and Walhout, A. J. (2009). A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. *Cell* **138**, 314–327.
- Hammell, M., Long, D., Zhang, L., Lee, A., Carmack, C. S., Han, M., Ding, Y., and Ambros, V. (2008). mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat. Methods* **5**, 813–819.
- Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000). DNA cloning using *in vitro* site-specific recombination. *Genome Res.* **10**, 1788–1795.
- Hollenhorst, P. C., Shah, A. A., Hopkins, C., and Graves, B. J. (2007). Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the *ETS* gene family. *Genes Dev.* **21**, 1882–1894.
- Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H. M., Pan, J., Schulz, A. K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel, R., Jones, S. J., Marra, M. A., Baillie, D. L., and Moerman, D. G. (2007). High-throughput *in vivo* analysis of gene expression in *Caenorhabditis elegans*. *PLoS Biol.* **5**, e237.
- Kato, M., de Lencastre, A., Pincus, Z., and Slack, F. J. (2009). Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. *Genome Biol.* **10**, R54.
- Kwon, E. S., Narasimhan, S. D., Yen, K., and Tissenbaum, H. A. (2010). A new DAF-16 isoform regulates longevity. *Nature* **466**, 498–502.
- Lall, S., Grun, D., Krek, A., Chen, K., Wang, Y. -L., Dewey, C. N., Sood, P., Colombo, T., Bray, N., MacMenamin, P., Kao, H. -L., Gunsalus, K. C., Pachter, L., Piano, F., and Rajewski, N. (2006). A genome-wide map of conserved microRNA targets in *C. elegans*. *Curr. Biol.* **16**, 460–471.
- Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R., Bosak, S., Doucette-Stam, L., Vandenhaute, J., Hill, D. E., and Vidal, M. (2004). *C. elegans* ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res.* **14**, 2064–2069.
- Lehrbach, N. J., and Miska, E. A. (2008). Functional genomic, computational and proteomic analysis of *C. elegans* microRNAs. *Brief. Funct. Genomic. Proteomic.* **7**, 228–235.

- Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P., and Bork, P. (2004). SMART 4.0: towards genomic data integration. *Nucleic Acids Res.* **32**, D142–D144.
- Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20.
- Li, J. J., and Herskowitz, I. (1993). Isolation of the ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* **262**, 1870–1874.
- Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. -O., Han, J. -D. J., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. -F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Bauemgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., van den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stam, L., Gunsalus, K., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D., and Vidal, M. (2004). A map of the interactome network of the metazoan *C. elegans*. *Science* **303**, 540–543.
- Li, X. Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D. A., Iyer, V. N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C. L., Chu, H. C., Ogawa, N., Inwood, W., Sementchenko, V., Beaton, A., Weiszmann, R., Celniker, S. E., Knowles, D. W., Gingeras, T. R., Speed, T. P., Eisen, M. B., and Biggin, M. D. (2008). Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**, e27.
- Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., and Bartel, D. P. (2003). The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991–1008.
- Mangone, M., Macmenamin, P., Zegar, C., Piano, F., and Gunsalus, K. C. (2008). UTRome.org: a platform for 3'UTR biology in *C. elegans*. *Nucleic Acids Res.* **36**, D57–D62.
- Mangone, M., Manoharan, A. P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S., Mis, E., Zegar, C., Gutwein, M. R., Khivansara, V., Attie, O., Chen, K., Salehi-Ashtiani, K., Vidal, M., Harkins, T. T., Bouffard, P., Suzuki, Y., Sugano, S., Kohara, Y., Rajewsky, N., Piano, F., Gunsalus, K. C., and Kim, J. K. (2010). The landscape of *C. elegans* 3'UTRs. *Science* **329**, 432–435.
- Martinez, N. J., Ow, M. C., Barrasa, M. I., Hammell, M., Sequerra, R., Doucette-Stamm, L., Roth, F. P., Ambros, V., and Walhout, A. J. M. (2008a). A *C. elegans* genome-scale microRNA network contains composite feedback motifs with high flux capacity. *Genes Dev.* **22**, 2535–2549.
- Martinez, N. J., Ow, M. C., Reece-Hoyes, J., Ambros, V., and Walhout, A. J. (2008b). Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity. *Genome Res.* **18**, 2005–2015.
- Mukhopadhyay, A., Deplancke, B., Walhout, A. J., and Tissenbaum, H. A. (2008). Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in *Caenorhabditis elegans*. *Nat. Protoc.* **3**, 698–709.
- Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Barrell, D., Bateman, A., Binns, D., Biswas, M., Bradley, P., Bork, P., Bucher, P., Copley, R. R., Courcelle, E., Das, U., Durbin, R., Falquet, L., Fleischmann, W., Griffiths-Jones, S., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lopez, R., Letunic, I., Lonsdale, D., Silventoinen, V., Orchard, S. E., Pagni, M., Peyruc, D., Ponting, C. P., Selengut, J. D., Servant, F., Sigrist, C. J., Vaughan, R., and Zdobnov, E. M. (2003). The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res.* **31**, 315–318.
- Oh, S. W., Mukhopadhyay, A., Dixit, B. L., Raha, T., Green, M. R., and Tissenbaum, H. A. (2006). Identification of direct targets of DAF-16 controlling longevity, metabolism and diapause by chromatin immunoprecipitation. *Nat. Genet.* **38**, 251–257.
- Ow, M. C., Martinez, N. J., Olsen, P., Silverman, S., Barrasa, M. I., Conradt, B., Walhout, A. J. M., and Ambros, V. R. (2008). The FLYWCH transcription factors FLH-1, FLH-2 and FLH-3 repress embryonic expression of microRNA genes in *C. elegans*. *Genes Dev.* **22**, 2520–2534.
- Pagano, J. M., Farley, B. M., Essien, K. I., and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proc. Natl Acad. Sci. U. S. A.* **106**, 20252–20257.

- Pagano, J. M., Farley, B. M., McCoig, L. M., and Ryder, S. P. (2007). Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J. Biol. Chem.* **282**, 8883–8894.
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabasi, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555.
- Reboul, J., Vaglio, P., Rual, J. F., Lamesch, P., Martinez, M., Armstrong, C. M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson Jr., J. R., Hartley, J. L., Brasch, M. A., Vandenhaute, J., Boulton, S., Endress, G. A., Jenna, S., Chevet, E., Papatotiropoulos, V., Tolia, P. P., Ptacek, J., Snyder, M., Huang, R., Chance, M. R., Lee, H., Doucette-Stamm, L., Hill, D. E., and Vidal, M. (2003). *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* **34**, 35–41.
- Reece-Hoyes, J. S., Deplancke, B., Barrasa, M. I., Hatzold, J., Smit, R. B., Arda, H. E., Pope, P. A., Gaudet, J., Conradt, B., and Walhout, A. J. (2009). The *C. elegans* Snail homolog CES-1 can activate gene expression *in vivo* and share targets with bHLH transcription factors. *Nucleic Acids Res.* **37**, 3689–3698.
- Reece-Hoyes, J. S., Deplancke, B., Shingles, J., Grove, C. A., Hope, I. A., and Walhout, A. J. M. (2005). A compendium of *C. elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome Biol.* **6**, R110.
- Reece-Hoyes, J. S., Shingles, J., Dupuy, D., Grove, C. A., Walhout, A. J., Vidal, M., and Hope, I. A. (2007). Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics* **8**, 27.
- Rehmsmeier, M., Steffen, P., Hochsmann, M., and Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507–1517.
- Ruvkun, G., and Hobert, O. (1998). The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* **282**, 2033–2041.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504.
- Tabuchi, T., Deplancke, B., Osato, N., Zhu, L. J., Barrasa, M. I., Harrison, M. M., Horvitz, H. R., Walhout, A. J. M., and Hagstrom, K. (2011). Chromosome-biased binding and gene regulation by the *Caenorhabditis elegans* DRM complex. *PLoS Genet.* **7**, e1002074.
- Vermeirssen, V., Barrasa, M. I., Hidalgo, C., Babon, J. A. B., Sequerra, R., Doucette-Stam, L., Barabasi, A. L., and Walhout, A. J. M. (2007a). Transcription factor modularity in a gene-centered *C. elegans* core neuronal protein-DNA interaction network. *Genome Res.* **17**, 1061–1071.
- Vermeirssen, V., Deplancke, B., Barrasa, M. I., Reece-Hoyes, J. S., Arda, H. E., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Brent, M., and Walhout, A. J. M. (2007b). Matrix and Steiner-triple-system smart pooling assays for high-performance transcription regulatory network mapping. *Nat. Methods* **4**, 659–664.
- Walhout, A. J. M. (2011). What does biologically meaningful mean?. A perspective on gene regulatory network validation. *Genome Biol.* **12**, 109.
- Walhout, A. J. M., Reboul, J., Shtanko, O., Bertin, N., Vaglio, P., Ge, H., Lee, H., Doucette-Stam, L., Gunsalus, K. C., Schetter, A. J., Morton, D. G., Kemphues, K. J., Reinke, V., Kim, S. K., Piano, F., and Vidal, M. (2002). Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline. *Curr. Biol.* **12**, 1952–1958.
- Walhout, A. J. M., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., and Vidal, M. (2000a). Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science* **287**, 116–122.
- Walhout, A. J. M., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000b). GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* **328**, 575–592.
- Walhout, A. J. M., and Vidal, M. (1999). A genetic strategy to eliminate self-activator baits prior to high-throughput yeast two-hybrid screens. *Genome Res.* **9**, 1128–1134.
- Walhout, A. J. M., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. *Methods* **24**, 297–306.

- Whittle, C. M., Lazakovitch, E., Gronostajski, R. M., and Lieb, J. D. (2009). DNA-binding specificity and *in vivo* targets of *Caenorhabditis elegans* nuclear factor I. *Proc. Natl Acad. Sci. U. S. A.* **106**, 12049–12054.
- Zisoulis, D. G., Lovci, M. T., Wilbert, M. L., Hutt, K. R., Liang, T. Y., Pasquinelli, A. E., and Yeo, G. W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **17**, 173–179.

---

---

## CHAPTER 11

# Affinity Purification of Protein Complexes in *C. elegans*

**Esther Zanin<sup>\*</sup>, Julien Dumont<sup>†</sup>, Reto Gassmann<sup>\*</sup>,  
Iain Cheeseman<sup>§</sup>, Paul Maddox<sup>‡</sup>, Shirin Bahmanyar<sup>\*</sup>,  
Ana Carvalho<sup>\*</sup>, Sherry Niessen<sup>¶</sup>, John R. Yates III<sup>¶</sup>,  
Karen Oegema<sup>\*</sup> and Arshad Desai<sup>\*</sup>**

<sup>\*</sup>Ludwig Institute for Cancer Research and Department of Cellular & Molecular Medicine, University of California San Diego, La Jolla, California, USA

<sup>†</sup>Institut Curie-UMR144, Paris, France

<sup>‡</sup>Institute for Research in Immunology and Cancer, Department of Pathology and Cell Biology, University of Montreal, Montreal, Quebec, Canada

<sup>§</sup>Whitehead Institute for Biomedical Research, MIT, Cambridge, MA, USA

<sup>¶</sup>The Skaggs Institute for Chemical Biology and Department of Chemical Physiology, The Center for Physiological Proteomics, The Scripps Research Institute, La Jolla, California, USA

---

### Abstract

- I. Introduction
- II. Generating Tools for Biochemistry
  - A. Generating a Polyclonal Antibody
  - B. Assessing Antibody Specificity by Immunoblotting after RNAi
  - C. Tandem Affinity Purification Tags
  - D. Introduction of Transgenes for Expression of Tagged Proteins
  - E. Validating the Functionality of the Transgene-Encoded Tagged Protein
- III. The Optimal Starting Material for Protein Purification
  - A. Growing Worms in Large-Scale Liquid Culture
  - B. Unsynchronized Liquid Starter Culture
  - C. Monitoring Worm Cultures
  - D. Harvesting Worms from Liquid Cultures
  - E. Isolation of Embryos and Synchronization as Starved L1 Larvae
  - F. Seeding Synchronized Cultures Using Starved L1 Larvae
  - G. Freezing Adult Worms for Immunoprecipitation
  - H. Freezing Embryos for Immunoprecipitation

- I. Enriching for Specific Age Embryos
- J. Enriching for Meiosis I Arrested Embryos
- IV. Isolation of Nuclei from Embryos
- V. Preparing Worm and Embryonic Extract
- VI. Single-step Immunoprecipitation
  - A. Covalent Coupling of Antibodies to Protein A Beads
  - B. Immunoprecipitation
  - C. Sample Buffer Elution: For Silver Staining & Immunoblotting
  - D. Glycine Elution: For Mass Spectrometry
  - E. Urea Elution: For Mass Spectrometry
- VII. Tandem Affinity Purification Using a LAP Tag
  - A. TEV Cleavage
  - B. S Protein Agarose
  - C. Sample Buffer Elution: For Silver Staining & Immunoblotting
  - D. Urea Elution: For Mass Spectrometry
- VIII. Mass Spectrometry & Prioritization for Follow-up Experiments
- IX. Summary
- X. Solutions and Media
  - A. Worm Reagents
  - B. Large-Scale Liquid Culture
  - C. Isolation of Nuclei from Embryos
  - D. Single-step Immunoprecipitation
  - E. Tandem Affinity Purification using a LAP Tag
- XI. Equipment
- Acknowledgments
- References

---

---

---

## Abstract

*C. elegans* is a powerful metazoan model system to address fundamental questions in cell and developmental biology. Research in *C. elegans* has traditionally focused on genetic, physiological, and cell biological approaches. However, *C. elegans* is also a facile system for biochemistry: worms are easy to grow in large quantities, the functionality of tagged fusion proteins can be assessed using mutants or RNAi, and the relevance of putative interaction partners can be rapidly tested *in vivo*. Combining biochemistry with function-based genetic and RNA interference screens can rapidly accelerate the delineation of protein networks and pathways in diverse contexts. In this chapter, we focus on two strategies to identify protein–protein interactions: single-step immunoprecipitation and tandem affinity purification. We describe methods for growth of worms in large-scale liquid culture, preparation of worm and embryo extracts, immunoprecipitation, and tandem affinity purification. In addition, we describe methods to test specificity of antibodies, strategies for optimizing starting material, and approaches to distinguish specific from non-specific interactions.

## ABBREVIATIONS

APC, Anaphase-promoting complex; CBP, Calmodulin-binding peptide; ChIP, Chromatin Immunoprecipitation; DMP, Dimethylpimelidimate; dsRNA, Double stranded RNA; GST, Glutathione S-transferase; h, Hour; IP, Immunoprecipitation; LAP, Localization and Affinity Purification; MosSCI, Mos1 mediated Single Copy transgene Insertion; nAChR, Nicotinic acetylcholine receptor; RNAi, RNA interference; RT, Room temperature; TAP Tandem Affinity Purification; TEV, Tobacco etch virus.

---

---

---

## I. Introduction

*Caenorhabditis elegans* is widely recognized as a powerful model system for cell and developmental biology. The landmark work that described the cell lineage from embryo to adult provided the foundation to study cell biology in the context of development in *C. elegans* (Sulston *et al.*, 1983). Research in *C. elegans* has traditionally emphasized genetic and physiological approaches to elucidate gene function. Classical epistasis analysis groups genes isolated by mutagenesis screens into distinct pathways (Huang and Sternberg, 2006). In the past decade, genome-wide RNAi screens have greatly accelerated the annotation of gene functions (Fernandez *et al.*, 2005; Kamath *et al.*, 2003; Piano *et al.*, 2000, 2002; Sonnichsen *et al.*, 2005). Until recently, biochemical studies have lagged behind, primarily due to the historical trajectory of *C. elegans* research. However, *C. elegans* is a facile system for biochemical approaches as it is straightforward to grow worms in large quantities, assess the functionality of tagged fusion proteins using mutants or RNAi, and test the relevance of any identified interacting protein rapidly through *in silico* analysis and *in vivo* methods (Audhya and Desai, 2008; Moresco *et al.*, 2010).

In this chapter, we focus on methods in *C. elegans* for isolating protein complexes and identifying new interacting proteins using mass spectrometry. In addition, we describe cloning vectors that are useful for protein tagging and methods to assess antibody specificity. To identify new interaction proteins we employ two major strategies. In the first strategy, the target protein is purified using single-step immunoprecipitation (IP) with an affinity purified polyclonal antibody and the entire immunoprecipitate is subjected to mass spectrometric analysis. Immunoprecipitation of the endogenous protein has several advantages: protein expression is controlled by the endogenous promoter and protein function is not altered by addition of a tag. The drawback of this approach is that a large number of proteins are detected using current highly sensitive mass spectrometry methods. This poses a challenge for discriminating between relevant and non-specific interactions and therefore the significance of co-purified proteins needs to be carefully evaluated in follow-up work. A potential additional drawback is that binding of the primary antibody may block association with a subset of interacting components. In the second strategy, we use tandem affinity purification (TAP) to isolate high

affinity complexes (Rigaut *et al.*, 1999). Two sequential affinity purification steps significantly reduce background and lead to clean isolation of protein complexes. Two potential drawbacks of this approach are that the tag has the potential to alter protein function and that the stringency of the two-step purification procedure may cause loss of low affinity interacting proteins. As either strategy has drawbacks, whenever possible we conduct both in parallel. Such a dual approach was critical in defining the protein network that constitutes the core microtubule-binding site of the chromosome during cell division (Cheeseman *et al.*, 2004; Desai *et al.*, 2003).

Below we discuss first the tools necessary for biochemical analysis of a protein of interest followed by detailed methods for large-scale worm culture, extract preparation, and protein complex isolation (Fig. 1). We additionally profile methods to assess specificity of antibodies and optimize starting material for complex isolations.

---

---

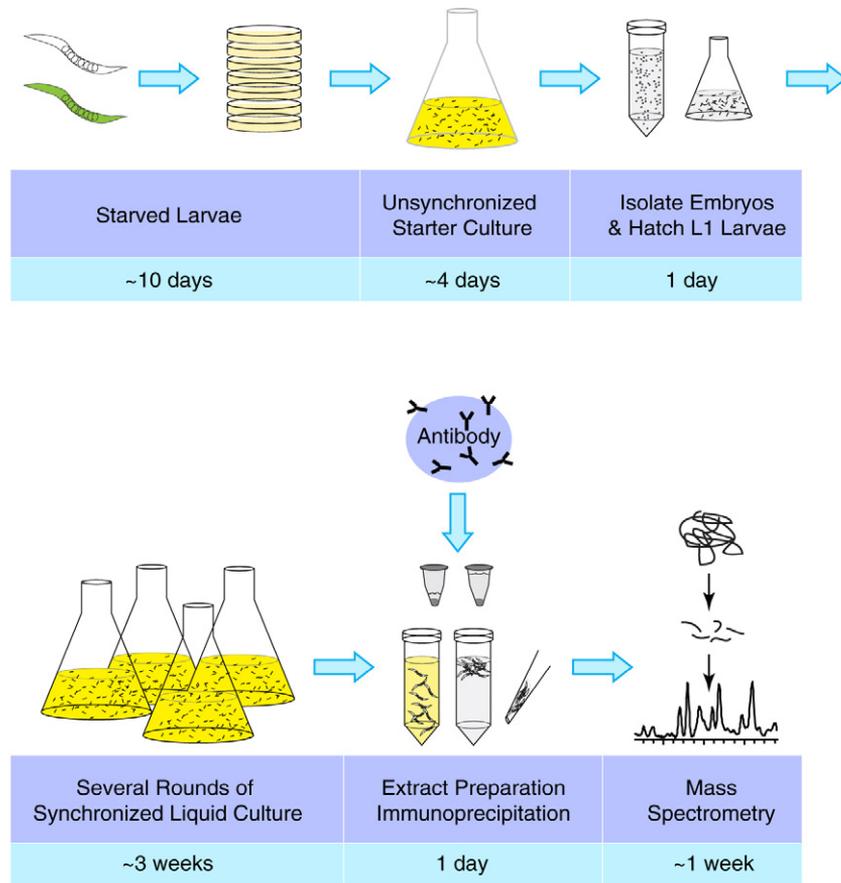
---

## II. Generating Tools for Biochemistry

### A. Generating a Polyclonal Antibody

We highly recommend that an affinity-purified antibody be generated to every protein of interest. An antibody that recognizes its target with high specificity is an invaluable reagent, and in most cases will be suitable for immunoprecipitation, localization studies, and immunoblotting. Ideally, two independent antibodies should be generated against non-overlapping epitopes. We have had good success with soluble fusion protein fragments of 100–200 amino acids expressed in bacteria, purified, and sent to a commercial vendor for antisera production (almost always in rabbits). The antibody is affinity purified from the antisera using columns with immobilized antigen that lack the fusion tag used for the initial antigen purification. Alternatively, antibodies against the tag can be first depleted from the antisera prior to affinity purification. Peptide antibodies can also be produced but, in our experience, have significantly lower rates of success than fusion protein antibodies. A newer option is DNA-based immunization, which does not require antigen purification for immunization but still requires purified protein for affinity purification of the antisera (Chowdhury, 2003).

The specificity of the affinity-purified antibody must be validated by immunoblot and immunofluorescence using wild-type and mutant backgrounds. If a mutant is not available, RNAi targeting the protein of interest should be performed. The most common cross-reactivity we have observed in fusion protein injection-generated antibodies is to *E. coli* proteins that are present in both the injected antigen and in the antigen preparations used for affinity purification. Because worms eat *E. coli*, bacterial protein epitopes are difficult to eliminate. If the protein of interest is present in embryos, the use of embryo isolated by bleaching avoids this problem as these do not have bacterial epitopes. Contaminating antibodies to bacterial proteins can be depleted using immobilized *E. coli* lysate (Thermoscientific Cat.



**Fig. 1** Experimental outline for protein complex identification in *C. elegans*. Wild-type or LAP-tagged strain is grown on NGM plates until larvae are starved. With starved larvae an unsynchronized starter culture is inoculated. Embryos are isolated by bleaching and hatched in the absence of food to obtain starved L1 larvae. Starved L1 larvae are used to set up six synchronized liquid cultures. After several rounds of synchronized liquid culture, when sufficient amounts of worms/embryos are obtained, the extract is prepared and protein complexes are purified by immunoprecipitation and analyzed by mass spectrometry. Approximate time for each experiment is indicated. (See color plate.)

#44938) and also blocked by adding an unrelated fusion protein preparation that harbors similar contaminants. For both immunofluorescence and immunoblotting, we typically use 0.5–1  $\mu\text{g}/\text{mL}$  affinity-purified antibody in the primary antibody incubation step. To deplete contaminating anti-bacterial protein antibodies, we incubate 500  $\mu\text{L}$  of 10  $\mu\text{g}/\text{mL}$  affinity purified antibody (diluted in AbDil: PBS

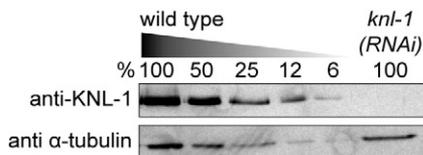
+ 2% BSA + 0.1% Triton X-100 + 0.1% sodium azide) with 100  $\mu$ L of *E. coli* protein agarose for 1–2 h at room temperature. The supernatant is then mixed with 50  $\mu$ g/mL final concentration of an unrelated fusion protein purified from bacteria (harboring the same tag as the antigen) and then used for immunoblotting. This procedure eliminates anti-bacterial protein antibody cross-reactivity even in highly sensitive chemiluminescent detection.

## B. Assessing Antibody Specificity by Immunoblotting after RNAi

The following protocol describes how to prepare worm extract for immunoblot analysis to test antibody specificity. It is important to wash the worms thoroughly to remove bacteria. Worms can be washed for up to 2 h in M9 containing 0.1% Triton X-100 if bacterial contamination remains a problem. As noted above, a good way to prevent bacterial epitopes in the sample is to use embryonic extract. However, a tradeoff with using embryo extract is that RNAi has to be performed by feeding, which might be less penetrant than injection. Therefore, we routinely perform RNAi blots using worms and, if necessary, treat the primary antibody to remove/block antibodies to bacterial epitopes.

To determine RNAi efficiency, a serial dilution of extract prepared from wild-type worms should be analyzed on the same blot as the RNAi extract sample (Fig. 2). As a loading control, a primary antibody of a different species should be used – we typically use anti  $\alpha$ -tubulin antibody that was generated in mouse (DM1A Sigma-Aldrich, Cat. #T9026).

RNAi-mediated depletion of gene products has shown to be effective for a large number of genes, including essential genes (Green *et al.*, 2011; Kamath *et al.*, 2003; Sonnichsen *et al.*, 2005). If reduction of the band detected by western blotting is not observed after RNAi this may be due to low RNAi efficacy or due to the antibody recognizing a non-specific band of similar molecular weight as the gene product of interest. In this case, alternate RNAi conditions (feeding, injection, soaking, temperature, and time) or, ideally, null mutants should be analyzed by western blotting to assess antibody specificity.



**Fig. 2** Immunoblot of wild-type and *knl-1*(RNAi) adult extract probed using anti-KNL-1 and anti  $\alpha$ -tubulin antibodies. A serial dilution of wild-type extract was loaded to determine the efficiency of KNL-1 depletion.

*Before starting*

Pipet 15  $\mu\text{L}$  water into two screw-cap tubes, mark the liquid level, and remove the water.

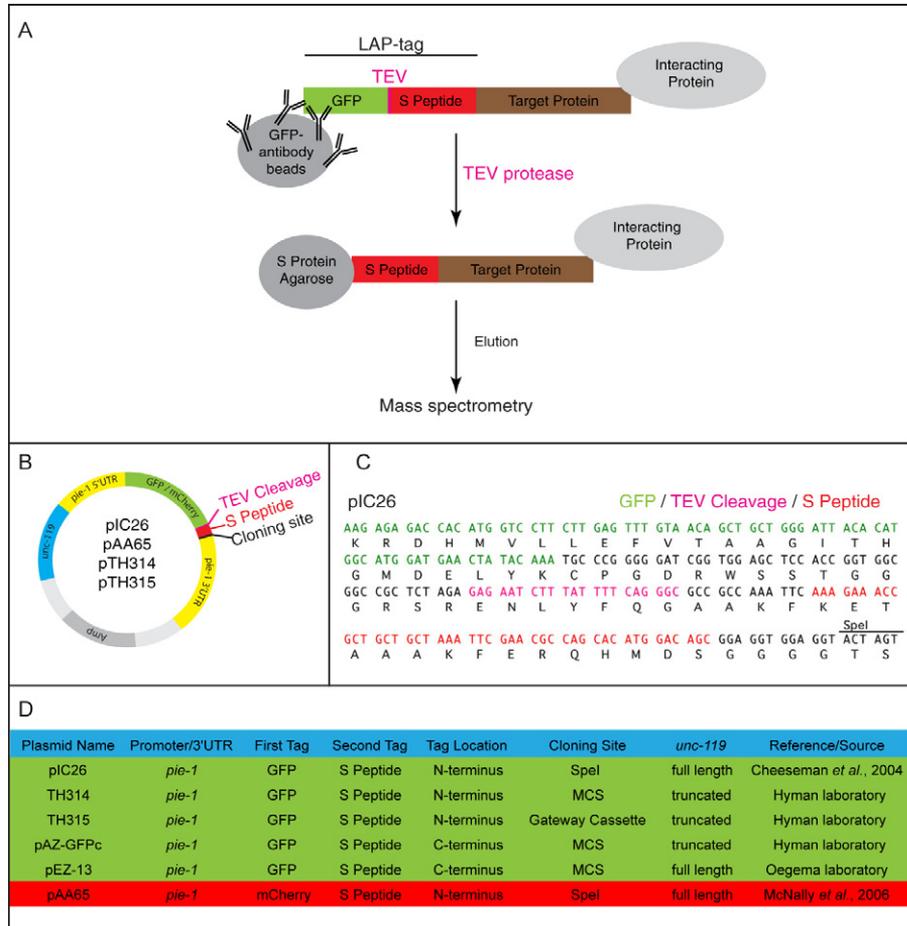
Put distilled water in sonicating water bath (Bransonic Ultrasonic Cleaner 3510) and turn heating to 80 °C or to maximum. If the sonicating water bath does not have heating capability, boiling water should be added prior to use.

1. Transfer 30 RNAi and 30 control worms into 0.5 mL M9 in marked tubes. We typically use injection as a method for RNAi as it has the best penetrance for early embryonic depletions, but soaking or feeding can also be used.
2. Add 1 mL M9 and pellet at 200g in a microcentrifuge. Carefully remove supernatant leaving the worms undisturbed.
3. Repeat step 2 two times.
4. After last wash, remove all liquid down to 15  $\mu\text{L}$  mark and add 15  $\mu\text{L}$  2 $\times$  Sample Buffer.
5. Place in sonicating water bath at 80 °C. Sonicate on maximum setting for 15–20 min.
6. Microcentrifuge at 200g and check that worms have dissolved (you should not see a pellet). If a significant pellet remains boil the samples at 95 °C in a heat block with intermittent vortexing.
7. Lightly centrifuge, mix by flicking, and load  $\sim$ 10  $\mu\text{L}$ /lane for immunoblots (aim for one worm per microliter of final sample).

**C. Tandem Affinity Purification Tags**

Tandem Affinity Purification allows the isolation of protein complexes in high purity. A composite tag is fused to the protein of interest containing two different purification tags separated by a protease cleavage site (Fig. 3). The original TAP tag used a domain of protein A that binds to IgG and a calmodulin-binding peptide (CBP) separated by a highly specific tobacco etch virus (TEV) protease site (Rigaut *et al.*, 1999). In *C. elegans*, a version of this approach that we have used with significant success is the localization and affinity purification (LAP) tag (Cheeseman and Desai, 2005). The LAP tag is a modification of the TAP tag that can be used to both affinity purify the fusion protein and study its *in vivo* localization dynamics. The LAP tag contains GFP (or mCherry) and S peptide (that binds S protein with high affinity) separated by a TEV cleavage site (Fig. 3A). The LAP-tag fusion protein is first purified using anti-GFP-coupled beads, released from the beads by TEV protease cleavage and further purified in a second step over S protein agarose. The LAP tag has been successfully used in several studies to isolate new protein complexes (Audhya *et al.*, 2005; Cheeseman *et al.*, 2004; Dammermann *et al.*, 2009; Gassmann *et al.*, 2008). When using this tag for analysis of protein complexes containing RNA, it should be kept in mind that the binding of S peptide to S protein reconstitutes an active ribonuclease.

Several LAP-tag containing vectors are available: pIC26 allows fusing the target protein at its N-terminus to the LAP tag using a SpeI restriction site (Fig. 3B–D)



**Fig. 3** TAP strategy and cloning vectors. A) The LAP-tagged target protein is first purified with anti-GFP antibody coupled beads. LAP-tagged target protein and interacting proteins are released from the beads by TEV protease cleavage and purified over S protein agarose. Interacting proteins are eluted from the beads and analyzed by mass spectrometry. B) Schematic vector map illustrating the common features of the standard LAP cloning vectors. C) Partial sequence of pIC26: in green 3' end of GFP, pink TEV protease cleavage site, red S peptide. The SpeI cloning site is indicated. D) Overview of the different features of the TAP cloning vectors. (See color plate.)

(Cheeseman *et al.*, 2004). pIC26 contains *pie-1* regulatory sequences to express the LAP-tagged protein in the germline and embryo and can be integrated by ballistic bombardment of the strain DP38 using *unc-119* as a transformation marker (see Section II.D.). pAA65 contains the same features as pIC26 but has mCherry instead of GFP as a fluorescent tag (Fig. 3B, D) (McNally *et al.*, 2006). *C. elegans* LAP vectors were reduced in size by introducing a truncated version of *unc-119* in TH314

and TH315 (Fig. 3D). However, incomplete rescue of the Unc phenotype makes the identification of transformants more challenging. The target gene can be inserted in these vectors at the 3' end of GFP by either Gateway cloning (TH315) or conventional cloning (TH314). To fuse the protein of interest at the C-terminus to the LAP tag either pAZ-GFPc (truncated *unc-119*) or pEZ-13 (full length *unc-119*) can be used (Fig. 3D). The LAP tag can easily be transferred into other vectors using the cassette present in pIC26.

Instead of fusing both tags to one protein, it is also possible to fuse them to different members of the same protein complex. Such a “split” TAP tag was used to isolate new binding partners of the integral membrane nicotinic acetylcholine receptor (nAChR) (Gottschalk *et al.*, 2005). Additional epitopes have been implemented for tandem affinity purifications (Polanowska *et al.*, 2004; Schaffer *et al.*, 2010).

#### D. Introduction of Transgenes for Expression of Tagged Proteins

For somatic expression of transgenes, heritable and repetitive extrachromosomal arrays are often sufficient; for example, Gottschalk *et al.* used an array to express TAP-tagged nAChR subunits (Gottschalk *et al.*, 2005). Injecting DNA in the *C. elegans* germline will generate extrachromosomal arrays (Mello and Fire, 1995; Mello *et al.*, 1991). However, a transgene in an array is typically overexpressed in somatic cells and rapidly silenced in germ cells (Kelly and Fire, 1998; Seydoux and Strome, 1999). The variable degree of heritability of the arrays can also make it difficult to obtain sufficient material from large-scale cultures.

An alternative to arrays is ballistic bombardment where small transgene-coated gold particles are introduced into the worm tissue at high speed (Praitis *et al.*, 2001). Bombardments are performed in the DP38 strain that carries a mutation in the *unc-119* gene. The DP38 strain is unable to move and does not transition to the dauer stage. A copy of the *unc-119(+)* gene is introduced in the same vector as the transgene and transformants are identified by wild-type movement and dauer formation. Ballistic bombardment yields low-copy number integrations at random sites in the genome. Bombarded transgenes may not be expressed at the endogenous level nor at all relevant developmental stages. Another drawback of generating transgenic lines by ballistic bombardment is that the integration sites are different for each transgene making it difficult to compare wild-type and engineered mutants. Detailed procedure for ballistic bombardment is described in Green *et al.* (2008) and on the Seydoux laboratory website (<http://www.bs.jhmi.edu/MBG/SeydouxLab/vectors/index.html>).

A recent technique, MosSCI (Mos1 mediated Single Copy transgene Insertion) circumvents the problems associated with arrays and bombarded lines by directing the transgene at a fixed locus in the genome (Frokjaer-Jensen *et al.*, 2008). Transformants are identified using the same strategy as for ballistic bombardment: the injected strain contains an *unc-119(ed3)* mutation that is rescued by introducing the wild-type *unc-119(+)* gene on the vector harboring the transgene. Description of

the vectors and a detailed protocol for MosSCI can be found on the webpage of the Jorgensen laboratory that developed this method (<http://sites.google.com/site/jorgensenmossci/Home>).

### E. Validating the Functionality of the Transgene-Encoded Tagged Protein

Before using the tagged protein for biochemical studies it is important to validate its functionality. The most straightforward test is whether the fusion protein can rescue a mutant phenotype. If a mutant is not available, functional tests can be conducted using RNAi. This requires that the endogenous messenger RNA is specifically targeted by a dsRNA. Re-encoding transgenes or using a different 3'UTR on the transgene are two strategies that have been used for this purpose (Audhya *et al.*, 2005; Dammermann *et al.*, 2008). A detailed description of how we re-encode transgenes was recently presented (Green *et al.*, 2008).

---

---

---

## III. The Optimal Starting Material for Protein Purification

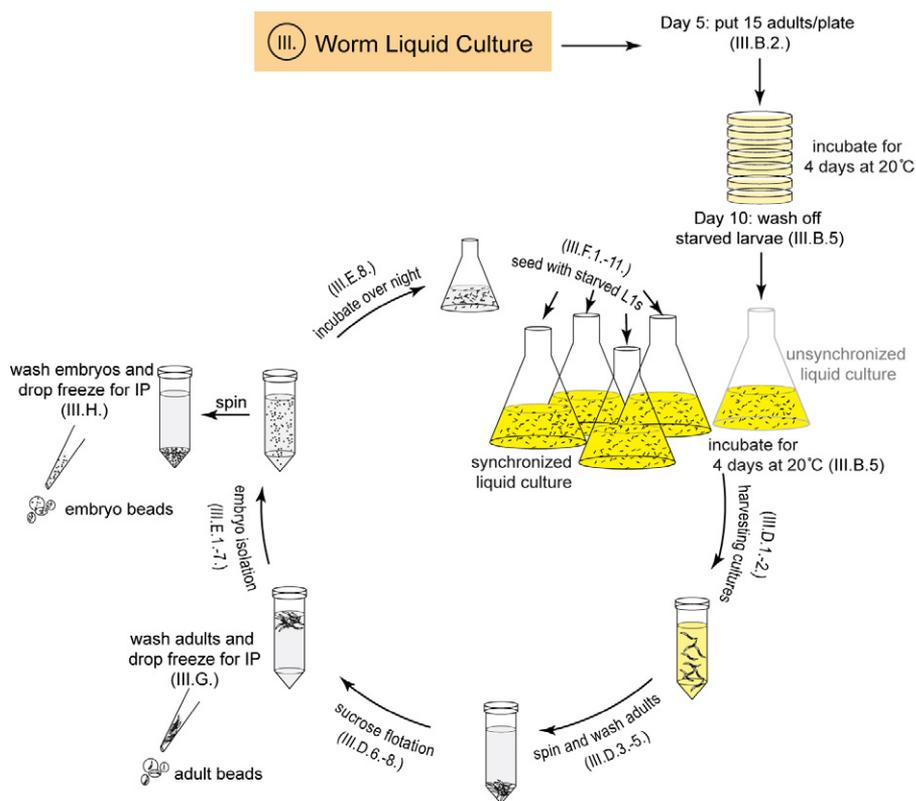
For proteins that are widely expressed, adult worms are the most straightforward starting material to use since they are easily grown in large-scale using liquid culture. If the protein of interest is enriched in embryos, embryonic extracts can be made by bleaching adult hermaphrodites to isolate embryos. By carefully choosing the incubation temperature it is also possible to enrich the embryonic extract either for young (<50 cells) or old embryos (>200 cells) (see Section III.I.). If the target protein is present only in specific cell types, one can try to enrich for these cell types by using conditional loss of function mutations. For example, by using a temperature-sensitive mutant that does not form a germline one can enrich for somatic cells (Beanan and Strome, 1992). We describe in Section III.I. a strategy that uses a temperature-sensitive meiotic arrest mutant to enrich for embryos in meiosis I. While flow cytometry sorting methods have been developed to isolate specific cell types for expression and small RNA analysis, the amounts enriched are not yet sufficient for biochemical methods (Cinar *et al.*, 2005; Colosimo *et al.*, 2004; Fernandez *et al.*, 2010; Zhang *et al.*, 2002). Future miniaturization of protein isolation methods is likely to enable proteomic characterization of individual cell types sorted by flow cytometry.

### A. Growing Worms in Large-Scale Liquid Culture

Worms are relatively straightforward to grow in biochemical quantities using liquid cultures. Growing worms on egg plates is an alternative approach and a detailed protocol for their use is described by Mains & McGhee (Hope, 1999). The following protocols describe growing worms in large-scale liquid culture to obtain sufficient amount of either adult worms or embryos to perform

immunoprecipitations or TAPs for mass spectrometry. The required buffers, solutions, and equipment are listed in Sections X and XI.

The standard procedure for liquid cultures is outlined in Fig. 4. First, adult worms are grown on OP-50 seeded 100 mm NGM agarose plates to obtain starved larvae that are used to initiate an unsynchronized liquid starter culture. Adults are harvested, embryos isolated by bleaching and hatched as L1s in the absence of food. Synchronized L1s are used to set up the second round of liquid culture and worms or embryos are harvested for extract generation. Worms and embryos can be stored at



**Fig. 4** Schematic overview of worm liquid culture (see Section III.). Fifteen axenized adults are grown ~4 days on 100 mm NGM agarose plates until larvae are starved (Section III.B.2.). Starved larvae are washed off the NGM plates and used to inoculate an unsynchronized liquid culture (Section III.B.5.). After 4 days, adult worms with 10–15 embryos are harvested, washed with M9, and cleaned by sucrose flotation (Section III.D.1.-8.). Embryos are isolated by bleaching and incubated overnight for L1 larvae to hatch (Section III.E.1.-8.). Starved L1 larvae are used to inoculate synchronized liquid culture (Section III.F.1.-11.). Harvested worms (Section III.G.) and embryos (Section III.H.) are drop frozen in liquid nitrogen for subsequent extract preparation. Roman numeral indicates corresponding section. (See color plate.)

–80 °C until sufficient amounts for immunoprecipitation are obtained. 500 mL liquid worm culture yields about 7 g of worms and 0.7 g of embryos.

Growth of worms in liquid culture should be carefully and continuously monitored. It is important to determine under a dissecting microscope the age of the worms, whether they have sufficient amounts of food, and if contamination is present. Cultures must be handled under sterile conditions, preferably in a dedicated laminar flow hood to avoid contamination. Growth temperature can be adjusted between 16 °C and 25 °C and it is recommended to use a cooling shaker that reliably holds temperature.

## B. Unsynchronized Liquid Starter Culture

Since contamination is a serious problem in liquid culture, we recommend starting with freshly axenized worms. These are then amplified on 100 mm NGM agarose plates by putting 15 cleaned worms onto each plate (Fig. 4). After incubating the plates for 4–5 days at 20 °C the just starved larvae are used to start the unsynchronized liquid culture.

1. *Day 1:* Pipet 5  $\mu$ L 2 M NaOH and 5  $\mu$ L of bleach at edge of bacteria on a 60 mm NGM agarose plate. Transfer 10 adult worms into NaOH/bleach drop. Wait until L1 larvae hatch and transfer L1 larvae onto new plates and grow until they are adults.
2. *Day 5:* Seed 100 mm NGM agarose plates with 15 cleaned adults. You will need 7–8 plates per 500 mL of liquid culture. Incubate plates at 20 °C for 4–5 days until worms are just starved.
3. *Day 8:* Start a 50 mL overnight of OP-50-1 in LB + 50  $\mu$ g/mL streptomycin. OP-50-1 is a streptomycin resistant *E. coli* strain.
4. *Day 9:* Start 1.5 L culture of OP-50-1 in LB + 50  $\mu$ g/mL streptomycin. You will need 1.5 L of bacterial culture as food for 500 mL of *C. elegans* culture.
5. *Day 10:*
  - Harvest bacterial culture at 4200 rpm for 15 min in sterile 1 L centrifuge bottles. Pour off LB. Bacterial food can be stored in 50 mL conicals at 4 °C for several weeks.
  - Make 500 mL Complete S Basal.
  - Resuspend bacterial pellet in 20 mL Complete S Basal and transfer into a sterile 2.8 L wide bottom Fernbach flask with 500 mL Complete S Basal.
  - Rinse 7–8 plates of just starved larvae with 10 mL of sterile M9. To collect as many worms as possible from plates repeat wash with 5 mL of M9. Check plates to make sure most worms were washed off. Collect in a 50 mL conical.
  - Pellet worms at 600g in centrifuge for 3 min with slow deceleration. Remove supernatant with a sterile pipet and discard. Resuspend worms in 5 mL of fresh M9 and add to flask containing the Complete S Basal with bacteria. Shake the inoculated culture at 20 °C at 200–230 rpm. Adjust growth time by varying temperature between 16 °C and 25 °C.

6. *Days 11–13*: Growth of the culture should be carefully followed (see Section III C.) and worms harvested when majority are adults. The culture will take about 3–3.5 days depending on age of worms on starter plates and the desired state of the final culture.

*Note*: The bacterial food for liquid culture can be obtained from fermentor facilities. However, we have had contamination as well as growth problems with externally supplied food and, consequently, prefer growing up our own bacterial cultures.

### C. Monitoring Worm Cultures

Worm liquid cultures should be monitored under a dissecting microscope once a day. This allows assessment of the developmental stage and health of the worms, ensures that worms have enough food, and confirms that there is no significant contamination.

1. In a laminar flow hood transfer 1 mL of the culture with a sterile pipet into 1.5 mL microcentrifuge tubes.
2. Spin in microcentrifuge 600g for 3 min.
3. Carefully pipet ~20  $\mu$ L worm slurry using a cut-off tip onto a glass slide. Place cover slip on top and look at the worms under dissecting microscope.

### D. Harvesting Worms from Liquid Cultures

Worms are harvested by settling under gravity and cleaned by flotation on a sucrose cushion, which separates healthy worms from bacteria and debris (Fig. 4). The sucrose floated worms are washed and used to isolate embryos.

1. When the majority of worms in the culture have 10–15 embryos, transfer the cultures into 500 mL graduated cylinders. Settle worms in ice water bath for 1 h.
2. Aspirate off media using a sterile 5 mL pipet and transfer slurry (brown film at bottom) to two 50 mL conicals per L culture.
3. Pellet in centrifuge at 600g for 3 min with slow deceleration.
4. Aspirate off supernatant, collect worms into 50 mL conical, and add cold M9.
5. Pellet in centrifuge at 600g for 3 min and aspirate off supernatant.
6. Resuspend pelleted worms by adding cold M9 to the 25 mL mark. To this add 25 mL cold 60% (w/v) sucrose. Mix and centrifuge immediately at 1500g for 5 min.
7. After the spin, adult worms will form a layer at the top of the tube. Remove adults down to the 35 mL mark with 5 mL pipet and transfer to a new conical.
8. Wash worms by adding cold M9 to 50 mL mark. Pellet the worms at 600g for 3 min and carefully remove supernatant.

*Note*: It is important to work rapidly during the sucrose flotation step. Do not leave the worms for too long in sucrose and wash them out of sucrose rapidly after collection from the top of the cushion.

### E. Isolation of Embryos and Synchronization as Starved L1 Larvae

The washed worms are bleached to isolate embryos (Fig. 4). Good bleaching efficiency depends on small volumes (maximum 5 mL worm pellet per 50 mL conical) and freshness of bleach. Bleaching should be followed by eye under a dissecting microscope. Bleaching sterilizes the culture; after bleaching, sterile technique becomes critical again so that no contaminants are introduced into the material that will be used to inoculate the synchronized liquid culture.

1. To worm pellet add 25 mL 0.1 M NaCl and mix by pipeting up and down twice. Settle worms on ice for 5 min.
2. Aspirate off supernatant including worms that have not settled to the bottom and add 0.1 M NaCl up to a volume of 30 mL.
3. Mix 5 mL 5 N NaOH with 10 mL bleach in conical. Immediately add NaOH/bleach mix to 30 mL worm suspension.
4. Vortex at maximum speed for 5 s and stand tube at RT for 2 min. Repeat four times for a total bleaching time of 7–9 min. Follow bleaching by examining samples on a glass slide under a dissecting microscope. Stop bleaching when only embryos remain. The color of the bleach mixture will change to a burnt orange as worms are dissolved.
5. Immediately centrifuge at 800g for 1 min at 4 °C. Aspirate off supernatant.
6. Add sterile water to a total volume of 50 mL, mix by inverting the tube, and centrifuge at 800g for 2 min.
7. Repeat step 6. Washed embryos from synchronized liquid culture can be frozen for immunoprecipitation as described in Section III.H.
8. Add 35 mL M9 and transfer to a 50 mL flask. Rinse the conical with an extra 10 mL of sterile M9 and add to flask. Shake at 20 °C until embryos hatch and are starved L1 larvae (~18–20 h).

### F. Seeding Synchronized Cultures Using Starved L1 Larvae

Starved L1 larvae from 500 mL starter flask can be expanded to inoculate up to six synchronized liquid cultures.

1. *2 Days before:* Start 300 mL culture of OP-50-1 in LB + 50 µg/mL streptomycin.
2. *Day before:* Start six 1.5 L cultures of OP-50-1 in LB + 50 µg/mL streptomycin.
3. Harvest bacterial cultures at 4200 rpm for 15 min in sterile 1 L centrifuge bottles. Pour off LB.
4. Make 3 L Complete S Basal and distribute into six 2.8 L Fernbach flasks.
5. Resuspend bacterial pellet of each 1.5 L culture in 20 mL Complete S Basal and transfer into sterile 2.8 L Fernbach flask with 500 mL Complete S Basal.
6. When flasks with bacterial food are ready, start processing starved L1s. In the hood, transfer L1s to a 50 mL conical. Chill on ice for 5–10 min.
7. Spin at 600g for 3 min. Carefully remove supernatant. Bring up to 50 mL with sterile cold M9.

8. Repeat step 7.
9. Add sterile M9 such that total volume is ~5 mL, transfer L1s to 15 mL conical, and pellet by spinning at 600g for 2 min. Immediately after the spin, use a pen to mark the volume of the pellet on the side of the tube. Estimate the volume of pellet by adding known volumes of water to a separate conical.
10. Resuspend L1s to total volume of 12 mL using sterile M9 and look at a sample under dissection microscope; estimate percentage of L1s relative to dead embryos/worm parts/clumps.
11. Seed each flask with equivalent of 50  $\mu$ L pure L1 pellet (e.g., if pellet volume is 0.6 mL and % L1 in the resuspension is 70% then seed with 1.4 mL of the resuspended pellet). Avoid overseeding or cultures will starve.
12. Put flasks at 20 °C at 230 rpm. Grow for 48 h while monitoring cultures under dissecting microscope.

### G. Freezing Adult Worms for Immunoprecipitation

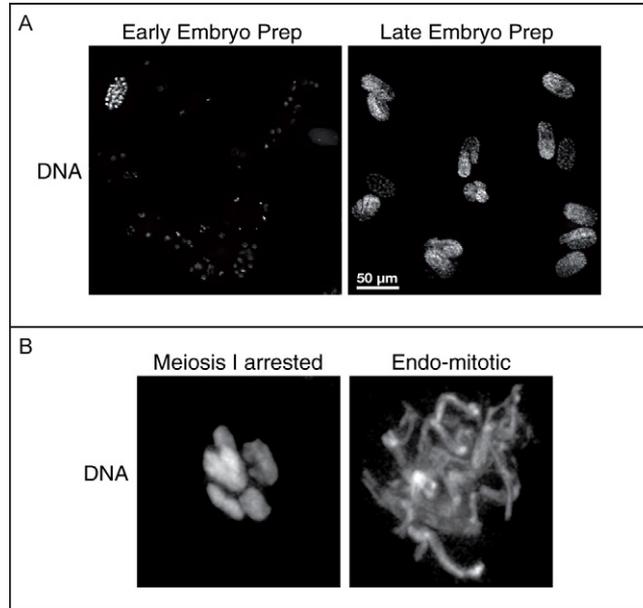
Once the liquid cultures are ready collect and wash worms as described in Section III.D. Wash worms by adding 50 mL cold 1 $\times$  Lysis Buffer (with protease inhibitors). Remove Lysis Buffer until only a small amount remains. Freeze adult worms by dispensing from a pipette drop by drop in liquid nitrogen, which will form small beads (Fig. 4). Store at  $-80^{\circ}\text{C}$ .

### H. Freezing Embryos for Immunoprecipitation

Bleach adult worms as described in Section III.E. and continue after step 7. Wash embryo pellet with 50 mL cold 1 $\times$  Lysis Buffer (with protease inhibitors). Freeze embryos by dispensing from a pipette drop by drop in liquid nitrogen. Store embryo beads at  $-80^{\circ}\text{C}$ .

### I. Enriching for Specific Age Embryos

While precise synchronization of embryos is not possible, it is straightforward to enrich for old or young embryo populations by varying growth conditions and carefully monitoring worms in the culture under a dissecting microscope. Generate synchronized starved L1 larvae (see Section III.E.) and inoculate worm cultures (see Section III.F.). To obtain worms that just started embryo production, incubate the flask for ~64 h at 17 °C. Using these growth conditions worms typically contained up to five embryos, the majority of which have <50 cells (Fig. 5A). If several flasks of worm culture are grown simultaneously, it is necessary to monitor each flask separately, as the time at which embryo production begins may vary between flasks. It is also critical to avoid contamination, which may adversely affect synchronous growth of the worms. To obtain worms with mostly old embryos (>200 cells), cultures are incubated for ~64 h at 19 °C (Fig. 5A). Embryos are frozen as described in Section



**Fig. 5** A) Early and late embryo preparation stained with Hoechst. B) Meiosis I arrested and endo-mitotic chromosomes stained with Hoechst.

III.H. To determine the age of the embryos, about  $\sim 5 \mu\text{L}$  of packed embryos are fixed in 1 mL of cold methanol for at least 30 min, then incubated with  $1 \mu\text{g/mL}$  Hoechst stain in PBS+0.1% Triton X-100 for 10 min. After three washes with PBS+0.1% Triton X-100, embryos are suspended in  $30 \mu\text{L}$  PBS, and  $5 \mu\text{L}$  are mixed on a  $18 \times 18$  mm coverslip with  $15 \mu\text{L}$  of mounting medium. The coverslip is carefully placed on a slide, sealed with nail polish, and the number of nuclei per embryo is determined using a fluorescence microscope (Fig. 5A).

### J. Enriching for Meiosis I Arrested Embryos

Adult worms contain about 15 mitotically dividing embryos and at most two meiotic embryos because meiosis is completed 30 min after fertilization. To enrich for meiosis I metaphase, we used a mutation in a subunit of the anaphase-promoting complex (APC). We chose a temperature-sensitive allele (*g48*) of *emb-27* (Cassada *et al.*, 1981; Golden *et al.*, 2000). At permissive temperature ( $16^\circ\text{C}$ ) *emb-27(g48)* worms contain about 15 mitotically dividing embryos, similar to wild type. At the restrictive temperature ( $24\text{--}25^\circ\text{C}$ ), *emb-27(g48)* embryos arrest at metaphase of meiosis I and worms accumulate meiotically-arrested fertilized embryos. Mutant worms are initially grown at  $16^\circ\text{C}$  as described in Section III.

A.-E. It is important to keep the temperature at or below 16 °C to prevent metaphase arrest. Bleached larvae are used to inoculate a synchronized liquid culture (see Section III.F.) and are shifted to 24 °C when the majority of worms are at the L4 stage. The time point of temperature shifting has to be chosen carefully because an early shift causes larval arrest and a late shift will “contaminate” the extract with mitotic embryos. Worms are harvested when the majority contains 4–8 one-cell meiotic-arrested embryos and worm extract is prepared as described in Section V. Timing is once again critical since after 4–5 h the embryos overcome the meiotic arrest and start cycling endo-mitotically (DNA is replicated but cell division does not occur). The quality of the culture can be assayed by cutting 2–3 worms in a drop of L-15 blastomere culture medium (Edgar, 1995) containing 1 µg/mL Hoechst 33342, and analyzing the shape of the chromosomes. Metaphase I arrested embryos display six maternal chromosomes that have a typical oval shape (Fig. 5B). Endo-mitotic embryos contain decondensed chromosomes that have a fibrous-like appearance and tend to detach from the anterior cortex of the embryo (Fig. 5B).

The above two examples are specific to our research interests but related strategies can be used to enrich for cell types or stages of interest prior to protein complex isolations.

---

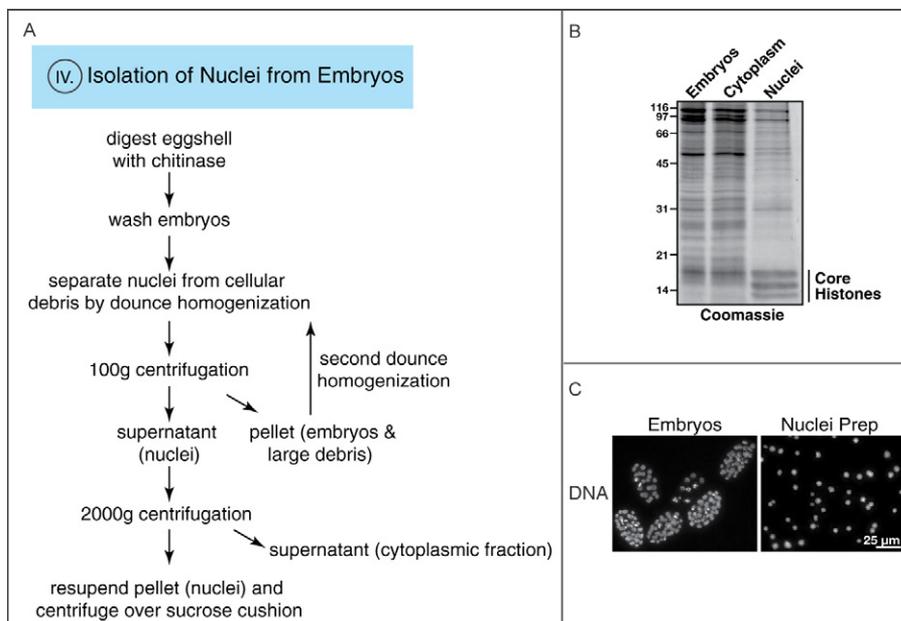
---

---

#### IV. Isolation of Nuclei from Embryos

If the protein of interest is a nuclear protein or if DNA–protein interactions are being analyzed as in Chromatin Immunoprecipitation (ChIP) experiments, nuclei can be isolated from embryos prior to further analysis. Embryos are isolated by bleaching as described in Section III.E Steps 1–7; however it is important that embryos are not frozen prior to nuclei isolation. To isolate nuclei, the embryo eggshell is digested using chitinase (Fig. 6A). Dounce homogenization in hypotonic buffer liberates intact nuclei. A 100g centrifugation removes large debris and a subsequent 2000g centrifugation pellets the nuclei. A final centrifugation step over a sucrose cushion is used to separate the nuclei from membrane contaminants. Enrichment of nuclei can be easily followed by the emergence of the core histone bands on Coomassie-stained gels (Fig. 6B). Because the chitinase step is performed at room temperature, nuclei must be purified from freshly isolated embryos to avoid protein degradation. 1 mL of embryo pellet typically yields ~100 µL of nuclei.

1. Harvest embryos by bleaching as described in Section III.E. until step 7.
2. Add two embryo pellet volumes of Embryo Buffer and 250 µL of chitinase stock solution per mL of embryo pellet.
3. Rotate at RT for ~30 min. Monitor eggshell integrity under dissecting microscope. Embryos without an eggshell will lose their oval shape and fall apart into clumps of cells.



**Fig. 6** A) Experimental outline for nuclei isolation. Eggshell of the embryos is digested by chitinase treatment and embryos are washed. Dounce homogenization in hypotonic buffer separates the nuclei from cellular debris. Hundred  $\times$  g centrifugation pellets embryos and cellular debris, which are subject to a second dounce homogenization. The supernatant containing the nuclei is centrifuged at 2000g to pellet the nuclei. Nuclei are finally purified by centrifugation over a sucrose cushion. Roman numeral indicates corresponding section. B) Embryos, cytoplasmic fraction, and isolated nuclei stained with Coomassie. Histones are enriched in the isolated nuclei. C) Embryo and nuclei preparation stained with Hoechst. (For color version of this figure, the reader is referred to the web version of this book.)

4. Spin at 1000g for 3 min.
5. Wash pellet twice with 50 mL cold Embryo Buffer. All subsequent steps are performed on ice.
6. After final wash, remove supernatant and resuspend embryos in 10 mL 0.5 $\times$  Nuclei Buffer. Incubate for 15 min on ice to swell cells.
7. Add 10 mL 1 $\times$  Nuclei Buffer (with 0.1% digitonin and protease inhibitors) and immediately dounce with  $\sim$ 50 strokes in a 15 mL Wheaton dounce homogenizer with a B pestle (tight fit).
8. Spin at 100g for 3 min. This centrifugation pellets large debris.
9. Remove supernatant that contains the nuclei and keep on ice.
10. Resuspend pellet in 10 mL Nuclei Buffer (with 0.1 % digitonin and protease inhibitors) and dounce pellet as described in step 7.
11. Spin at 100g for 3 min and combine two supernatants.
12. Spin at 2000g for 15 min to pellet nuclei. Supernatant is cytoplasmic fraction.

13. Resuspend pellet in 500  $\mu\text{L}$  Nuclei Buffer (with 0.1% digitonin and protease inhibitors) and mix with 5 mL 30% (w/v) sucrose cushion in Nuclei Buffer with 0.1% digitonin.
14. Centrifuge at 2000g for 15 min. The pellet will be enriched for nuclei.
15. Check integrity of the nuclei under a microscope after incubation with 1  $\mu\text{g}/\text{mL}$  Hoechst stain (Fig. 6C).



## V. Preparing Worm and Embryonic Extract

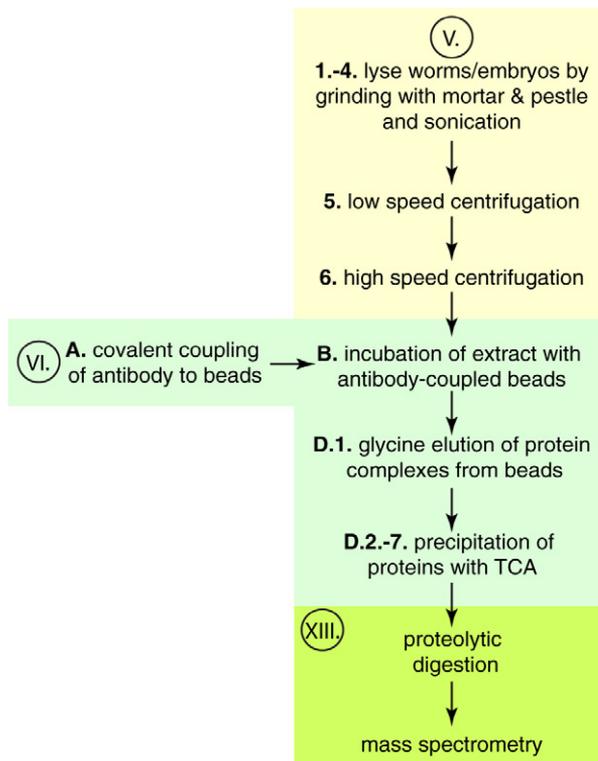
Extract can be prepared using whole worms harvested from liquid culture or from embryos isolated by bleaching. Lysis of worms and embryos is performed in an isotonic buffer. Subsequent to lysis, the salt concentration can be raised to 300 mM KCl to enhance stringency of protein complex isolation. Worms and embryos are lysed by grinding in liquid nitrogen followed by sonication (Fig. 7). Two consecutive centrifugations remove membranes and lipids and the supernatant is used for complex isolation. The lysis conditions described here are not well-suited for membrane-associated proteins. A protocol for isolating membrane proteins is described in Gottschalk *et al.* (2005).

1. Pre-cool a mortar and pestle by filling with liquid nitrogen for at least 5 min.
2. Weigh out frozen adult/embryo beads. For one immunoprecipitation use  $\sim 1$  g of frozen worm/embryo beads. Grind frozen beads to a fine powder: initially break down worm/embryo beads by gentle tapping (try not to lose too many beads as they have a tendency to jump out of the mortar), then grind. Keep mortar cold by adding more liquid nitrogen as necessary and waiting for it to evaporate. For more than 40 g of worms, one can also use a warring blender cooled with liquid nitrogen. For embryo extracts, the freeze-grinding step may be skipped, as sonication is sufficient to release cytoplasmic contents.
3. Add an equal volume of  $1.5\times$  Lysis Buffer (with protease inhibitors) to each gram of adult/embryo beads. Keep 20  $\mu\text{L}$  sample for gel analysis.
4. Set up ice-water bath and sonicate with a tip sonicator (e.g., Branson Digital Sonifier). It is critical to prevent heating of the sample during sonication. For a Branson Digital Sonifier with a microtip use the following settings:
  - 30% amplitude for 3 min total (15 s on; 45 s off—after each 1 min of sonication wait  $\sim 2$  min to chill)
  - 40% amplitude for 30 s (15 s on; 45 s off) Save 20  $\mu\text{L}$  sample.

We recommend optimizing the sonication protocol by two methods:

Use a dissecting microscope to monitor worm/embryo lysis. At least 80–90% of the worms/embryos must be lysed.

Use Bradford reagent or UV absorbance to directly monitor lysis. As cells lyse, the protein concentration and  $A_{260}$  absorbance (due to nucleic acid release) will



**Fig. 7** Experimental outline for protein complex purification. Roman numerals indicate the corresponding sections. A crude extract is generated by grinding and sonicating the worms and embryos (Section V.1.-4.). Two consecutive centrifugations clear the extract from debris and membranes (Section V.5.-6.). Antibodies are covalently coupled to beads (Section VI.A.) and extract is incubated with antibody-coupled beads (Section VI.B.). Protein complexes are eluted with glycine (Section VI.D.1.) and precipitated with TCA (Section VI.D.2.-7.). Prior to mass spectrometry the eluate is proteolytically digested (Section VIII.). (For color version of this figure, the reader is referred to the web version of this book.)

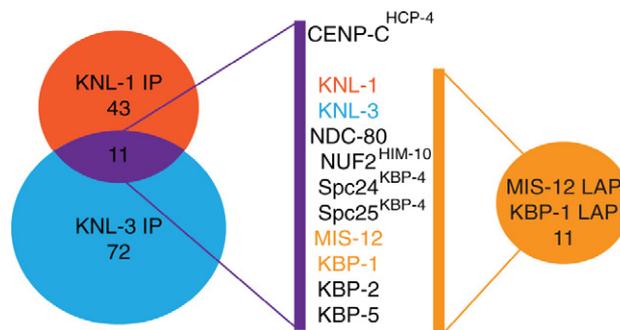
increase. A test sonication should be performed to define conditions at which one or both of these values begin to plateau.

5. Transfer crude extract to TLA100.3 tube and spin at 20,000g for 10 min at 2 °C, with medium deceleration. Save 20  $\mu$ L sample.
6. Transfer supernatant to new tube and pellet at 100,000g for 20 min at 2 °C. Try to avoid lipid and re-pellet if too cloudy. Save 20  $\mu$ L sample.
7. Transfer supernatant to a tube on ice. This extract will be used for purifications. All samples and the extract may be frozen in liquid nitrogen and stored at  $-80$  °C until further use.
8. Use Bradford reagent to measure protein concentration in extract.

## VI. Single-step Immunoprecipitation

Immunoprecipitation using specific antibodies is a powerful method for analyzing protein–protein interactions and identifying protein complexes. Single-step purification using polyclonal antibodies followed by mass spectrometry on the entire eluate commonly identifies several hundred to a thousand proteins making it challenging to distinguish between signal and noise. Background can arise from general non-specific binding to beads, to constant regions of antibody chains, and to partially denatured antibodies. Background can also be specific to individual affinity-purified antibodies – consequently a random IgG control cannot be used to discriminate between true signal and noise. We recommend using as a control a polyclonal rabbit antibody raised against GST (glutathione S-transferase) that is affinity-purified using the same procedures used for the antibody to the target protein. In addition, whenever possible, we recommend parallel immunoprecipitations with either two antibodies to the same protein (preferable with non-overlapping epitopes) or one antibody and one tagged fusion protein. Purification from a mutant strain is an ideal negative control but is not feasible if the mutation is lethal. If a tagged protein is immunoprecipitated, an untagged strain can be used as a negative control.

The following two examples illustrate how redundant purification strategies and suitable negative controls helped pinpoint *bona fide* complex members. Polyclonal rabbit antibodies to two essential chromosome segregation proteins, KNL-1 and KNL-3, identified in an RNAi screen, were used to purify complexes containing these proteins (Cheeseman *et al.*, 2004; Desai *et al.*, 2003). Each protein was present in the other immunoprecipitation, allowing cross-referencing of the two immunoprecipitations to identify 11 proteins in common (Fig. 8) (in total the two



**Fig. 8** Identification of 10-protein kinetochore complex. Immunoprecipitation of KNL-1 and KNL-3 isolated more than 130 interacting proteins with 11 proteins common in both immunoprecipitations. MIS-12 and KBP-1, two common interactors, were LAP tagged and purified isolating 11 common interactors of which 10 were found also in KNL-1 and KNL-3 immunoprecipitations. (For color version of this figure, the reader is referred to the web version of this book.)

immunoprecipitations identified over 130 proteins). MIS-12 and KBP-1, two of the newly discovered proteins that were in the common set were LAP tagged and tandem affinity purified revealing a 10-protein complex containing both KNL-1 and KNL-3 (Fig. 8). Functional validation confirmed that all of the proteins that co-purified were involved in chromosome segregation.

In a second example, novel interaction partners of *C. elegans* Dicer (DCR-1), a RNA endoribonuclease that is required in mammalian cells for small RNA generation, were identified by single-step purification (Duchaine *et al.*, 2006). Two single-step DCR-1 immunoprecipitations were performed: one using adult worms with a polyclonal anti-DCR-1 antibody and one using embryos that express HA-tagged DCR-1. Negative control purifications were performed in a *dcr-1* mutant strain, an untagged wild-type strain and with unrelated antibodies. Co-purified proteins were considered high confidence interactors if they were reproducibly identified in at least two independent purifications and not present in the negative control purifications. Twenty high confidence interactors were analyzed further with biological assays. Twelve of the 20 interactors could be linked to DCR-1 activities or its small-RNA products.

These two examples illustrate the importance of experimental design that incorporates redundancy at the outset. The need for employing such strategies has been magnified by the significant technical advances in protein mass spectrometry – currently, one-step immunoprecipitations with an affinity-purified antibody frequently yield between 100 and 1000 distinct proteins in the immunoprecipitate. We note that existing datasets can provide computational tools for filtering the output of mass spectrometric analysis. For example, over the years we have conducted a large number of protein isolation and mass spectrometry experiments in *C. elegans* embryos. The targets for this analysis span a number of cellular processes. This cumulative dataset helps identify potentially interesting candidates from a new purification experiment as opposed to frequently observed co-purifying “sticky” proteins. New label free quantitative mass spectrometry methods can also be employed to help discriminate true signal from noise (Hubner *et al.*, 2010).

We note that, prior to mass spectrometry, it is important to assess the solubility of the protein target and the efficiency of the immunoprecipitation using immunoblotting. For this purpose, samples of the crude extract, low-speed supernatant, high-speed supernatant, the supernatant after incubation with antibody beads, and the immunoprecipitate are important to prepare and store, keeping track of the volumes at each step. Such an effort can quickly identify potential confounding problems (e.g., low solubility or poor immunoprecipitation efficiency) and guide approaches to overcome them.

Below, we describe how to covalently couple the affinity-purified antibody to protein A beads, use the antibody-coupled beads for immunoprecipitation, and elute bound proteins from the beads for analysis by mass spectrometry, silver staining and immunoblotting (Fig. 7).

### A. Covalent Coupling of Antibodies to Protein A Beads

Immunoprecipitation and tandem affinity purification both start with coupling the antibody to Protein A resin. For antibodies generated in species with weak affinity to protein A, protein G resin should be used. Coupling greatly reduces antibody leaching during elutions. The tradeoff is a reduction in antibody efficacy because coupling is performed using a bifunctional amine crosslinker that to a variable extent will react with the antigen-binding site. The use of the crosslinker dimethylpimelidate (DMP) for coupling is based on the protocol described in Harlow and Lane (Harlow *et al.*, 1999).

Amounts indicated are for one tube of coupled beads

- For one TAP purification prepare four tubes.
  - For one-step immunoprecipitation used for mass spectrometry prepare one tube.
  - Volume of Protein A beads indicated is for settled beads, material received from Biorad is a 1:1 slurry.
1. Equilibrate ~120  $\mu$ L Affi-Prep Protein A beads into PBST (PBS + 0.1% Tween-20) by washing the beads three to four times with 1 mL PBST. This should give about 60  $\mu$ L of packed beads. Wash beads by gentle inversion. Briefly pellet them using a pulse in a microcentrifuge (30 s at 3000g) and remove supernatant avoiding bead pellet.
  2. Resuspend beads in 500  $\mu$ L PBST and add 10–50  $\mu$ g of affinity-purified antibody. Mix for 45 min – 1 h at RT on a rotor to allow antibodies to bind to resin.
  3. Wash beads three times with 1 mL PBST as described in step 1.
  4. Wash beads three times with 1 mL 0.2 M sodium borate (pH 9.0) (dilute from a stock of 1 M sodium borate (pH 9.0)). After the final wash, add 900  $\mu$ L of the 0.2 M sodium borate (pH 9.0) to bring the final volume to ~1 mL.
  5. To initiate coupling add 100  $\mu$ L of 220 mM DMP. Rotate tubes gently at RT for 30 min.
  6. *To make DMP:* Let bottle sit tightly closed at RT for 20 min before opening. Weigh out DMP and leave dry until just before use. Resuspend in appropriate volume of 0.2 M sodium borate (pH 9.0) and add immediately to the bead suspension (e.g., for 34 mg DMP add 596  $\mu$ L sodium borate).
  7. After incubation with DMP, wash beads two times with 1 mL 0.2 M ethanolamine, 0.2 M NaCl (pH 8.5) to inactivate the residual crosslinker. Resuspend in 1 mL of the same buffer and rotate for 1 h at RT. Resuspend beads in 500  $\mu$ L of the same buffer. Leave the beads in 0.2 M ethanolamine, 0.2 M NaCl (pH 8.5) at 4 °C until use. Beads are stable at 4 °C for at least one month.

### B. Immunoprecipitation

To prevent proteolysis it is important to keep the beads on ice and cool all buffers and tubes before use. Using higher stringency conditions (300 mM KCl) reduces

background although there is also an increased likelihood of losing meaningful low affinity interactions.

1. Pre-elute 100  $\mu$ L of antibody-coupled beads three times with 1 mL of 100 mM glycine (pH 2.6) to remove antibody that is not covalently coupled to the beads. Do not leave beads for a long time in glycine or the antibody will denature.
2. Wash beads three times with 1 mL cold Lysis Buffer (with 0.5 mM DTT) to neutralize glycine and prepare the beads for immunoprecipitation.
3. Mix beads with 900  $\mu$ L extract for at least 1 h at 4 °C on rotating platform.
4. Rinse beads two times with 1 mL cold Lysis Buffer (with 0.5 mM DTT and protease inhibitors).
5. On rotator in cold room wash beads two times for 5 min with 1 mL cold Lysis Buffer (with 0.5 mM DTT and protease inhibitors).
6. Wash five times with 1 mL Lysis Buffer (with 0.5 mM DTT) without detergent (NP-40) or protease inhibitors. Remove as much supernatant as possible.

*Note:* The presence of detergents can interfere with mass spectrometry. Therefore, it is important to wash the sample thoroughly in detergent-free buffer after immunoprecipitation.

### C. Sample Buffer Elution: For Silver Staining & Immunoblotting

1. Elute beads by heating in 100  $\mu$ L of 2 $\times$  Sample Buffer without DTT for 10 min at 70 °C.
2. Pellet beads, transfer supernatant to a new tube, and add DTT to 100 mM (1/9 supernatant volume of 1 M DTT stock; Elution 1).
3. Add 100  $\mu$ L 2 $\times$  Sample Buffer with DTT to pelleted beads (Elution 2).
4. Boil both elution samples for 5 min and analyze by silver staining or immunoblot. Both elutions will contain immunoprecipitated proteins although amounts in each may vary; Elution 2 will have more IgG contamination than Elution 1.

*Note:* For immunoblots conducted using rabbit primary antibodies, the secondary antibody will detect any IgG released by the elution from the beads. For silver staining, load 5–10  $\mu$ L directly. For immunoblots, load 5–10  $\mu$ L of a 1/10 dilution made in Sample Buffer.

### D. Glycine Elution: For Mass Spectrometry

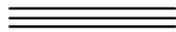
1. After standard immunoprecipitation (see Section VI.B), elute beads three times with 150  $\mu$ L 100 mM glycine (pH 2.6). Pool elutions and neutralize by adding 150  $\mu$ L 2 M Tris (pH 8.5). Neutralize the beads by washing two times with 150  $\mu$ L Lysis Buffer (without detergent) and pool with eluate. Total volume will be 900  $\mu$ L. Make sure you remove all the beads.

2. Add 1/5th volume 100% trichloroacetic acid (TCA; ~200  $\mu$ L).
3. Leave samples at 4 °C overnight.
4. Spin for 30 min at maximum speed in microcentrifuge. Remove supernatant and spin again for 1 min. Remove any residual supernatant with gel loading tip, leaving 5–10  $\mu$ L behind.
5. Wash twice with 500  $\mu$ L cold acetone. Spin 10 min at maximum speed in microcentrifuge.
6. Dry the protein pellet by spinning briefly in a speed vac.
7. Freeze in liquid nitrogen and store at –80 °C. The protein pellet is suitable for direct mass spectrometric analysis.
8. After elution with glycine, the beads should be boiled in Sample Buffer and analyzed by silver staining/immunoblotting to assess elution efficiency.

### E. Urea Elution: For Mass Spectrometry

Urea elution can also be used for mass spectrometry. In practice, we find that glycine elution works better for elution of the antigen from the antibody and for detection of the purified antigen in mass spectrometric analysis.

1. After standard immunoprecipitation (see Section VI.B), wash beads with Pre-urea Wash Buffer. Remove all residual supernatant.
2. Add 75  $\mu$ L Urea Elution Buffer and rotate for 30 min at RT.
3. Pellet beads and transfer eluate to a new tube. Re-pellet to ensure removal of all protein A beads.
4. Remove 50  $\mu$ L of elution and drop freeze in liquid nitrogen to send for mass spectrometry. Add Sample Buffer to the rest to run on a gel.



## VII. Tandem Affinity Purification Using a LAP Tag

As a first step, the extract is incubated with anti-GFP antibody-coupled protein A beads (Fig. 3A). For this purpose, we use in-house rabbit polyclonal anti-GFP antibodies generated by injecting purified GFP in rabbits. Recombinant GFP-binding domains from single-chain antibodies have also been used successfully for affinity purification of GFP-tagged proteins (Trinkle-Mulcahy *et al.*, 2008). After immunoprecipitation of GFP, the fusion protein is released by TEV protease cleavage (Fig. 3A). The subsequent purification on S Protein agarose further enriches for complexes containing the fusion protein. Extract prepared from an untagged strain can serve as negative control for the TAP procedure.

### A. TEV Cleavage

Perform immunoprecipitation using anti-GFP antibody-coupled beads as described in Section VI.B.

*Note:* At this stage, it is possible to elute bound material using glycine as described in Section VI.D. and compare the elution of the one-step GFP immunoprecipitation to the two-step LAP.

1. Pool beads into a single tube and fill with Lysis Buffer (with 300 mM KCl, 0.5 mM DTT).
2. Add ~30  $\mu$ L of purified TEV protease (1 mg/mL) and rock tubes for 4 h at 4 °C. Add an additional 30  $\mu$ L of TEV and rock tubes overnight.
3. Pellet beads and transfer supernatant to a new tube. Add 350  $\mu$ L Lysis Buffer to beads to remove any residual cleaved protein.

### **B. S Protein Agarose**

1. Wash tube of 85  $\mu$ L S protein agarose three times with 1 mL Lysis Buffer (with 300 mM KCl).
2. Add TEV protease eluted supernatant to S protein agarose and rock for 3 h at 4 °C.
3. Pellet beads and wash three times with Lysis Buffer (with 300 mM KCl).
4. Wash one time with Lysis Buffer with 100 mM KCl without detergent (NP-40).

### **C. Sample Buffer Elution: For Silver Staining & Immunoblotting**

Perform Sample Buffer Elution as described in Section VI.C.

### **D. Urea Elution: For Mass Spectrometry**

Perform Urea Elution as described in Section VI.E.

## **VIII. Mass Spectrometry & Prioritization for Follow-up Experiments**

Protein mass spectrometry has made remarkable advances in the recent decade. We will not discuss the details of the methodology, which are extensively reviewed elsewhere (Cravatt *et al.*, 2007; Yates *et al.*, 2009). We typically do not separate proteins from immunoprecipitates on gels prior to mass spectrometry. Instead, the entire eluate is proteolytically digested and the peptide mixture is separated by multidimensional liquid chromatography. The mass/charge ratio of ionized peptides is determined in the first mass analyzer, and then the peptides are fragmented in the collision cell and passed through the second mass analyzer to determine amino acid sequence. As mentioned above, current methods can yield up to 1000 proteins in one-step immunoprecipitation performed using an antibody that passes generally accepted antibody specificity criteria. This abundance of information makes it

challenging to sort relevant hits from background and necessitates redundant strategies. New technical developments, including labeling-based and label-free methods, may also aid in this effort. In practical terms, the challenge in analyzing protein mixtures by mass spectrometry lies primarily in developing a good working relationship with a mass spectrometry-focused laboratory or core facility.

Prioritization of potential interactors from lists generated by mass spectrometry of immunoprecipitates/tandem affinity purifications for follow-up experiments is perhaps the greatest challenge faced in biochemical analysis of protein complexes. In this regard, it is helpful to think of the initial list of proteins identified as hits from a primary screen, with a variety of secondary screens being necessary to separate the wheat from the chaff. Redundant strategies and elimination of common contaminants are important means for filtering such lists. In addition, the extensive genomic resources available in *C. elegans* that are archived on Wormbase provide an invaluable tool. Finally, the ability to rapidly test gene function is perhaps the most important, as it motivates in-depth analysis to validate the initially observed putative physical interaction.

---

---

---

## IX. Summary

Although genetic and cell biological analysis continue to be central to elucidating gene function in *C. elegans*, identifying protein–protein interactions is increasingly being employed to develop comprehensive understanding of cellular pathways. In this chapter, we discussed biochemical methods and outlined protocols currently used to isolate protein complexes in *C. elegans*.

---

---

---

## X. Solutions and Media

### A. Worm Reagents

- *NGM agarose plates (100 mm and 60 mm)*

3 g NaCl

25 g agarose

2.5 g peptone

975 mL ddH<sub>2</sub>O

Autoclave 35 min and place in 55 °C water bath.

When cooled sterilely add:

1 mL cholesterol (5 mg/mL in EtOH)

1 mL 1 M CaCl<sub>2</sub>

1 mL 1 M MgSO<sub>4</sub>

25 mL 1 M KH<sub>2</sub>PO<sub>4</sub> (pH 6.0)

- *M9*
  - 10 g NaCl
  - 12 g Na<sub>2</sub>HPO<sub>4</sub>
  - 6 g KH<sub>2</sub>PO<sub>4</sub>
  - 0.5 g MgSO<sub>4</sub>·7H<sub>2</sub>O
  - Add ddH<sub>2</sub>O to 2 L and autoclave 35 min.
- *2 M NaOH*
- *Bleach*
  - Fischer Cat. #SS290 (parafilm when not in use to minimize exposure to air and store at 4 °C)
- *OP-50*

## B. Large-Scale Liquid Culture

- *60% (w/v) sucrose* in *M9*
- *0.1 M NaCl*
- *OP-50-1*
  - streptomycin resistant *E. coli* strain
- *LB (2 x)*
  - 100 g of LB
  - 1800 mL ddH<sub>2</sub>O
  - Autoclave 35 min
- *LB + 50 µg/mL streptomycin*
- *S Basal (1 L)*
  - 5.9 g NaCl
  - 50 mL of 1 M KH<sub>2</sub>PO<sub>4</sub> (pH 6.0)
  - Add ddH<sub>2</sub>O to 1 L
  - Split into two 500 mL bottles; to each bottle add 0.5 mL cholesterol (5 mg/mL in EtOH; should form a light cloudy precipitate).
  - Autoclave and store at RT.
- *Trace Metals Solution*

Disodium EDTA	1.86 g	(5 mM)
FeSO <sub>4</sub> ·7H <sub>2</sub> O	0.69 g	(2.5 mM)
MnCl <sub>2</sub> ·4H <sub>2</sub> O	0.2 g	(1 mM)
ZnSO <sub>4</sub> ·7H <sub>2</sub> O	0.29 g	(1 mM)
CuSO <sub>4</sub> ·5H <sub>2</sub> O	0.025 g	(0.1 mM)

Dissolve in 1 L water; filter sterilize, aliquot into 50 mL conicals, and store in dark.

- *1 M potassium citrate (pH 6.0)*

268.8 g tripotassium citrate

26.3 g citric acid monohydrate

Add ddH<sub>2</sub>O to 900 mL, adjust pH to 6.0 using 10 N KOH and add ddH<sub>2</sub>O to 1 L. Autoclave and store at RT.

- *1 M MgSO<sub>4</sub>*
- *1 M CaCl<sub>2</sub>*
- *Complete S Basal*

Working under sterile conditions in the laminar flow hood add the following to 500 mL S Basal:

5 mL 1 M potassium citrate (pH 6.0)

5 mL Trace Metals Solution

1.5 mL 1 M MgSO<sub>4</sub>

1.5 mL 1 M CaCl<sub>2</sub>

### C. Isolation of Nuclei from Embryos

- *30% (w/v) sucrose in Nuclei Buffer*
- *Embryo Buffer*

25 mM Hepes KOH (pH 7.6)

118 mM NaCl

48 mM KCl

2 mM CaCl<sub>2</sub>

2 mM MgCl<sub>2</sub>

- *Nuclei Buffer*

10 mM Tris:HCl (pH 8.0)

80 mM KCl

2 mM EDTA

0.75 mM spermidine

0.3 mM spermine

Spermine and spermidine are prepared as 1 M stocks in ddH<sub>2</sub>O and stored at – 20 °C. Add immediately before use.

Add 1 tablet Complete EDTA-Free Protease Inhibitor Cocktail Tablets (Roche Applied Science, Cat. #1873580) per 10 mL buffer just prior to use.

- *10% (w/v) digitonin*

(Sigma, Cat. #37006) stock solution is prepared in ddH<sub>2</sub>O by boiling and filtering. Store aliquots at – 20 °C. Briefly boil, chill on ice, and add to Nuclei Buffer immediately before use.

- *Chitinase*

Dilute in Embryo Buffer at 2 U/mL and store aliquots at  $-20\text{ }^{\circ}\text{C}$ . The efficiency of chitinase treatment varies and the optimal length of digestion should be re-determined for each new batch.

- *Hoechst 33342*

## D. Single-step Immunoprecipitation

### Antibody coupling

- *10–50  $\mu\text{g}$  of affinity-purified antibody*
- *Affi-Prep Protein A beads (Biorad Cat. #156-0006)*
- *PBST (PBS + 0.1% Tween-20)*
- *1 M sodium borate (pH 9.0)*

Dissolve 61.8 g of boric acid in 800 mL of ddH<sub>2</sub>O and adjust the pH to 9.0 with NaOH pellets. Bring volume to 1 L with ddH<sub>2</sub>O and filter to sterilize. Store at RT.

- *0.2 M sodium borate (pH 9.0)*

Dilute from a stock of 1 M sodium borate (pH 9.0).

- *220 mM DMP*

(Sigma Cat. #D8388; FW 259.2; stored in a desiccated box at  $-20\text{ }^{\circ}\text{C}$ ) Let bottle sit tightly closed at RT for 20 min before opening. Weigh out DMP and leave dry until just before use. Resuspend in appropriate volume of 0.2 M sodium borate (pH 9.0) and add immediately to the bead suspension (e.g., for 34 mg DMP add 596  $\mu\text{L}$  sodium borate).

- *0.2 M ethanolamine, 0.2 M NaCl (pH 8.5)*

Dissolve 12.2 g of ethanolamine and 11.7 g of NaCl in ddH<sub>2</sub>O, adjust the pH to 8.5 with HCl. Add ddH<sub>2</sub>O to 1 L and filter to sterilize. Store at RT.

### Immunoprecipitation

- *100 mM glycine (pH 2.6)*

Dissolved 7.5 g glycine in 800 mL of ddH<sub>2</sub>O and adjust pH to 2.6 with HCl. Add ddH<sub>2</sub>O to 1 L. Filter to sterilize and store at  $4\text{ }^{\circ}\text{C}$ .

- *Lysis Buffer*

50 mM HEPES (pH 7.4)

1 mM EGTA

1 mM MgCl<sub>2</sub>  
100 mM KCl  
10% glycerol  
0.05% NP-40

DTT goes off with time, so add it to the buffer just before the experiment.

Just prior to use add one tablet Complete EDTA-Free Protease Inhibitor Cocktail Tablets (Roche Applied Science, #1873580) to 12 mL Lysis Buffer. Use Lysis Buffer with 300 mM KCL as indicated in text.

- *3x Sample Buffer*

6% SDS  
240 mM Tris (pH 6.8)  
30% Glycerol  
~0.04% (w/v) Bromophenol blue  
Add 50 µL 100% 2-ME to 1 mL just before use.

- *1 M DTT*

Dissolve 7.7 g DTT in 50 mL sterile ddH<sub>2</sub>O. Store 1 mL aliquots at -20 °C.

- *Pre-urea Wash Buffer*

50 mM Tris (pH 8.5)  
1 mM EGTA  
75 mM KCl  
Filter to sterilize and store at RT.

- *Urea Elution Buffer*

50 mM Tris (pH 8.5)  
8 M urea (Invitrogen Cat. # 15505-035)  
Store at RT and make fresh on the day of use.

- *2 M Tris (pH 8.5)*

Dissolve 121.1 g Tris base in 800 mL ddH<sub>2</sub>O and adjust pH to 8.5 with HCl. Add ddH<sub>2</sub>O to 1 L. Sterilize by autoclaving and store at RT.

- 100% TCA (Sigma Aldrich Cat. # T0699100 mL)

## E. Tandem Affinity Purification using a LAP Tag

- *GFP antibody*
- *TEV protease 6His-TEV protease* (1 mg/mL stock), purified using nickel-nitrilotriacetic acid (Ni-NTA) agarose and gel filtration (Kapust *et al.*, 2001; Parks *et al.*, 1995), or available from Invitrogen (Cat. #12575-015)
- *S protein agarose* (Novagen Cat. #69704)

---

---

---

## XI. Equipment

- 2.8 L Fernbach flasks
- microcentrifuge
- 0.5 mL and 1.5 mL microcentrifuge tubes
- 1 L centrifuge bottles
- 1.5 mL screw-cap tubes
- 15 mL and 50 mL conicals
- Dissecting microscope
- Graduated cylinders
- Mortar and pestle
- Ultracentrifuge with TLA100.3 rotor (Beckman, #349622)
- Sonicator with pulse capacity (e.g., Branson Digital Sonifier)
- Wheaton dounce homogenizer with pestle B
- Cooling shaker (Kuhner Shaker)
- Bransonic Ultrasonic Cleaner 3510
- Liquid nitrogen dewar flask

### Acknowledgments

We acknowledge support of grants from the NIH to A.D. (GM074215) and K.O. (GM074207), and by funding from the Ludwig Institute for Cancer Research to A.D. and K.O. This work is supported by the National Center for Research Resources of the National Institutes of Health by a grant to T.N.D. entitled “Comprehensive Biology: Exploiting the Yeast Genome” (P41 RR1 1823). E.Z. is supported by a research fellowship from the DFG (ZA619/1-1).

### References

- Audhya, A., and Desai, A. (2008). Proteomics in *Caenorhabditis elegans*. *Brief Funct. Genomic Proteomic* **7**, 205–210.
- Audhya, A., Hyndman, F., McLeod, I. X., Maddox, A. S., Yates III, J. R., Desai, A., and Oegema, K. (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*. *J. Cell. Biol.* **171**, 267–279.
- Beanan, M. J., and Strome, S. (1992). Characterization of a germ-line proliferation mutation in *C. elegans*. *Development* **116**, 755–766.
- Cassada, R., Isnenghi, E., Culotti, M., and von Ehrenstein, G. (1981). Genetic analysis of temperature-sensitive embryogenesis mutants in *Caenorhabditis elegans*. *Dev. Biol.* **84**, 193–205.
- Cheeseman, I. M., and Desai, A. (2005). A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci. STKE* **2005**, 11.
- Cheeseman, I. M., Niessen, S., Anderson, S., Hyndman, F., Yates III, J. R., Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18**, 2255–2268.
- Chowdhury, P. S. (2003). DNA immunization as a means to generate antibodies to proteins. *Methods Mol. Biol.* **207**, 57–62.
- Cinar, H., Keles, S., and Jin, Y. (2005). Expression profiling of GABAergic motor neurons in *Caenorhabditis elegans*. *Curr. Biol.* **15**, 340–346.

- Colosimo, M. E., Brown, A., Mukhopadhyay, S., Gabel, C., Lanjuin, A. E., Samuel, A. D., and Sengupta, P. (2004). Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types. *Curr. Biol.* **14**, 2245–2251.
- Cravatt, B. F., Simon, G. M., and Yates III, J. R. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991–1000.
- Dammermann, A., Maddox, P. S., Desai, A., and Oegema, K. (2008). SAS-4 is recruited to a dynamic structure in newly forming centrioles that is stabilized by the gamma-tubulin-mediated addition of centriolar microtubules. *J. Cell. Biol.* **180**, 771–785.
- Dammermann, A., Pemble, H., Mitchell, B. J., McLeod, I., Yates, J. R., Kintner, C., Desai, A. B., and Oegema, K. (2009). The hydrolethalus syndrome protein HYLS-1 links core centriole structure to cilia formation. *Genes Dev.* **23**, 2046–2059.
- Desai, A., Rybina, S., Muller-Reichert, T., Shevchenko, A., Hyman, A., and Oegema, K. (2003). KNL-1 directs assembly of the microtubule-binding interface of the kinetochore in *C. elegans*. *Genes Dev.* **17**, 2421–2435.
- Duchaine, T. F., Wohlschlegel, J. A., Kennedy, S., Bei, Y., Conte Jr, D., Pang, K., Brownell, D. R., Harding, S., Mitani, S., and Ruvkun, G., *et al.* (2006). Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124**, 343–354.
- Edgar, L. G. (1995). Blastomere culture and analysis. *Methods Cell Biol.* **48**, 303–321.
- Fernandez, A. G., Gunsalus, K. C., Huang, J., Chuang, L. S., Ying, N., Liang, H. L., Tang, C., Schetter, A. J., Zegar, C., and Rual, J. F., *et al.* (2005). New genes with roles in the *C. elegans* embryo revealed using RNAi of ovary-enriched ORFeome clones. *Genome Res.* **15**, 250–259.
- Fernandez, A. G., Mis, E. K., Bargmann, B. O., Birnbaum, K. D., and Piano, F. (2010). Automated sorting of live *C. elegans* using laFACS. *Nat. Methods* **7**, 417–418.
- Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383.
- Gassmann, R., Essex, A., Hu, J. S., Maddox, P. S., Motegi, F., Sugimoto, A., O'Rourke, S. M., Bowerman, B., McLeod, I., and Yates, J. R., *et al.* (2008). A new mechanism controlling kinetochore-microtubule interactions revealed by comparison of two dynein-targeting components: SPDL-1 and the Rod/Zwilch/Zw10 complex. *Genes Dev.* **22**, 2385–2399.
- Golden, A., Sadler, P. L., Wallenfang, M. R., Schumacher, J. M., Hamill, D. R., Bates, G., Bowerman, B., Seydoux, G., and Shakes, D. C. (2000). Metaphase to anaphase (mat) transition-defective mutants in *Caenorhabditis elegans*. *J. Cell. Biol.* **151**, 1469–1482.
- Gottschalk, A., Almedom, R. B., Schedletzky, T., Anderson, S. D., Yates III, J. R., and Schafer, W. R. (2005). Identification and characterization of novel nicotinic receptor-associated proteins in *Caenorhabditis elegans*. *EMBO J.* **24**, 2566–2578.
- Green, R. A., Audhya, A., Pozniakovsky, A., Dammermann, A., Pemble, H., Monen, J., Portier, N., Hyman, A., Desai, A., and Oegema, K. (2008). Expression and imaging of fluorescent proteins in the *C. elegans* gonad and early embryo. *Methods Cell Biol.* **85**, 179–218.
- Green, R. A., Kao, H. L., Audhya, A., Arur, S., Mayers, J. R., Fridolfsson, H. N., Schulman, M., Schloissnig, S., Niessen, S., and Laband, K., *et al.* (2011). A high-resolution *C. elegans* essential gene network based on phenotypic profiling of a complex tissue. *Cell* **145**, 470–482.
- Harlow, E. L., David, J., Harlow, E., and Lane, D. (1999). *Using Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- Hope, I. A. (1999). *C. elegans: A Practical Approach*. Oxford University Press, Oxford & New York.
- Huang, L. S., and Sternberg, P. W. (2006). Genetic dissection of developmental pathways. *WormBook* **14**, 1–19.
- Hubner, N. C., Bird, A. W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals *in vivo* protein interactions. *J. Cell Biol.* **189**, 739–754.

- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., and Sohrmann, M., *et al.* (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237.
- Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., and Waugh, D. S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14**, 993–1000.
- Kelly, W. G., and Fire, A. (1998). Chromatin silencing and the maintenance of a functional germline in *Caenorhabditis elegans*. *Development* **125**, 2451–2456.
- McNally, K., Audhya, A., Oegema, K., and McNally, F. J. (2006). Katanin controls mitotic and meiotic spindle length. *J. Cell Biol.* **175**, 881–891.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Moresco, J. J., Carvalho, P. C., and Yates III, J. R. (2010). Identifying components of protein complexes in *C. elegans* using co-immunoprecipitation and mass spectrometry. *J. Proteomics* **73**(11), 2198–2204.
- Parks, T. D., Howard, E. D., Wolpert, T. J., Arp, D. J., and Dougherty, W. G. (1995). Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology* **210**, 194–201.
- Piano, F., Schetter, A. J., Mangone, M., Stein, L., and Kempthues, K. J. (2000). RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. *Curr. Biol.* **10**, 1619–1622.
- Piano, F., Schetter, A. J., Morton, D. G., Gunsalus, K. C., Reinke, V., Kim, S. K., and Kempthues, K. J. (2002). Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Curr. Biol.* **12**, 1959–1964.
- Polanowska, J., Martin, J. S., Fisher, R., Scopa, T., Rae, I., and Boulton, S. J. (2004). Tandem immunofluorescence purification of protein complexes from *Caenorhabditis elegans*. *Biotechniques* **36**, 778–780.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032.
- Schaffer, U., Schlosser, A., Muller, K. M., Schafer, A., Katava, N., Baumeister, R., and Schulze, E. (2010). SnAvi – a new tandem tag for high-affinity protein-complex purification. *Nucleic Acids Res.* **38**, e91.
- Seydoux, G., and Strome, S. (1999). Launching the germline in *Caenorhabditis elegans*: regulation of gene expression in early germ cells. *Development* **126**, 3275–3283.
- Sonnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A. M., Artelt, J., Bettencourt, P., and Cassin, E., *et al.* (2005). Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
- Trinkle-Mulcahy, L., Boulton, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., and Leonhardt, H., *et al.* (2008). Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* **183**, 223–239.
- Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009). Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **11**, 49–79.
- Zhang, Y., Ma, C., Delohery, T., Nasipak, B., Foat, B. C., Bounoutas, A., Bussemaker, H. J., Kim, S. K., and Chalfie, M. (2002). Identification of genes expressed in *C. elegans* touch receptor neurons. *Nature* **418**, 331–335.

---

---

**PART III**

Development



---

---

## CHAPTER 12

# Cell Identification and Cell Lineage Analysis

**Claudiu A. Giurumescu and Andrew D. Chisholm**

Section of Cell and Developmental Biology, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA

---

- Abstract
- I. Introduction
- II. Rationale
  - A. Analysis of Mutant Phenotypes
  - B. Cell Division Pattern as the Focus of Interest
  - C. Cellular Patterns of Gene Expression
  - D. Cell Killing by Laser Ablation
  - E. Genetic Mosaic Analysis
  - F. Comparative Developmental Biology
- III. Resources
- IV. Nomenclature and Conventions
- V. Cell Identification and Lineage Analysis
  - A. Embryonic Cell Identification and Lineage Analysis
  - B. Postembryonic Cell Identification and Lineage Analysis
- VI. Materials, Methods, and Protocols
  - A. Protocol 1: Analysis of Embryonic Cell Lineages
  - B. Protocol 2: Post-Embryonic Cell-Lineage Analysis
- VII. Discussion
  - Acknowledgments
  - References

---

---

### Abstract

*Caenorhabditis elegans* is uniquely suited to the analysis of cell lineage patterns. *C. elegans* has a small number of somatic cells whose position and morphology are almost invariant from animal to animal. Because *C. elegans* is virtually transparent, cells can be identified in live animals using a simple bright-field microscopy technique, Nomarski differential interference contrast

(DIC), or by expression of transgenic fluorescent reporter genes. The small size and rapid development of *C. elegans* mean that animals can develop while under continuous observation, allowing cell lineages to be analyzed throughout embryonic and postembryonic development. Embryonic cell lineages can also be traced semiautomatically using timelapse imaging of GFP-labeled nuclei. Analysis of mutant cell lineages remains important for defining the roles of developmental control genes.

---

---

---

## I. Introduction

Cell lineage analysis refers to the tracing of cellular genealogies by following cell divisions and migrations over time, beginning with specific progenitor cells and ending with their postmitotic descendants. The development of almost all metazoan animals can in principle be described as a lineage tree whose origin is the single-cell zygote. However, the variability of normal development means that cell lineage relationships can in general only be described in probabilistic terms. In contrast, for some animal groups, including nematodes, molluscs, and tunicates, the pattern of cell divisions throughout development is highly invariant between individuals. In such animals, the invariant lineage constitutes a complete fate map of development with single-cell resolution.

The first descriptions of nematode cell lineages began in the late 19th century and were based on a series of fixed specimens. These studies established that the pattern of embryonic cell divisions was virtually invariant from animal to animal. In some cases, the cell lineage was thought to generate a fixed number of cells in the adult (“cell constancy” or eutely), or at least in certain tissues (“partial constancy”) (van Cleave, 1932). However, it was not until the development of Nomarski DIC microscopy in the late 1960s (Allen *et al.*, 1969; Padawer, 1968) that it became feasible to observe cell divisions in live animals.

Using Nomarski DIC microscopy of live animals the complete cell lineage of *C. elegans* from zygote to adult was delineated in a series of classic studies, culminating in the complete description of the embryonic cell lineage in 1983 (Sulston *et al.*, 1983). All these descriptions were based on direct observation of live animals, without significant use of recording technology. Since then, cell lineages have been traced in over ten other nematode species (see Table I). The *C. elegans* “lineage papers” (Table II) remain an essential resource for learning cell identification and lineage analysis. For historical accounts of the early days of lineage analysis see Horvitz and Sulston (1990) and John Sulston’s Nobel Lecture (Sulston, 2003).

With the advent of green fluorescent proteins (GFP) in the early 1990s (Chalfie *et al.*, 1994), cell identification entered a new phase. Now specific cells or cell types could be identified more rapidly, without the need for meticulous drawing out of cells and their positions. Expression of a fluorescent marker provided an unambiguous measure of cell fate. For example, cells could be identified even

**Table I**

Cell-lineage analysis in other nematode species

Species	Lineages studied	Reference
<i>Panagrellus redivivus</i>	Postembryonic lineages	(Sternberg and Horvitz, 1981, 1982)
<i>Turbatrix aceti</i>	Early embryo	(Sulston <i>et al.</i> , 1983)
<i>Mesorhabditis</i> etc.	Vulval lineages	(Sommer <i>et al.</i> , 1994)
<i>Pristionchus pacificus</i>	Vulval lineages	(Sommer, 1997)
<i>Acrobeloides</i>		(Wiegner and Schierenberg, 1998)
<i>Oscheius tipulae</i>	Vulval cell lineage	(Delattre and Felix, 2001)
<i>Pellioditis marina</i>	Complete embryonic lineage	(Houthoofd <i>et al.</i> , 2003)
<i>Halicephalobus gingivalis</i>		(Houthoofd and Borgonie, 2007)
<i>Rhabditophanes</i>	Embryonic lineage	(Houthoofd <i>et al.</i> , 2008)
<i>C. briggsae</i>	Embryonic lineage; automated lineage tracing	(Zhao <i>et al.</i> , 2008)
<i>Diploscapter coronatus</i>	Embryonic lineage	(Lahl <i>et al.</i> , 2009)
<i>Romanomermis culicivorax</i>	Embryonic lineage	(Schulze and Schierenberg, 2008, 2009)

**Table II**Key publications describing *C. elegans* lineages

Lineages	Reference
Ventral nerve cord	(Sulston, 1976)
Postembryonic nongonadal lineages (hermaphrodite)	(Sulston and Horvitz, 1977)
Gonadal lineages (both sexes)	(Kimble and Hirsh, 1979)
Postembryonic nongonadal lineages (male)	(Sulston <i>et al.</i> , 1980)
Embryonic lineage	(Deppe <i>et al.</i> , 1978; Sulston <i>et al.</i> , 1983)

when misplaced in aberrant locations as a result of a cell migration defect. Most importantly, whereas under DIC observation only cell nuclei are typically resolved, GFP markers can be used to visualize the entire cell, or specific subcellular compartments. The ability to see axons, muscle arms, and other structures in live animals opened up whole new areas of analysis.

GFP transgenic markers have in many cases replaced DIC for cell identification. Nevertheless there are still reasons to learn the DIC anatomy. First, transgenic markers are not available for all cell types or subsets of cell types. Second, care is needed to ensure that the GFP marker (often a high-copy number transgene) does not itself interfere with cell differentiation. Issues of photobleaching or phototoxicity often limit the amount of observation possible, although this has been to some extent overcome in the automated analysis of embryonic cell divisions. Finally, analyzing a number of markers can involve considerable strain construction, work that can be avoided if the cells can be identified by DIC.

There has been rapid recent progress in automation of cell identification and lineage analysis. Cell lineage analysis in embryos can be partly automated as discussed in detail below. Automatic cell lineage analysis in larval stages has so far not been possible, largely due to the difficulty in immobilizing larvae in a way that allows normal development. However, digital atlases of cell positions at defined stages can be generated, allowing gene expression patterns to be mapped semiautomatically (Long *et al.*, 2009). At present such atlases represent ~65% of the nuclei in the L1 stage. However, many neuronal nuclei are too closely spaced to be reliably identified by automated analysis. Thus, to identify specific neuronal expression patterns a knowledge of the anatomy remains indispensable.

---

---

---

## II. Rationale

### A. Analysis of Mutant Phenotypes

One of the most frequent goals in cell-lineage analysis is to address the developmental basis of a specific phenotype, whether caused by mutation, RNAi, or some other perturbation. For example, using cell type specific GFP reporters it is straightforward to screen for mutants affecting the number of cells that express a given reporter (Doitsidou *et al.*, 2008; Kanamori *et al.*, 2008). A change in the number of expressing cells could have a number of causes and lineage analysis can resolve these possibilities. For example, do excess GFP-expressing cells arise from ectopic expression of the reporter or from an overproliferation of specific precursors? Does failure to generate a given cell type reflect a cell fate transformation or an earlier defect in the lineage?

### B. Cell Division Pattern as the Focus of Interest

In some cases, the pattern of cell divisions itself is the focus of interest, especially where no other molecular markers are available. For example, the role of the Wnt ligand LIN-44 in cell polarity was deduced from its effects on the polarity of certain cell divisions in the male tail (Herman and Horvitz, 1994). The stage specificity of cell-division patterns was critical in inferring the genetic control of developmental timing in larval development (Ambros and Horvitz, 1984). The regulative ability of certain tissues to undergo compensatory growth after damage was studied using cell-lineage analysis (Sulston and White, 1980).

Stem-cell-like division is inherently polarized. The stem-cell-like behavior of larval seam cells has been extensively analyzed by direct lineage analysis (e.g. Nimmo *et al.*, 2005). Analysis of cell-lineage mutants has also been important in understanding the genetic basis of cell cycle control (e.g. Kostić and Roy, 2002; Fukuyama *et al.*, 2003).

### C. Cellular Patterns of Gene Expression

A common goal in cell identification is to define the cellular expression patterns of reporter genes. Some genes are expressed in relatively clear-cut patterns, for example, intestinal cells, all body wall muscles, all GABAergic neurons, and so on (cf. Chapter by Yan and Jin?). It is essential to learn to correlate such simple expression patterns first before attempting more complex patterns.

The nervous system poses the most daunting challenge for identification of gene-expression patterns. This is especially so in the large anterior ganglia, each of which contains 10–20 closely packed nuclei. However, the relative positions of most neuronal nuclei are fairly reproducible and can be learnt by reference to the maps in the L1 stage (Sulston *et al.*, 1983). Maps of these nuclei in the adult are based on serial section EM reconstruction (White *et al.*, 1986) and can be found in WormAtlas. Some neuronal nuclei display natural variability in location, and so cannot be conclusively identified based on position. Fortunately, identification of neurons is often made considerably easier by the distinctive disposition of the axons and dendrites of individual cell types.

### D. Cell Killing by Laser Ablation

Cell ablation has been an important technique to define the developmental and physiological functions of cells (Bargmann and Avery, 1995). Individual cells can be killed with a laser microbeam focused on the cell nucleus. This depends on accurate cell identification, for which both DIC and GFP are now used. Ablation during development can be used to test the extent of replacement regulation by other cells. Physiological functions of cells can be addressed by ablation unless they are subject to replacement or compensation.

### E. Genetic Mosaic Analysis

In *C. elegans*, genetic mosaic analysis relies on spontaneous loss of unstable extrachromosomal arrays or chromosomal duplications during development. By identifying the pattern of cells in which the array or duplication has been lost, the “loss point” in the early lineage can be deduced. Such patterns are generally examined in late larval stages, that is, after the cell divisions are largely complete.

### F. Comparative Developmental Biology

Cell lineages can be traced using Nomarski DIC in any optically transparent organism that can develop under continuous observation. As a result, direct cell-lineage analysis has now been undertaken in over a dozen different nematode species (Table I). Embryonic lineages have now been traced in several species. Although initial studies suggested a high degree of conservation in early embryonic lineages (Sulston *et al.*, 1983), subsequent studies of other species have revealed a remarkable

diversity in the patterns of cell division within nematode embryos. Studies of vulval cell lineages in several species have been critical to our understanding of evolution of developmental mechanisms. As transgenic tools are now being developed in other nematode species, their use in automated analysis is likely to increase; the embryonic cell lineage of *C. briggsae* has already been followed using automated histone-GFP lineage tracing (Zhao *et al.*, 2008).

---

---

### III. Resources

The descriptions of cell lineages from the late 1970s remain the definitive descriptions of the cellular anatomy (Table II). In learning the anatomy an important initial goal is to compare one's own drawings with the diagrams in the following papers. In particular, the description by Sulston *et al.* (1983) remains the best resource for learning embryonic anatomy; an “embryo” section of WormAtlas is currently under construction.

WormAtlas ([www.wormatlas.org](http://www.wormatlas.org)) and the *C. elegans* Atlas book (Hall and Altun, 2008) are invaluable for understanding adult anatomy and for correlating cellular anatomy with electron micrographs. The web site contains a small section on cell identification. A good online guide to cell identification is in Wormbook (Yochem, 2006), with plentiful Nomarski DIC images of “landmark” cells. This is an important addition to the original lineage papers. However, in our experience the only way to successfully learn cell identification is to sit at the microscope and draw what one sees.

---

---

### IV. Nomenclature and Conventions

The nomenclature for cells was set out by Sulston and Horvitz (1977) and systematized by Sulston *et al.* (1983). Every cell in *C. elegans* can be named according to its ancestry, for example, ABpla. Terminally differentiated cells also have “functional” names that are either semiarbitrary (e.g., ASEL) or descriptive of terminal fate (hyp 7). For example, the cell ABalppppppaa is the neuron ASEL.

Embryonic cells are named beginning with one of the five early embryonic “founder cells”: AB, E, MS, C, D. The cells P<sub>0</sub> through P<sub>4</sub> denote the zygote and the precursors of the germ line, and should not be confused with the postembryonic blast cells P1–P12. Cells that go on to divide in postembryonic stages are renamed with a blast cell name (e.g., ABplapapaaa=QL), and their progeny named according to similar rules.

The suffixes in lineage names refer to the approximate orientation of the cell division relative to the overall axes of the embryo or larva: anterior/posterior, dorsal/ventral, left/right. Almost all cell divisions in *C. elegans* have a clear anterior–posterior orientation; indeed only ~8 embryonic cell divisions are predominantly in the transverse (left–right) axis. Cells are named according to the relative position of the daughters at the time of division, even if the daughters subsequently change



(~240 min post first cleavage) cellular differentiation begins. Cells can be identified based on position relative to landmarks such as the first apoptotic cell deaths; maps of cell nuclei at 270 nuclei are in Figure 7 of Sulston *et al.* (1983).

Cell and nuclear identification is more or less straightforward using the available maps up to the 1.75-fold stage when body wall muscle contractions begin. By the twofold stage the rolling activity of the embryo severely hampers attempts to identify cells in live specimens. To prevent this rapid movement the microscope objective can be transiently and reversibly cooled (Sulston *et al.*, 1983). Overall the last 3 h of embryonic development remain relatively little studied.

The cell lineage reported in Sulston *et al.* (1983) was a composite built up from partial lineages of hundreds of embryos. With the improvement of imaging and storage technology by the late 1980s it became possible to record the complete development of single embryos using timelapse imaging in multiple focal planes. Such “four dimensional” (4D) microscopy was developed by J. G. White and used for studies of early *C. elegans* embryos (Hird and White, 1993). Several groups developed software for 4D acquisition and playback (Fire, 1994; Thomas *et al.*, 1996). Commercial imaging software suites such as Improvion or Amira include options for automated 4D capture, as do the software suites for most commercial compound microscopes. Finally the Open Source software suite Micro-Manager (<http://valelab.ucsf.edu/~MM/MMwiki/index.php/Micro-Manager>) can be used for automated 4D acquisition.

Cell-lineage tracing from DIC 4D movies requires manual identification of nuclei at each time point. Software such as SIMI Biocell (Schnabel *et al.*, 1997) (Table III)

**Table III**

Comparison of some methods for manual and semiautomatic embryonic lineage recording

Software name	Data	Maximum nuclei tracked	Method	Time required	Error rate	Reference
SIMI Biocell	DIC	385 cells full/114 cells up to 1.5 fold	Manual	N/A	N/A	(Schnabel <i>et al.</i> , 1997)
Angler	DIC	end of gastrulation/ 300 cells	Manual	N/A	N/A	(Martinelli <i>et al.</i> , 1997)
N/A	DIC	24 cells	Automated	N/A	N/A	(Hamahashi <i>et al.</i> , 2005)
StarryNite	GFP	350 cells	Automated	8 h up to 350 cells	1–3% up to 194 cells, larger up to 350 cells	(Bao <i>et al.</i> , 2006)
Endrov	DIC/GFP	150 cells full/partial later stage	Manual	N/A	N/A	(Hench <i>et al.</i> , 2009)
NucleiTracker4D	GFP	566 live/ 669 total cells	Semiautomated	For GFP confocal = 4 h up to 350 cells 8–16 h to 525 cells	Up to 3% to 350 cells, up to 15% to 525 cells (at each time-point).	CAG & ADC, unpublished

allows lineages to be generated in a straightforward manner through most of embryogenesis. The image degradation in deeper DIC focal planes limits complete lineaging of single embryos from DIC. Although computer image analysis can identify nuclei in DIC images (Hamahashi *et al.*, 2005) the low contrast of DIC images makes this computationally challenging.

## 2. Semiautomated Lineage Analysis

With the advent of histone-GFP reporters marking all embryonic nuclei (Ooi *et al.*, 2006), imaging and lineage analysis of embryo development could now be studied using the laser confocal microscope. The high fluorescence contrast of nuclear-localized GFP allows identification of nuclei by computer image segmentation in a more reliable fashion than from DIC data. Although photobleaching and phototoxicity limit the image quality, it is possible to image histone-GFP throughout embryonic development. The histone-GFP itself is presumably synthesized anew during each cell cycle, partly compensating for photobleaching effects. The Waterston lab implemented the first histone::GFP imaging platform for following embryo development through gastrulation and morphogenesis (Murray *et al.*, 2006). To track moving and dividing nuclei the Starrynite software was developed (Bao *et al.*, 2006). Additional visualization software allows curation of the tracking data and extraction of lineage information (Boyle *et al.*, 2006).

The accuracy of the original Starrynite software declined in embryos with >350 nuclei. To allow nuclear tracking in later embryos several approaches are being tested. One approach is to optimize the segmentation algorithm for images of optically sectioned nuclei (Santella *et al.*, 2010). Another approach, currently implemented in our laboratory (see below), is to curate nuclear identification at each time point, such that the automated tracking at the next time point  $t+1$  starts with corrected information. A second modification to the search algorithm is to constrain the search for a particular nucleus (or its daughters) at  $t+1$  only in the local neighborhood of its previous position at  $t$ . Rather than performing *de novo* segmentation, in this approach the maximum amount of information available at  $t$  is used for performing the segmentation and tracking for  $t+1$  (Table III).

## 3. Extent of Variation in the Wild-Type Lineage

The ability to completely track all nuclei in individual embryos has prompted further examination of the degree of variability in wild-type development. Automated lineage tracing has confirmed the high degree of invariance in the assignment of cell fates, with the known exceptions of the midline cell pairs mentioned above. Some cell-division axes in the C lineage display variability. Cell division times can vary by a factor of 10% between embryos (Sulston *et al.*, 1983), but within individual embryos the relative timing of cell divisions is highly consistent. The high degree of correlation of cell cycles within an individual embryo

suggests the existence of a “developmental clock” controlling the rate of embryogenesis; the overall standard deviation in this developmental clock has been estimated at 4.5% (Bao *et al.*, 2008).

Earlier lineage studies suggested that relative cell positions could show variability in midembryogenesis (Schnabel *et al.*, 1997), although such variability decreases by the premorphogenetic stages. More recent studies have suggested that some of this variability may result from the slight compression introduced when an embryo is mounted on an agar pad or between slide and coverslip using beads (Hench *et al.*, 2009). Compressed embryos display two stereotypical rotations. First, during gastrulation, the embryo turns from a left–right aspect to a dorsal–ventral aspect. Then, following epidermal enclosure, the embryo turns once again to display the left–right aspect and its comma shape. The increased variability in cell positions in such compressed embryos may reflect increased migration displacements in the flattened eggshell in conjunction with rotational movements. Unlike compressed embryos, freely mounted embryos attached to polylysine coated coverslips show less variability in cell positions (Hench *et al.*, 2009; Schnabel *et al.*, 2006) and do not display the typical left–right/dorsal–ventral rotations. However, the increased depth of the uncompressed embryo leads to a slight loss of optical quality, and many laboratories continue to use slightly compressed embryos for optimal imaging (see chapter by Hardin).

## B. Postembryonic Cell Identification and Lineage Analysis

Cell identification in larval and adult stages is facilitated by the increased separation of nuclei and the differentiation of cell types. However as development proceeds nuclei tend to have slightly less stereotyped positions. Accurate identification of cells and nuclei is also complicated by the tendency of worms to move out of the field of view; at present there is no anesthetic or physical restraint that is compatible with long-term development.

Most cell types are readily identified by position and nuclear morphology. Complex cell groups such as the anterior ganglia require practice and tracing of cell positions from multiple animals. To begin identifying cells it is essential to start with simple easily recognized stages and tissues such as the 12-cell stage of vulval development. A novel approach to identifying new expression patterns is to analyze their intersection with previously characterized patterns using “split GFP” (Zhang *et al.*, 2004).

Paralleling the automated lineaging efforts in the embryo, Long *et al.* (2009) have recently constructed a 3D atlas of nuclear positions in L1 larvae. Generating a standard 3D representation of the L1 larvae nuclei is instrumental for mapping gene expression patterns or high-throughput computer-controlled functional screens. Atlas building depends on (1) reliable identification of larval nuclei, (2) registration of multiple larval samples into the same standard representation, and (3) mapping of novel samples onto this standard representation. To achieve

these goals Long *et al.* used DAPI stained worms to mark all cell nuclei, followed by several image processing steps. First, the images of larvae are straightened to a rod shape (Peng *et al.*, 2008). Next, nuclei are automatically identified by adaptive thresholding and rule- and training-based segmentation. Nuclei can be validated and curated using the volume-object image annotation system (Peng *et al.*, 2009). In this way, 357 out of the 558 L1 larval nuclei could be faithfully identified.

Using fiduciary muscle nuclei GFP markers, Long *et al.* registered as many as 40 L1 larvae samples into a 3D standard representation. However, as few as 15 samples were sufficient to infer correct nuclear positions along the body. A comparison of cell positions in the atlas samples along the three axes of the body showed that nucleus-to-nucleus spatial relationships are invariant, especially among cells belonging to the same tissue. After building the standard representation, Long *et al.* developed an automated procedure for mapping and annotating novel samples, such as expression patterns, onto this reference (Liu *et al.*, 2009). Overall, their automated segmentation can identify nuclei in certain tissues with >80% accuracy. Improving accuracy and the ability to segment all of the 558 nuclei of larvae will entail using higher resolution microscopy methods like selective plane illumination or stimulated emission depletion.

---

---

---

## VI. Materials, Methods, and Protocols

A general protocol for mounting *C. elegans* on agar pads for live analysis is provided in the chapter by Shaham in Worm Methods ([http://www.wormbook.org/toc\\_wormmethods.html](http://www.wormbook.org/toc_wormmethods.html)); see also the methods appendix to the *C. elegans* I book.

### A. Protocol 1: Analysis of Embryonic Cell Lineages

#### 1. Mounting

Detailed protocols for mounting *C. elegans* embryos are provided in the chapter by Hardin. Traditionally *C. elegans* embryos have been mounted on agar pads with buffer and a coverslip. Although embryos are completely viable under such conditions, it is clear that this method compresses the egg and eggshell. An uncompressed egg mounted in an aqueous medium is 50  $\mu\text{m}$  long and 30  $\mu\text{m}$  in diameter (Deppe *et al.*, 1978). Also Ref. Blanchoud *et al.*, 2010; *Dev Dyn* 239: 3285–96. Embryos mounted on agar pads are compressed to a thickness of  $\sim 20$   $\mu\text{m}$  (Schnabel *et al.*, 1997). Mounting using spacer beads can also compress the embryo, depending on the bead size used. The lateral compression is helpful in reducing the number of optical sections needed for 4D lineage analysis and constraining the embryo to a fixed orientation for observation. However as mentioned in the text, compression may contribute to the variability in cell positions in early embryogenesis.

## 2. DIC Cell-Lineage Analysis and 4D Recording

Procedures for manual lineaging of embryos are described by Sulston *et al.* (1983). A number of software tools have been described over the years to allow time-lapse recording in multiple focal planes (4D recording) (see above). At present most commercial microscope vendors include 4D acquisition as an option. The minimal requirements are microscope equipped for Nomarski DIC optics, a motorized z-drive, a camera, and a computer workstation that controls the z-drive. A high-N.A. DIC objective (e.g., Zeiss Plan-Neofluar 100 $\times$ ) is essential for any lineage studies. Cell lineages can be traced manually from DIC 4D data sets. The software package SIMI Biocell is specifically designed to facilitate lineage construction from DIC 4D movies.

## 3. 4D Lineaging Using Histone-GFP

Procedures for 4D imaging of histone-GFP marked embryos are defined in Murray *et al.* (2006). Briefly, single transgenic HIS-72::GFP(*zuls178*) embryos are mounted between coverslips in 8  $\mu\text{L}$  of a mixture of 20  $\mu\text{m}$  polystyrene beads (Polysciences Inc., Warrington, PA) in 1% methylcellulose in M9 (15% v/v beads, 85% v/v 1% methylcellulose in M9). The coverslip sandwich can be flipped to display the desired late development dorsal or ventral aspect and then attached to a slide and sealed.

We use a Zeiss LSM510 confocal with a 30 mW Argon laser. We acquire confocal z-stacks of size  $64 \times 35 \times 30 \mu\text{m}^3$  with resolution of  $0.125 \times 0.125 \times 0.85 \mu\text{m}^3$ /voxel every minute for the first 300 min of development then every 2 min for the next 180 min. Two-color (GFP/mCherry) movies can be acquired to correlate cell-specific expression patterns with the ubiquitously expressed histone-GFP. Laser power, detector, and acquisition configurations are loaded through the MultiTime macro in the Zeiss LSM software.

Precise temperature control is extremely important to maintain embryo viability over prolonged periods of confocal imaging. Although embryos will survive 4D DIC imaging throughout embryogenesis at 25  $^{\circ}\text{C}$ , we find the upper limit for confocal imaging is 24  $^{\circ}\text{C}$ ; the viability of imaged embryos should be checked whenever 4D imaging is being set up for the first time. There are several options for control of specimen temperature; we have used a custom-designed aluminum casing for the objective. The casing is cooled or heated by a small Peltier element and a liquid cooling system designed for computer chips.

The analysis of 4D LSM data sets by Starrynite and Acetree is described in detail by (Murray *et al.*, 2006). We provide here a brief overview of our nuclear tracking approach (Giurumescu *et al.*, in preparation). We analyze 4D LSM data sets with a user interface that combines the automated tracking and user-selected curation. At the first time point (usually 4–6 nuclei) the user manually identifies nuclei and names them according to the canonical wild-type lineage. For subsequent time z-stacks, the software first performs an automatic segmentation and

tracking step using a minimal movement algorithm and local neighborhood search. Nuclei that do not satisfy the strict minimal movement condition (i.e., those that move less than their radius from  $t$  to  $t + 1$ , usually less than 5% of nuclei), are flagged for manual curation. Nuclear divisions are also curated manually. At each time point the correct set of nuclei is annotated, preventing the accumulation of annotation errors. Our software does not search for nuclear radii as an additional free parameter. Our initial manual lineaging confirmed the initial observation (Bao *et al.*, 2006) that all nuclei in each of the major sublineages (AB, C, D, E, MS, and P4) show distinct radii values that linearly decrease with each round of division. Hence, our software prescribes nuclear radii values to all nuclei depending on their position in the lineage. Using this semiautomated approach it is possible to lineage essentially all nuclei up to the 1.5-fold stage (566 live nuclei, 103 cell deaths).

### B. Protocol 2: Post-Embryonic Cell-Lineage Analysis

1. Worms to be lineaged must be in healthy, unstarved condition.
2. Prepare a standard slide mount agar pad (cf. Sulston and Hodgkin methods appendix in *C. elegans* I). The agar should have been freshly prepared or melted.
3. Using a drawn-out capillary and mouth pipette pick up the worm(s) to be lineaged in a few microliters of M9 or S basal. Deposit the larva onto the agar pad together with a small volume of buffer. [If you are very dextrous, it is possible to do this with a worm pick, but small larvae are very easily injured. We recommend the mouth pipette.] Remove excess buffer by wicking with lens paper.
4. Using lens paper, wipe clean a small coverslip (18×18 mm OK, 12×12 mm best but can be hard to find). Using a worm pick, smear a small amount of OP50 *E. coli* onto the center of the coverslip. Place the coverslip gently over the buffer + worm so that the bacterial blob is within a couple of mm of the worm.
5. After 1–5 min the worm should become active and head toward the bacteria and start browsing contentedly. Under optimal circumstances, the worm will continue eating for hours, with occasional bouts of movement. If the worm does not move or begin eating within 10 min of mounting, it may be damaged.
6. If the worm appears healthy, trim the agar around the coverslip with a razor blade. Seal the edges of the coverslip with immersion oil or vaseline. Some brands of immersion oil are toxic and can interfere with long-term observation. Vaseline works fine unless your worm swims into it or you get some on the objective.
7. Find the area of interest in the animal and draw out everything you see as often as you can, identifying nuclei by reference to the standard maps in the papers in. With practice, multiple animals can be lineaged at a time, depending on the complexity of the lineage being traced. It is usually best to keep one animal per slide to avoid confusion.

## 1. Troubleshooting

The joy of observing a well-behaving worm is balanced by the frustration of a badly behaved worm that persistently heads for the edge of the agar pad, only to end up in the immersion oil or vaseline. To avoid such frustration it is important to check your worm frequently (every 10 min) and to learn how to recover the worm intact from the slide mount. Practice sliding off the coverslip and getting the worm into a buffer-drop from which you can suck it back into your capillary.

If the worm stops moving and cells lose contrast, the animal may be dying, or it may be entering lethargus, the 1–2 h period of inactivity that precedes each molt. If the developmental stage makes the latter explanation unlikely, there may be too much bacteria under the slide, leading to hypoxia. The worm can be revived by removing it from the slide (slide off the coverslip and use mouth pipette + drawn-out capillary to retrieve the worm). Place the worm on an NGM agar plate to recover for a few minutes, then remount.

The microscope DIC optics should be optimized (Köhler illumination). A heat filter must be used to prevent specimen heating under the prolonged observation. Immersion oil should be used between the objective and coverslip and between slide and condenser top lens. Ensure there are no bubbles or debris in the agar pad; once a worm crawls into a bubble, it will not come out again.

---

---

---

## VII. Discussion

Cell-lineage analysis allows rigorous definition of cell ancestries and positions, mutant phenotypes, and gene expression patterns to single-cell resolution. The labor-intensive nature of lineage tracing from live samples has tended to limit its popularity. The recent development of automated lineage analysis promises to reduce the effort needed for early embryonic lineage studies, but lineage tracing in later embryos and in larvae remains a labor of love. Further computational advances may help to return lineage studies to the center of *C. elegans* developmental biology.

## Acknowledgments

The authors thank members of their lab for comments and the *C. elegans* community for their development and sharing of the techniques mentioned here. Our work on embryonic development has been supported by an NIH NRSA Award to CAG and NIH R01 GM54657 to ADC.

## References

- Allen, R. D., David, G. B., and Nomarski, G. (1969). The Zeiss-Nomarski differential interference equipment for transmitted-light microscopy. *Z Wiss Mikrosk* **69**, 193–221.
- Ambros, V., and Horvitz, H. R. (1984). Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226**, 409–416.
- Bao, Z., Murray, J. I., Boyle, T., Ooi, S. L., Sandel, M. J., and Waterston, R. H. (2006). Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2707–2712.

- Bao, Z., Zhao, Z., Boyle, T. J., Murray, J. I., and Waterston, R. H. (2008). Control of cell cycle timing during *C. elegans* embryogenesis. *Dev. Biol.* **318**, 65–72.
- Bargmann, C. I., and Avery, L. (1995). Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol.* **48**, 225–250.
- Boyle, T. J., Bao, Z., Murray, J. I., Araya, C. L., and Waterston, R. H. (2006). AceTree: A tool for visual analysis of *Caenorhabditis elegans* embryogenesis. *BMC Bioinform.* **7**, 275.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Delattre, M., and Felix, M. A. (2001). Polymorphism and evolution of vulval precursor cell lineages within two nematode genera, *Caenorhabditis* and *Oscheius*. *Curr. Biol.* **11**, 631–643.
- Deppe, U., Schierenberg, E., Cole, T., Krieg, C., Schmitt, D., Yoder, B., and von Ehrenstein, G. (1978). Cell lineages of the embryo of the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 376–380.
- Doitsidou, M., Flames, N., Lee, A. C., Boyanov, A., and Hobert, O. (2008). Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nat. Meth.* **5**, 869–872.
- Fire, A. (1994). A four-dimensional digital image archiving system for cell lineage tracing and retrospective embryology. *Comput. Appl. Biosci.* **10**, 443–447.
- Fukuyama, M., Gendreau, S. B., Derry, W. B., and Rothman, J. H. (2003). Essential embryonic roles of the CKI-1 cyclin-dependent kinase inhibitor in cell-cycle exit and morphogenesis in *C. elegans*. *Dev. Biol.* **260**(1), 273–286.
- Hall, D. H., and Altun, Z. F. (2008). *C. elegans Atlas*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Hamahashi, S., Onami, S., and Kitano, H. (2005). Detection of nuclei in 4D Nomarski DIC microscope images of early *Caenorhabditis elegans* embryos using local image entropy and object tracking. *BMC Bioinform.* **6**, 125.
- Hench, J., Henriksson, J., Luppert, M., and Burglin, T. R. (2009). Spatio-temporal reference model of *Caenorhabditis elegans* embryogenesis with cell contact maps. *Dev. Biol.* **333**, 1–13.
- Herman, M. A., and Horvitz, H. R. (1994). The *Caenorhabditis elegans* gene *lin-44* controls the polarity of asymmetric cell divisions. *Development* **120**, 1035–1047.
- Hird, S. N., and White, J. G. (1993). Cortical and cytoplasmic flow polarity in early embryonic cells of *Caenorhabditis elegans*. *J. Cell Biol.* **121**, 1343–1355.
- Horvitz, H. R., and Sulston, J. E. (1990). Joy of the worm. *Genetics* **126**, 287–292.
- Houthoofd, W., and Borgonie, G. (2007). The embryonic cell lineage of the nematode *Halickephalobus gingivalis* (Nematoda: Cephalobina: Panagrolaimoidea). *Nematology* **9**, 573–584.
- Houthoofd, W., Jacobsen, K., Mertens, C., Vangestel, S., Coomans, A., and Borgonie, G. (2003). Embryonic cell lineage of the marine nematode *Pellioiditis marina*. *Dev. Biol.* **258**, 57–69.
- Houthoofd, W., Willems, M., Jacobsen, K., Coomans, A., and Borgonie, G. (2008). The embryonic cell lineage of the nematode *Rhabditophanes* sp. *Int. J. Dev. Biol.* **52**, 963–967.
- Kanamori, T., Inoue, T., Sakamoto, T., Gengyo-Ando, K., Tsujimoto, M., Mitani, S., Sawa, H., Aoki, J., and Arai, H. (2008). Beta-catenin asymmetry is regulated by PLA1 and retrograde traffic in *C. elegans* stem cell divisions. *EMBO J.* **27**, 1647–1657.
- Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev. Biol.* **70**, 396–417.
- Kostić, I., and Roy, R. (2002). Organ-specific cell division abnormalities caused by mutation in a general cell cycle regulator in *C. elegans*. *Development* **129**(9), 2155–2165.
- Lahl, V., Schulze, J., and Schierenberg, E. (2009). Differences in embryonic pattern formation between *Caenorhabditis elegans* and its close parthenogenetic relative *Diploscapter coronatus*. *Int. J. Dev. Biol.* **53**, 507–515.
- Liu, X., Long, F., Peng, H., Aerni, S. J., Jiang, M., Sanchez-Blanco, A., Murray, J. I., Preston, E., Mericle, B., Batzoglou, S., Myers, E. W., and Kim, S. K. (2009). Analysis of cell fate from single-cell gene expression profiles in *C. elegans*. *Cell* **139**, 623–633.
- Long, F., Peng, H., Liu, X., Kim, S. K., and Myers, E. (2009). A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Meth.* **6**, 667–672.

- Martinelli, S. D., Brown, C. G., and Durbin, R. (1997). Gene expression and development databases for *C. elegans*. *Semin. Cell Dev. Biol.* **8**, 459–467.
- Murray, J. I., Bao, Z., Boyle, T. J., and Waterston, R. H. (2006). The lineaging of fluorescently-labeled *Caenorhabditis elegans* embryos with StarryNite and AceTree. *Nat. Protoc.* **1**, 1468–1476.
- Nimmo, R., Antebi, A., and Woollard, A. (2005). *mab-2* encodes RNT-1, a *C. elegans* Runx homologue essential for controlling cell proliferation in a stem cell-like developmental lineage. *Development* **132** (22), 5043–5054.
- Ooi, S. L., Priess, J. R., and Henikoff, S. (2006). Histone H3.3 variant dynamics in the germline of *Caenorhabditis elegans*. *PLoS Genet.* **2**, e97.
- Padawer, J. (1968). The Nomarski interference-contrast microscope. An experimental basis for image interpretation. *J. R. Microsc. Soc.* **88**, 305–349.
- Peng, H., Long, F., Liu, X., Kim, S. K., and Myers, E. W. (2008). Straightening *Caenorhabditis elegans* images. *Bioinformatics* **24**, 234–242.
- Peng, H., Long, F., and Myers, E. W. (2009). VANO: A volume-object image annotation system. *Bioinformatics* **25**, 695–697.
- Santella, A., Du, Z., Nowotschin, S., Hadjantonakis, A. K., and Bao, Z. (2010). A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D. *BMC Bioinform.* **11**, 580.
- Schnabel, R., Bischoff, M., Hintze, A., Schulz, A. K., Hejnol, A., Meinhardt, H., and Hutter, H. (2006). Global cell sorting in the *C. elegans* embryo defines a new mechanism for pattern formation. *Dev. Biol.* **294**, 418–431.
- Schnabel, R., Hutter, H., Moerman, D., and Schnabel, H. (1997). Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: Variability of development and regional specification. *Dev. Biol.* **184**, 234–265.
- Schulze, J., and Schierenberg, E. (2008). Cellular pattern formation, establishment of polarity and segregation of colored cytoplasm in embryos of the nematode *Romanomermis culicivorax*. *Dev. Biol.* **315**, 426–436.
- Schulze, J., and Schierenberg, E. (2009). Embryogenesis of *Romanomermis culicivorax*: An alternative way to construct a nematode. *Dev. Biol.* **334**, 10–21.
- Sommer, R. J. (1997). Evolutionary changes of developmental mechanisms in the absence of cell lineage alterations during vulva formation in the Diplogastriidae (Nematoda). *Development* **124**, 243–251.
- Sommer, R. J., Carta, L. K., and Sternberg, P. W. (1994). The evolution of cell lineage in nematodes. *Dev. Suppl.* 85–95.
- Sternberg, P. W., and Horvitz, H. R. (1981). Gonadal cell lineages of the nematode *Panagrellus redivivus* and implications for evolution by the modification of cell lineage. *Dev. Biol.* **88**, 147–166.
- Sternberg, P. W., and Horvitz, H. R. (1982). Postembryonic nongonadal cell lineages of the nematode *Panagrellus redivivus*: Description and comparison with those of *Caenorhabditis elegans*. *Dev. Biol.* **93**, 181–205.
- Sulston, J. E. (1976). Post-embryonic development in the ventral cord of *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **275**, 287–297.
- Sulston, J. E. (2003). *Caenorhabditis elegans*: The cell lineage and beyond (Nobel lecture). *ChemBiochem.* **4**, 688–696.
- Sulston, J. E., Albertson, D. G., and Thomson, J. N. (1980). The *Caenorhabditis elegans* male: Postembryonic development of nongonadal structures. *Dev. Biol.* **78**, 542–576.
- Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
- Sulston, J. E., and White, J. G. (1980). Regulation and cell autonomy during postembryonic development of *Caenorhabditis elegans*. *Dev. Biol.* **78**, 577–597.
- Thomas, C., DeVries, P., Hardin, J., and White, J. (1996). Four-dimensional imaging: Computer visualization of 3D movements in living specimens. *Science* **273**, 603–607.

- van Cleave, H. J. (1932). Eutely or cell constancy in its relation to body size. *Q. Rev. Biol.* **7**, 59–67.
- White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1–340.
- Wiegner, O., and Schierenberg, E. (1998). Specification of gut cell fate differs significantly between the nematodes *Acrobeloides nanus* and *Caenorhabditis elegans*. *Dev. Biol.* **204**, 3–14.
- Yochem, J. Nomarski images for learning the anatomy, with tips for mosaic analysis (January 24, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.100.1, <http://www.wormbook.org>.
- Zhang, S., Ma, C., and Chalfie, M. (2004). Combinatorial marking of cells and organelles with reconstituted fluorescent proteins. *Cell* **119**, 137–144.
- Zhao, Z., Boyle, T. J., Bao, Z., Murray, J. I., Mericle, B., and Waterston, R. H. (2008). Comparative analysis of embryonic cell lineage between *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Dev. Biol.* **314**, 93–99.

---

---

## CHAPTER 13

# The Genetics and Cell Biology of Fertilization

**Brian D. Geldziler<sup>\*</sup>, Matthew R. Marcello<sup>\*,†</sup>, Diane C. Shakes<sup>†</sup>  
and Andrew Singson<sup>\*</sup>**

<sup>\*</sup>Waksman Institute, Rutgers University, Department of Microbiology and Molecular Genetics, Piscataway, New Jersey, USA

<sup>†</sup>College of William and Mary, Department of Biology, Williamsburg, Virginia, USA

<sup>‡</sup>UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey, USA

---

### Abstract

- I. General Introduction/Background
  - II. Finding Sterile Mutants
  - III. Maintaining Sterile Mutants
  - IV. Fecundity Analysis
  - V. Ovulation/Ovulation Rates
  - VI. Spermatogenesis
  - VII. Sperm Migratory Behavior/Tracking
  - VIII. Sperm Competition
  - IX. Oogenesis and Oocyte Maturation
  - X. Assessing Fertilization and Egg Activation in Egg-Sterile Mutants
  - XI. Eggshell Production
  - XII. Analysis of Fertilization-Specific Gene Products
  - XIII. Future Prospects/Issues
- Acknowledgments  
References

---

---

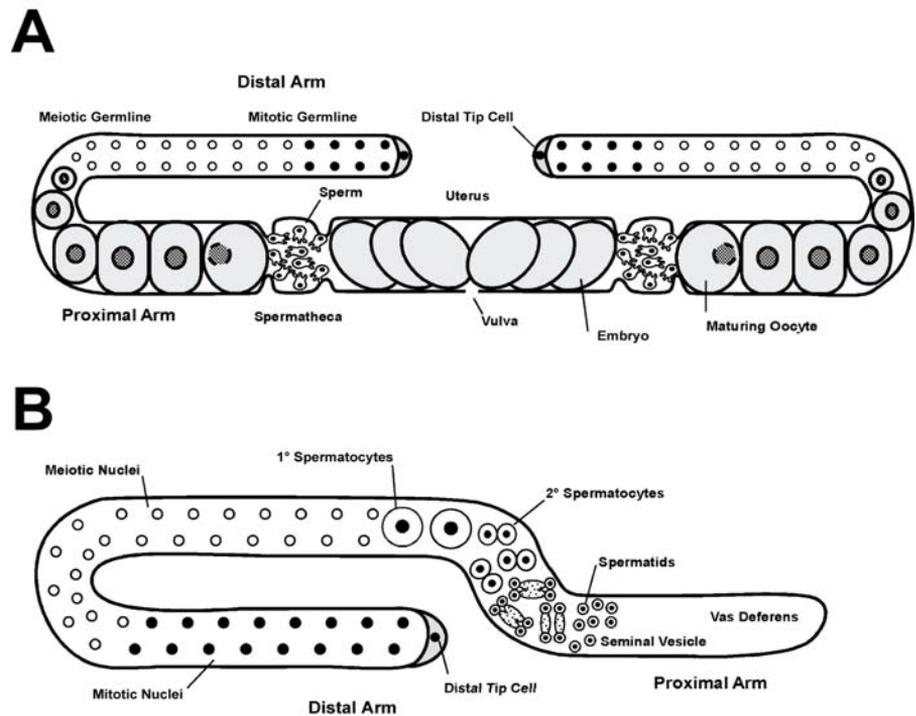
### Abstract

Although the general events surrounding fertilization in many species are well described, the molecular underpinnings of fertilization are still poorly understood. *Caenorhabditis elegans* has emerged as a powerful model system for addressing the molecular and cell biological mechanism of fertilization. A primary advantage is the ability to isolate and propagate mutants that effect gametes and no other cells. This chapter provides conceptual guidelines for the identification, maintenance, and experimental approaches for the study fertility mutants.

## I. General Introduction/Background

Successful fertilization is fundamentally important to a sexually reproducing species and requires a series of well-coordinated events including gamete activation, recognition, signaling, adhesion, and fusion (Primakoff and Myles, 2002; Singson *et al.*, 2001; Wassarman, 1999). Although our current understanding of these processes comes largely from work in marine invertebrates and vertebrate model systems, *Caenorhabditis elegans* has emerged as another powerful system for fertilization studies (Nishimura and L'Hernault, 2010; Singson, 2001; Singson *et al.*, 2008; Yamamoto *et al.*, 2006).

Fertilization in *C. elegans* takes place in the spermatheca, the site of sperm storage in the hermaphrodite. The hermaphrodite reproductive tract consists of a bilobed gonad (Fig. 1A) in which a separate spermatheca connects each lobe to the shared uterus (the male gonad is single-lobed [Fig. 1B]). Within the hermaphrodite gonad, both gamete types are produced in a sequential manner. Sperm are produced first during the last larval stage of development and stored in the spermatheca. The gonad switches to oocyte production in the adult hermaphrodite. Oocytes undergo meiotic maturation as they



**Fig. 1** *C. elegans* adult gonads. (A) Hermaphrodite gonad showing general scheme of oocyte development. (B) Male gonad showing general scheme of spermatogenesis.

move towards the uterus and are ovulated into the spermatheca, where they immediately contact multiple spermatozoa. Blocks to polyspermy exist, as only a single sperm fertilizes each oocyte (Parry *et al.*, 2009). The coordination of events leading to sperm/oocyte contact ensures highly efficient sperm utilization as virtually all functional sperm fertilize oocytes (Kadandale and Singson, 2004; Ward and Carrel, 1979). The hermaphrodite's own sperm can be supplemented by mating to males; male sperm are deposited in the uterus and immediately travel to the spermatheca to await oocyte passage. A sperm-sensing mechanism ensures that metabolically costly oocytes are not wasted; when hermaphrodites lack sperm, they ovulate at a very low basal level. Conversely, the presence of sperm within the hermaphrodite spermatheca causes a dramatic increase in ovulation rate (McCarter *et al.*, 1999; Miller *et al.*, 2001). After fertilization, the zygote secretes a multilayered eggshell and begins embryonic development. Eggs then pass through the uterus and are laid before hatching (Fig. 2).

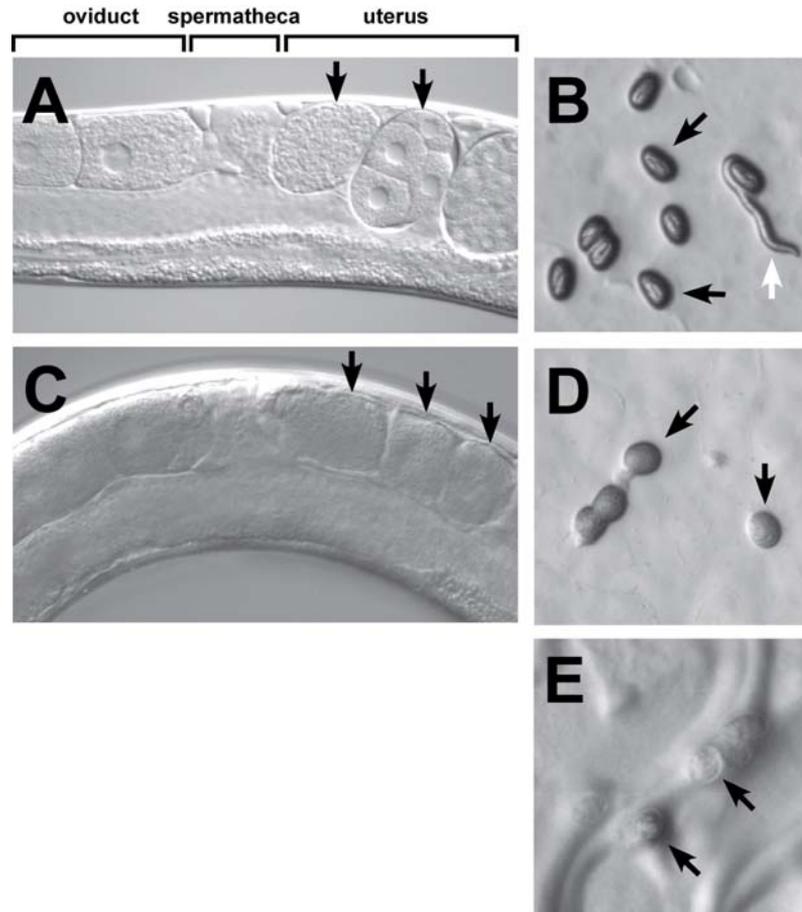
*C. elegans* offers several advantages over other model systems for studying fertilization. Although its amoeboid sperm possess neither a flagellum nor an acrosome (Fig. 3), these sperm successfully perform the same tasks required of all spermatozoa (e.g., migration to the fertilization site, species-specific oocyte recognition, fusion). In addition, the events of fertilization can be directly observed in living animals through the worm's transparent cuticle (McCarter *et al.*, 1999). It is also possible to isolate large quantities of sperm and oocytes, though this is more challenging to do than in some other model systems (Aroian *et al.*, 1997; L'Hernault and Roberts, 1995; Miller, 2006). Fertilization studies in *C. elegans* routinely use molecular and genetic tools that are unavailable or difficult to use in other systems. The complete sequencing of the worm genome and the availability of microarrays greatly simplifies the identification and analysis of genes required for fertility (Singson, 2001; Singson *et al.*, 2008). Perhaps the greatest advantage of *C. elegans* is the ease with which one can perform forward genetics to screen for fertilization-defective mutants (discussed below). Such screens have identified many of these mutants, which may be classified broadly into those mutations affecting sperm (*spe* or *fer* mutants, for *spermatogenesis* or *fertilization* defective) and those affecting eggs/oocytes (*egg* mutants). Note that the *fer* designation has been discontinued and all new sperm development or function mutants are now given the *spe* designation.

Although the majority of characterized *spe/fer* mutations affect sperm *development*, a subset of these mutations specifically affects sperm *function* (i.e., fertilization). *spe-9* class mutants, for example, produce sperm that are unable to fertilize oocytes despite exhibiting normal morphology, motility, and gamete contact (Chatterjee *et al.*, 2005; Kroft *et al.*, 2005; L'Hernault *et al.*, 1988; Singson *et al.*, 1998; Xu and Sternberg, 2003). Recent studies have also uncovered egg mutations that specifically influence fertilization and/or egg activation (Kadandale *et al.*, 2005b). A partial listing of characterized genes required for fertilization is given in Table I. Many of the experimental tools and techniques discussed in this chapter were developed for the study of these genes.

*C. elegans* also enables the evolutionary assessment of fertilization molecules and can help elucidate major molecular themes. For example, EGF-repeat-containing

molecules have been implicated in fertilization across a wide evolutionary spectrum, from HrVC70 in ascidians (Sawada *et al.*, 2004) to SPE-9 in worms (Singson *et al.*, 1998) and SED-1 in mammals (Ensslin and Shur, 2003).

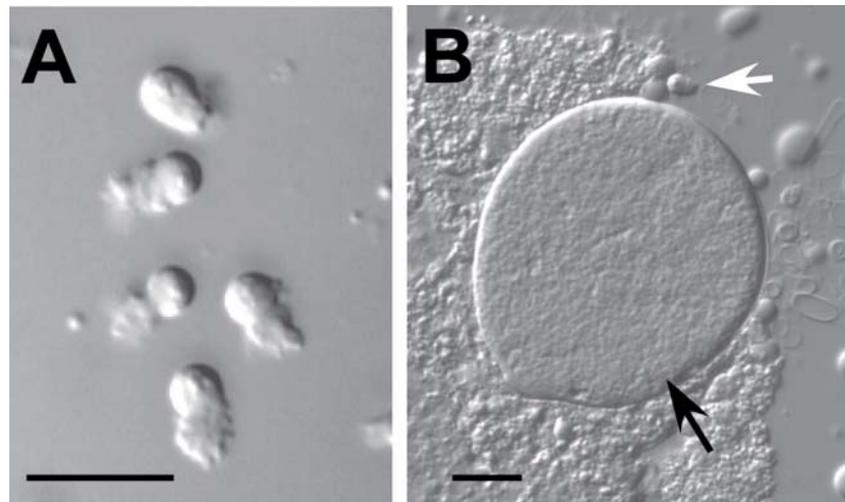
In this chapter, we introduce the major experimental approaches/implications to consider when using *C. elegans* to study fertilization. A general scheme for the study of sterility mutants in *C. elegans* is presented in Fig. 4. Detailed protocols can be found in the original literature and in L'Hernault and Roberts (1995).



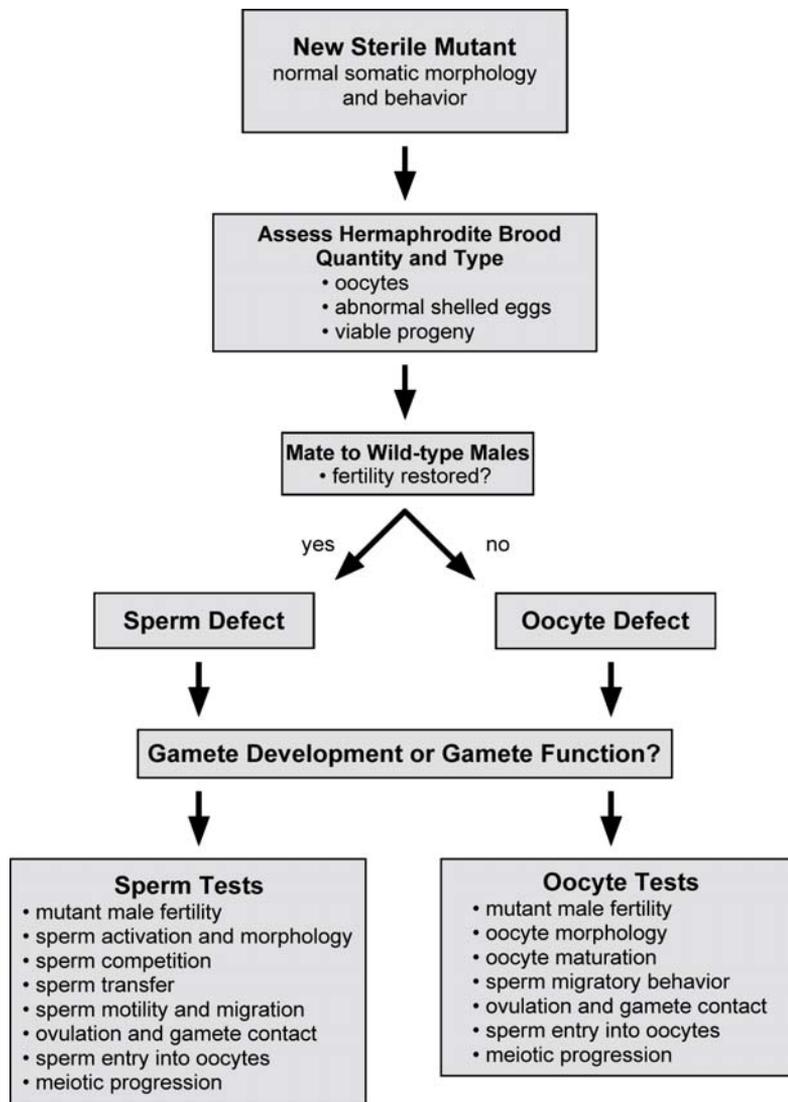
**Fig. 2** Examples of the sterile phenotypes observed when fertilization or proper egg activation does not occur. (A) In wild-type hermaphrodites, embryos (black arrows) can be observed developing in the uterus. (B) Oblong shelled embryos (black arrows) are laid at the 50+ cell stage and ultimately hatch into larvae (white arrow). (C) In sterile mutants (e.g., *spe-9(hc52)* shown here) unfertilized oocytes (black arrows) are seen in the uterus. (D) These soft, flattened, and opaque unfertilized oocytes (black arrows) are also laid. (E) Egg-activation mutants (e.g., *spe-11(hc77)* shown here) lay more rounded thinly shelled eggs or “pebbles” (black arrows) that neither develop nor hatch.

**Table I**Genes required for fertilization and egg activation in *C. elegans*.

Gene	Protein	Localization	Reference
<b>Fertilization</b>			
<i>spe-9</i>	Single-pass transmembrane protein with 10 EGF repeats	Spermatozoa Pseudopod	Singson <i>et al.</i> (1998)
<i>spe-38</i>	Novel four-pass integral membrane protein	Spermatozoa Pseudopod	Chatterjee <i>et al.</i> (2005)
<i>spe-41/trp-3</i>	TRPC channel subunit	Spermatozoa plasma membrane	Xu and Sternberg (2003)
<i>spe-42</i>	Novel seven-pass integral membrane protein	Unknown	Kroft <i>et al.</i> (2005)
<i>egg-1/2</i>	Type-II transmembrane protein with LDL-receptor repeats	Oocyte plasma membrane	Kadandale <i>et al.</i> (2005b)
<b>Egg activation</b>			
<i>spe-11</i>	Soluble protein	Perinuclear in sperm	Browning and Strome (1996)
<i>egg-3</i>	Protein tyrosine phosphatase-like (PTPL)	Oocyte cortex	Maruyama <i>et al.</i> (2007)
<i>egg-4/5</i>	Protein tyrosine phosphatase-like (PTPL)	Oocyte cortex	Parry <i>et al.</i> (2009)



**Fig. 3** Gamete morphology. (A) *In vitro* pronase-activated male-derived sperm. (B) Hermaphrodite dissection showing relative size of *in vivo* activated sperm and oocyte. The oocyte (black arrow) is approximately 160 times the volume of the sperm (white arrow). Bars = 10  $\mu\text{m}$



**Fig. 4** A general scheme for working with *C. elegans* sterility mutants. Once the general class of mutant is determined (*spe* or *egg*) by crossing a new sterile hermaphrodite to wild-type males, one can begin to assess whether the gene is specifically for fertilization, or more generally for gamete development. See text for information regarding specific tests.

---

---

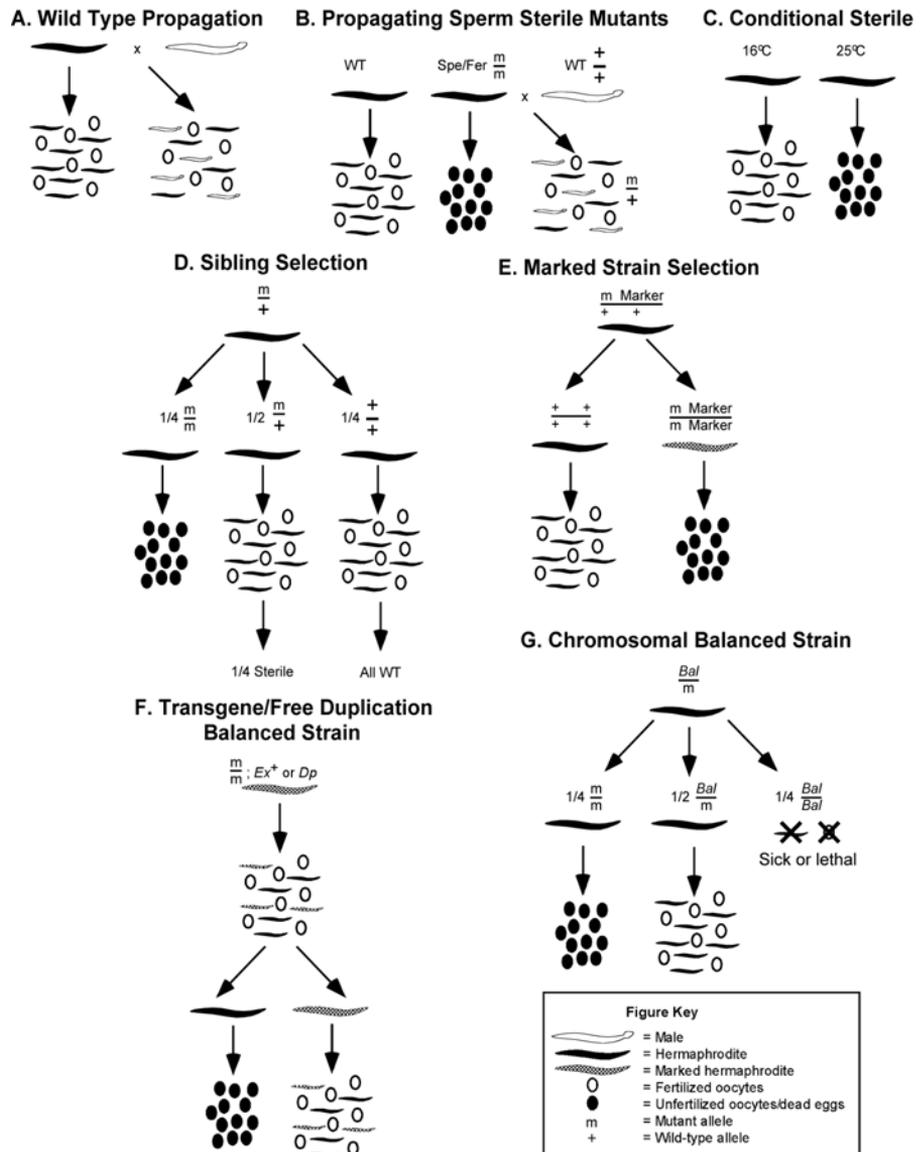
## II. Finding Sterile Mutants

A major advantage of using *C. elegans* for fertilization study is the ability to use forward genetic screens to isolate fertility mutants. Mutant screening in *C. elegans* is typically carried out using ethyl methane sulphonate (EMS) to induce mutations in the germ line of wild-type hermaphrodites. In the subsequent F<sub>2</sub> generation, homozygous mutants are screened for the phenotype of interest (L'Hernault *et al.*, 1988). Using standard concentrations of EMS, investigators can expect to find one null mutation per 2000 gene copies examined (Jorgensen and Mango, 2002) and approximately one in every 30 F<sub>2</sub> generation animals will display a Spe phenotype (S. L'Hernault, personal communication).

The power of these genetic screens lies in the ability to isolate mutations in sperm-specific genes by simply selecting hermaphrodites that are self-sterile but whose oocytes can be still be fertilized by wild-type male sperm (Fig. 5). More than 60 sperm-specific mutants have been identified this way, enabling the genetic delineation of sperm development and function pathways (L'Hernault *et al.*, 1988; L'Hernault, 1997, 2006; Nishimura and L'Hernault, 2010). *C. elegans* is particularly amenable to such sperm-specific screens because they can be carried out with hermaphrodites (rather than in males as in gonochoristic systems such as mice or flies) and thus remain independent of confounding issues such as male mating behavior or secondary sex characteristics (L'Hernault *et al.*, 1988; Lessard *et al.*, 2004; Wakimoto *et al.*, 2004).

Genetic screens for conditional egg-sterile mutants have likewise yielded many interesting candidate genes (G. Singaravelu, D. Shakes, and A. Singson, unpublished). These conditional screens are especially useful for the isolation of fertility mutants as they allow the propagation of *homozygous* mutations in fertility genes. Isolating egg-sterile mutants in standard nonconditional F<sub>2</sub> screens requires considerably more work since progeny cannot be recovered from homozygous egg-sterile mutants. As a result, the mutant chromosome must be recovered from heterozygous siblings.

Conditional fertility screens are similar to the conditional maternal effect lethal (*mel*) screens conducted in *C. elegans* (Jorgensen and Mango, 2002; O'Connell *et al.*, 1998). In a typical *mel* screen, worms with a pre-existing mutation that blocks normal egg-laying are mutagenized, and the uteri of the resulting F<sub>2</sub> hermaphrodites are examined for the presence of dead F<sub>3</sub> progeny. The screen is first carried out at 25 °C, and any F<sub>2</sub> animals that are filled with dead F<sub>3</sub> embryos are subsequently shifted to 16 °C to recover homozygous temperature-sensitive (*ts*) mutants. This same screen can be modified to isolate fertility mutants simply by looking for worms containing unfertilized oocytes rather than dead embryos. Oocytes that have been fertilized and have successfully completed egg-activation encase themselves in a rigid eggshell (Fig. 2A, B). Oocytes that remain unfertilized or that have defects in egg-activation lack this shell and are visibly distinct (the phenotype known as “squashy,” “mushy,” or “ugly brown *mels*”) (Fig. 2C, D). Because such “squashy” oocytes were routinely discarded in most large-scale *mel* screens, it is likely that new



**Fig. 5** Common schemes for the maintenance/propagation of sterile mutations in *C. elegans*. (A) Wild-type strain propagation. (B) Sterility mutant propagation. (C) Conditional sterility mutant propagation. (D) Sibling selection for nonconditional sterility mutants. (E) Marked strain selection. (F) Chromosomally balanced strains. (G) Transgene/free duplication-balanced strains. See text for details.

screens that focus specifically on this phenotype can identify many new fertility genes (G. Singaravelu, D. Shakes, and A. Singson, unpublished).

The *ts* alleles isolated in conditional screens can be used to determine the temporal requirements for a given protein's function or synthesis, and their availability can greatly simplify genetic manipulations (L'Hernault *et al.*, 1988; Putiri *et al.*, 2004). However, although their phenotypes may represent a wide range of severities, *ts* alleles are rarely amorphic (null) and investigators must therefore be cautious when relying on them to make inferences about gene functions/interactions. Null phenotypes are best defined using molecularly verified amorphic mutants.

Reverse genetic methods may also be used to study the function of fertility genes (Geldziler *et al.*, 2004). Microarray analysis in *C. elegans* has helped identify numerous candidates for reverse genetic approaches; 1343 spermatogenesis-enriched and 1652 oogenesis-enriched genes were initially identified (Reinke *et al.*, 2004), and subsequent studies have built on to this original list (Reinke and Cutter, 2009). Many of these genes may function in the events of fertilization.

RNAi is often used as a quick, loss-of-function, reverse genetic approach (Kamath and Ahringer, 2003; Maeda *et al.*, 2001) (also see chapter by Cipriani and Piano). Injecting, feeding, or soaking worms to introduce dsRNA can systemically trigger gene-specific silencing in both the treated animal and its F<sub>1</sub> progeny (Fire *et al.*, 1998; Kamath and Ahringer, 2003; Timmons and Fire, 1998; Timmons *et al.*, 2001). RNAi has been successfully used to study *C. elegans* genes with early sperm or oocyte development functions (Sumiyoshi *et al.*, 2002), but most *C. elegans* sperm genes have proven refractory to this method for unknown reasons (Geldziler *et al.*, 2004). Consequently, RNAi has not been a productive method for studying gene candidates with fertilization functions, and systematic RNAi screens have undoubtedly failed to identify many key molecules required for spermatogenesis and sperm function. Nevertheless, injection and soaking-based RNAi approaches may be useful in identifying at least some egg-sterile mutants (Kadandale *et al.*, 2005b; Maeda *et al.*, 2001).

RNAi and similar knockdown methods are relatively easy ways to assess gene function, but RNAi-induced phenotypes are not genetic mutations and therefore require careful interpretation. Genes differ greatly in the extent to which they can be functionally reduced by RNAi; RNAi phenotypes range from no effect to total loss of function. RNAi phenotypes are also highly sensitive to experimental subtleties such as genetic background and exposure method, time, and temperature. For instance, the effective RNAi knockdown of the *egg-1* and *egg-2* genes requires injection or soaking treatment at 25 °C to see a complete fertilization defect (Kadandale *et al.*, 2005b; Lee and Schedl, 2001; Maeda *et al.*, 2001). Feeding RNAi yields only incomplete knockdown phenotypes at best. Such variability can confound analyses and makes proper controls critically important.

Since RNAi methods do not provide actual germ line mutations, other methods are required for further genetic studies/manipulations (Liu *et al.*, 1999). Deletion library screening is one technique used extensively in *C. elegans* as a powerful large-scale method for deriving true knockout mutations in target genes known only via their

sequence (Jansen *et al.*, 1997). Pools of mutagenized worms are split for long-term viable freezing and DNA extraction. The DNA pools are then screened for deletions in the gene of interest via PCR and populations containing such deletions are sequentially subdivided to isolate a pure strain homozygous for the deletion. Such a strategy was used to demonstrate an essential role for various genes in fertilization (Kadandale *et al.*, 2005b; Parry *et al.*, 2009; Xu and Sternberg, 2003). New methods for gene knockout based on DNA transposition will also be important for generating fertility gene mutations (see chapter by Robert and Bessereau).

---

---

---

### III. Maintaining Sterile Mutants

Once sterility mutations have been generated and identified it becomes necessary to maintain them. Nonconditional recessive *spe* hermaphrodites cannot be maintained as self-fertile homozygotes, but the mutant chromosome can be propagated by crossing *spe* hermaphrodites with wild-type males. These crosses are referred to in the literature variously as “rescue,” propagation, complementation, or maintenance crosses (Fig. 5). Since these maintenance crosses are nothing more than standard genetic backcrosses (Fig. 5B), many *spe* mutant strains have been cleared of extraneous mutations. Nonconditional egg-sterile mutants cannot be complemented by mating, however, and must be maintained using labor-intensive “sib selection” (Fig. 5D) prior to genetic balancing (see below); the mutation of interest is recovered from a heterozygous sibling of the homozygous sterile mutant. Since sibling heterozygotes are often phenotypically wild-type, worms must be individually plated and followed for the segregation of sterile progeny to assure the correct genotype and maintain the strain. This must be done with each generation.

Because fertile males are required to maintain sperm-sterile mutants as described above, a ready supply is essential. In *C. elegans*, males naturally arise at too low a frequency (1 in 1000, due to nondisjunction of the X chromosome) to be useful for this purpose (although heat shocking hermaphrodites at 25–30 °C for >6 h will increase this frequency) (Hodgkin, 1999). Consequently, many researchers keep daily matings of wild-type worms (which generate 50% male progeny) to ensure a sufficient quantity (Fig. 5A). Alternatively, specific *him* (high incidence of males) strains may be used (Hodgkin, 1997). These mutants show increased X chromosome loss and therefore generate many males. *him-3*, *him-5*, and *him-8* do not affect spermatogenesis and are routinely used for this purpose (Nelson and Ward, 1980; Zannoni *et al.*, 2003). They do affect chromosome number and ploidy, however, so their use must be monitored to ensure they are not affecting the original function/process-of-interest. Although these are the most widely used methods for obtaining males, other methods such as heat-shock, ethanol treatment, and RNAi against *him* genes can generate males for crossing (Fay, 2006; Hodgkin, 1999).

Maintaining sterile strains as heterozygotes is tedious and requires manual selection, making it impractical for large-scale operations. Fortunately, a variety of genetic balancers allow mutants of interest to be maintained and propagated without

manual manipulation (Fig. 5 E–G). We will only briefly mention genetic balancers here, referring the reader to Chapter 7 in the previous edition of this text and the chapter by Jones *et al.* in this edition for a more thorough discussion (Edgley *et al.*, 1995).

Many types of balancers can be used to maintain sterile mutations, including chromosomal rearrangements (translocations, duplications, and inversions) and transgenes specifically constructed for this purpose. Regardless of type, all good balancers used for existing strain maintenance share common characteristics; heterozygotes (or transgene-containing animals) must be viable and fertile as well as have a unique, easy-to-score phenotype or mutations that effectively eliminate homozygote balancers from the population (see chapter by Jones *et al.*) (Edgley *et al.*, 1995).

When working with fundamental processes such as fertility, one must take care to ensure that balanced strains are not lost. Fertility is highly selected, and any reversions or mutations conferring a fertility advantage will quickly take over a stock. Further, we have observed weakly fertile stocks become more fertile over many generations (I. Chatterjee and A. Singson, unpublished). Conversely, spontaneous mutations that further negatively affect fertility may arise in the balancer strains, confounding analyses. In addition, balancers themselves can occasionally stop functioning to DNA rearrangements, resulting in the loss of balancing. Therefore, it is very important to carefully monitor balanced strains and to keep frozen stocks soon after their construction.

DNA injected into the worm gonad can form heritable extrachromosomal arrays (Mello and Fire, 1995) that can be used to balance mutations. Because these balancing transgenes can often be lost, mosaic animals can be studied to determine whether a given sterile gene is required in the germ line for its action. For example, *spe-19* is maintained via an extrachromosomal balancing array that also contains a GFP marker (*myo3::gfp*, which expresses in body wall muscle) (Geldziler *et al.*, 2005). Studies of mosaic (glowing sterile and nonglowing fertile) animals revealed that glowing Spe animals failed to carry the transgene in the germ line while rare nonglowing fertile animals sired glowing progeny (and therefore contained the transgene in the germ line). These results strongly suggest that *spe-19* is required in the germ line for function, but do not rule out an additional requirement in somatic tissue. Note that transgene expression is often repressed in the germ line (so-called “germ line silencing”), somewhat limiting the usefulness of this type of analysis for fertility study (Putiri *et al.*, 2004).

Maintenance of *spe-8* mutations provides another balancer example. Using the free duplication *sDp2* which contains wild-type copies of both *dpy-5* and *spe-8*, the strain is maintained as a homozygous *spe-8 dpy-5* double mutant and affords an easy visual assay for the *spe-8* mutation. Animals complemented by the duplication are phenotypically wild-type; animals that do not contain the duplication are Spe and Dpy (Edgley *et al.*, 2006; Singson *et al.*, 1999).

Of course, many sterile mutations are conditional and do not require balancers. Temperature-sensitive (*ts*) alleles of many sterility genes (both sperm and egg) exist

or can be generated by appropriate screening strategies (Kadandale *et al.*, 2005a; L'Hernault *et al.*, 1988; Putiri *et al.*, 2004). Although, by definition, *ts* sterility alleles produce progeny at the permissive temperature, the fertility of these strains may not be as robust as in wild-type animals; they are sometimes leaky, giving low brood sizes at the permissive temperature.

In some instances, sterile mutants can be stably maintained as homozygotes through mating because their mutant phenotype is sex-specific. All members of the *spe-8* class of genes, for example, are specifically required for hermaphrodite spermiogenesis (Geldziler *et al.*, 2005; L'Hernault, 1997; Shakes and Ward, 1989). However because the mutant males are fertile, homozygous populations can be maintained simply by crossing mutant males to mutant hermaphrodites. Mutants with male-specific spermatogenesis-defects also exist (Stanfield and Villeneuve, 2006). Although ideal for maintaining mutant populations, these types of sex-specific sterile mutations are rare.

---

---

#### IV. Fecundity Analysis

Once mutant strains of interest have been isolated, outcrossed, and genetically stabilized, they can be characterized in detail. Quantifying fertility using brood analysis, for example, is of fundamental importance for determining a mutation's functional severity. In hermaphrodites, this is done by allowing individually plated wild-type and mutant L4 animals to self-fertilize and then counting, averaging, and comparing the numbers of eggs/oocytes laid. In practice, eggs/oocytes are usually counted daily while the hermaphrodite is moved to a new plate to facilitate counting the next day's totals. Animals are typically transferred to new plates until no new eggs/oocytes are laid within a 24 h period. Alternatively, the eggs/oocytes may be removed each day to avoid damaging delicate mutants via unnecessary manipulations. Brood sizes may also be examined over shorter time periods (an approach that is less traumatic for the worm and less labor intensive for the investigator). The total number of lifetime progeny can be assessed using the "mate to death" assay (Kimble and Ward, 1988); the daily introduction of new males constantly replenishes the experimental hermaphrodite's sperm supply. Brood results are most usefully expressed as the percentage of wild-type numbers; absolute numbers are less informative due to the wide variation in actual brood sizes among animals resulting from unknown or difficult-to-control factors. In studies of maternal-effect lethal mutants, quantitative assessments of fertility should include the percentage hatched as well as distinct categories for non/weak-shelled "oocytes" (oocyte or egg steriles), hard-shelled dead embryos, and viable larvae (Fig. 2).

Male fertility is typically assessed by crossing mutant males to marked strains and counting the number of outcrossed (wild-type looking) progeny. *dpy-5* is often used as a marker as it is easily scorable, but any easily identified recessive morphological marker will suffice. Another method obviates marked strains; one simply counts the number of F<sub>1</sub> males and multiplies by two, since males sire 50% male progeny.

Markers/mutations may affect male mating behavior/efficiency; *some him* mutants, for example, have smaller broods than wild-type animals due to aneuploidy, and almost all morphologically marked males have at least partially compromised mating efficiencies (Hodgkin, 1997). Mating behavior/efficiency issues must therefore be distinguished from fertility issues when assessing overall fecundity.

Minimizing brood size variability due to environmental and genetic factors is important. Because fertility is age dependent, age-matched wild-type controls are essential. Nutritional effects can be minimized by using growth plates made at the same time and seeded from the same bacterial culture. Any plate contamination will complicate oocyte counts. For consistency of measurements, a single investigator should collect all data and proper statistical analysis must be performed to determine the significance of observed differences between groups. When assessing transgenic lines, it is important to remember that transgenic worms frequently produce smaller broods due to the unrelated effects of germ line silencing (Putiri *et al.*, 2004). The brood sizes of several different lines should be assessed as individual transgenic arrays frequently exhibit distinct levels of expression.

Directionality to sperm migration has been observed; sperm sometimes move to a single spermatheca rather than both, resulting in a mated Spe hermaphrodite laying a mix of eggs and unfertilized oocytes. If this is seen, the hermaphrodite should be examined to determine whether sperm are differentially localized in this manner.

---

---

---

## V. Ovulation/Ovulation Rates

Ovulation levels/rates are closely related to overall brood sizes and can be used to differentiate Spe mutant classes. Meiotic maturation/ovulation is stimulated by a sperm-derived signal: the major sperm protein (MSP) (Kosinski *et al.*, 2005; Miller *et al.*, 2001). On an average, sperm-containing wild-type hermaphrodites ovulate approximately 2.5 times per gonad arm per hour, while *fog-2* females (which lack sperm) ovulate only 0.09 times per gonad arm per hour (McCarter *et al.*, 1999; Miller *et al.*, 2003). Consequently, ovulation rates can be used to indirectly assess the presence of spermatids/spermatozoa. The ovulation rates of spermatocyte-arrest Spe mutants are significantly lower than those of Spe mutants whose genes affect spermiogenesis/sperm function (Kadandale and Singson, 2004; Singson *et al.*, 1998). Presumably only the spermatid/spermatozoa-producing Spe mutants have the ability to produce and respond to the MSP signal and ovulation rates have been used to distinguish *spe* mutant classes (Chatterjee *et al.*, 2005; Singson *et al.*, 1998, 2008).

Measurement of ovulation rates can be made in several ways. Paralyzed or anesthetized animals can be directly observed under a compound microscope using time-lapse video, for example (McCarter *et al.*, 1999; Ward and Carrel, 1979), or one can use a dissecting scope to count oocytes on single worm plates at set time intervals (Kadandale and Singson, 2004; Miller *et al.*, 2003). Reproducible plate counts require fresh uncontaminated bacterial lawns but are less traumatic for the worms

and yield results comparable to the more labor-intensive direct observations. In either method, age-matched controls are absolutely essential because the rate-limiting factor for ovulation rate changes with age; for young adults, it is oocyte growth (Miller *et al.*, 2003) while for older animals it is sperm. As with brood assessments, relative differences are more important than absolute numbers, and appropriate statistical analysis is mandatory.

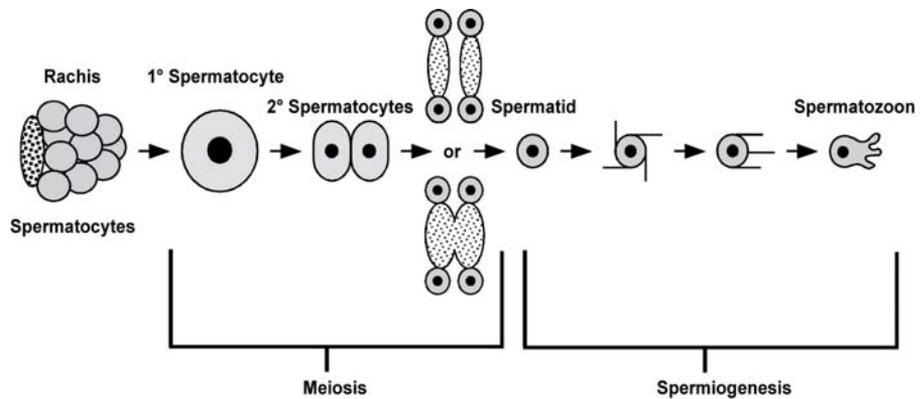
Ovulation rates can be measured in either self-crossed or male-mated hermaphrodites. Self-crossed hermaphrodites are easier to use but if they become sperm-depleted during the course of the assay, their ovulation rates will decrease over time and result in lower overall averages unrelated to the mutant-of-interest. To avoid sperm-depletion effects, mutant hermaphrodites may be plated with wild-type males to maintain nonbasal ovulation rates for the duration of the assay. Conversely, the ability of mutant sperm to induce ovulation can be assessed by measuring the ovulation rates of *fem* or *fog* “females” crossed with mutant males.

Many ovulation-defective mutants exhibit a secondary phenotype whereby oocytes undergo multiple rounds of DNA synthesis to become polyploid (Iwasaki *et al.*, 1996). The location of these endomitotic (EMO) oocytes offers clues to the nature of the mutation; EMO oocytes in the gonad arm suggest spermathecal and/or ovulation defects while EMO oocytes in the uterus suggest defects in fertilization itself or in egg activation following sperm entry. Note that in the absence of sperm, wild-type hermaphrodites also produce EMO oocytes in their uterus (Ward and Carrel, 1979).



## VI. Spermatogenesis

In *C. elegans*, gametogenesis and early spermatogenesis occur in a linear progression along the length of the tube-like gonad (Fig. 6, but also see Fig. 1). After



**Fig. 6** Spermatogenesis in *C. elegans*. Major developmental stages are indicated.

leaving the distal mitotic zone, germ cells simultaneously enter meiosis and become committed to spermatogenesis (Jaramillo-Lambert *et al.*, 2007). After homolog pairing, they then enter an extended period of pachytene. By mid to late pachytene, sperm-specific proteins begin to accumulate within the developing spermatocytes. As the most proximal spermatocytes exit pachytene, synaptonemal complexes between the paired homologs disassemble. Then global transcription ceases as the chromatin fully condenses and the spermatocytes enter a distinctive karyosome stage during which the chromosomes come together in a single mass within the still-intact nuclear envelope (Shakes *et al.*, 2009). Throughout this initial phase of spermatogenesis, the germ-cell nuclei are only partially encased within cellular membranes as the individual spermatocytes are syncytially connected to a common central cytoplasmic core called the rachis. Spermatocytes transition to the meiotic division stage of spermatogenesis as individual spermatocytes then detach from this rachis and their microtubule organization switches from a network to a centrosome-based pattern (Shakes *et al.*, 2009).

During the first meiotic division, 1° spermatocytes undergo a symmetrical, actin-based (and sometimes incomplete) division to form two 2° spermatocytes. The second round of meiotic chromosome segregation rapidly follows. Anaphase II is followed by the transient formation of a shallow cleavage furrow, but this furrow rapidly regresses. Cellular components unnecessary for subsequent sperm formation accumulate in a central residual body whereas individual spermatids bud and detach from this residual body using poorly understood, non-actin-based mechanisms of division (Shakes *et al.*, 2009; Ward *et al.*, 1981). Once spermatids detach from the residual body, they undergo a rapid maturation process that includes final compaction of the sperm chromatin, release of MSP from a paracrystalline assembly known as the fibrous body into the cytosol, and docking of the sperm-specific “membranous organelles (MOs)” with the plasma membrane (Shakes *et al.*, 2009; Ward *et al.*, 1981).

Although some investigators are beginning to include both residual body formation and initial events of spermatid maturation as part of “spermiogenesis” (Wu *et al.*, 2010), within the *C. elegans* literature, spermiogenesis is most frequently defined as the subsequent dramatic morphogenesis of spherical, sessile spermatids into amoeboid, motile, fertilization-competent spermatozoa (Shakes and Ward, 1989). Cellular changes include plasma membrane flow to the site of the newly developing pseudopod, fusion of caveolae-like MOs to the cell body plasma membrane, and the formation of a dynamic MSP pseudopod cytoskeleton. The entire transition takes less than 10 min, and the formation of filopodia-like spikes precedes the formation of fully motile pseudopods (Shakes and Ward, 1989; Singaravelu *et al.*, 2011). For male-derived spermatids, spermiogenesis occurs after ejaculation. For hermaphrodite-derived spermatids, the process begins as they are pushed into the spermatheca with the passage of the first oocyte (Singson, 2006). In males, premature spermiogenesis is actively repressed by a protease inhibitor known as SWM-1 (Stanfield and Villeneuve, 2006). Unlike many other examples of morphogenetic change, *C. elegans* spermiogenesis occurs without actin/tubulin cytoskeleton

regulation or new protein synthesis; ribosomes, actin, and tubulin are all discarded during the spermatid budding division.

It is important to distinguish between sperm *function* mutants (which affect fertilization directly via defects in sperm–egg interaction) and sperm *development* mutants (which affect the process only indirectly via defects in sperm morphogenesis or motility). Sperm development itself can be quickly assessed in microscope preparations of dissected worms. To study early spermatogenesis, “sperm squashes” are made: males are placed in sperm buffer and dissected at the pharynx–intestine junction. A coverslip is used to press the gonad into a monolayer of spermatocytes and spermatids that are then examined using DIC microscopy for spermatid number, shape, and size. Lipid-soluble dyes such as Hoechst 33342 may be included to detect DNA via epifluorescence, although UV exposure causes the cells to deteriorate rapidly. For further analysis, one can quickly freeze these sperm squash preparations on dry ice, remove the coverslip, and prepare the sample for immunofluorescence (for more details see the “Immunofluorescence” chapter in this volume by Shakes *et al.*, 2009). Useful markers for the general staging of spermatogenesis include combinations of DAPI (DNA) with antibodies against  $\alpha$ -tubulin (Sigma FITC-labeled DM1A), phosphorylated histone H3 (Ser10) (Upstate Biotechnology), and/or MSP (Developmental Studies Hybridoma Bank).

Spermatocytes/spermatids can be easily isolated from males, but naturally activated spermatozoa must usually be isolated from the gonads of mated hermaphrodites (Chatterjee *et al.*, 2005). The isolation of *in vivo* activated spermatozoa directly from self-fertile hermaphrodites is substantially more challenging as hermaphrodite spermatozoa are 50% smaller (Geldziler *et al.*, 2006; LaMunyon and Ward, 1998) and significantly less numerous.

Spermiogenesis may also be assessed *in vitro* via the addition of known activators to sperm media before worm dissection. Pronase (a nonspecific collection of proteases) can be used to examine activation using light microscopy, but other choices such as monesin (a cationic ionophore) or triethanolamine (a compound affecting intracellular pH) are more appropriate for studies using immunofluorescence (Chatterjee *et al.*, 2005; Roberts *et al.*, 1986; Singaravelu *et al.*, 2011; Zannoni *et al.*, 2003).

---

---

---

## VII. Sperm Migratory Behavior/Tracking

For successful *C. elegans* fertilization, male-derived spermatozoa must first migrate from the vulva (the site of insemination) to the spermatheca. Sperm (regardless of derivation) must then remain there, either by firmly attaching themselves to the spermatheca wall or by crawling back if dislodged by passing oocytes. Defects in hermaphrodite sperm retention can be assessed by comparing the number of sperm in whole-mount, DAPI-stained young adult hermaphrodites immediately before egg-laying with that of slightly older siblings that have already begun laying eggs

(L'Hernault *et al.*, 1988). The morphology of DAPI-stained sperm nuclei is quite distinctive and can be easily scored within the spermatheca and uterus. Alternatively, the hermaphrodite gonad can be directly observed using DIC microscopy and time-lapse photography. Animals can be dissected on glass microscope slides or anesthetized on agar pads to study sperm number and morphology. It can be challenging, however, to follow the four-dimensional movements of an individual spermatozoon using DIC optics alone.

Fluorescent methods for sperm-tracking experiments include the use of Nile Blue (Ward and Carrel, 1979), SYTO17 (Hill and L'Hernault, 2001; Singson *et al.*, 1998), or MITO tracker (Kubagawa *et al.*, 2006; Stanfield and Villeneuve, 2006). Males are labeled by soaking them in a dye-containing solution and are then mated to unlabeled hermaphrodites. These hermaphrodites are subsequently anesthetized and examined via fluorescent microscopy for the presence and location of sperm.

Care must be taken to ensure that the fluorescent dye used does not adversely affect the process/function under study. The use of mutant hermaphrodites with reduced gut autofluorescence (such as *daf-4* or *glo-1*) facilitates the identification of male-derived sperm within the female reproductive tract (Artal-Sanz *et al.*, 2003; Kroft *et al.*, 2005; Hill and L'Hernault, 2001).

---

---

---

## VIII. Sperm Competition

Although hermaphrodite self-fertilization is the primary mode of *C. elegans* reproduction, mated hermaphrodites produce predominantly outcrossed progeny because of the competitive superiority of male-derived sperm (termed “sperm competition”). This phenomenon is a primary reason that *C. elegans* is such a powerful genetic system. The transition towards outcross progeny typically occurs within 1 day after male-introduction; by day two, progeny are almost exclusively outcrossed (LaMunyon and Ward, 1995, 1997, 1998, 1999; Singson *et al.*, 1999; Ward and Carrel, 1979). Sperm competition is independent of the ability to fertilize; the sperm of fertilization-defective *C. elegans* mutants can effectively outcompete hermaphrodite sperm, reducing the self-fertility of these mated hermaphrodites (Singson *et al.*, 1999). The dominance of male-derived sperm may be partially due to size; *C. elegans* male-derived sperm are approximately 50% larger than hermaphrodite-derived sperm, move faster, and are thought to displace them from the distal end of the spermatheca (LaMunyon and Ward, 1998). However, size *per se* cannot be the complete explanation, as *C. remanei* male-derived sperm do not reduce self-fertility in mated *C. elegans* hermaphrodites, despite being two times larger in diameter (Hill and L'Hernault, 2001). *C. elegans* hermaphrodites may also have an independent mechanism for actively selecting functional sperm (Kadandale and Singson, 2004).

Sequential mating experiments with multiple males suggest sperm competition in *C. elegans* is limited to males versus hermaphrodites. This is perhaps not surprising given that in wild populations male sperm are more likely to encounter

hermaphrodite sperm within the spermatheca than sperm from another male (LaMunyon and Ward, 1999).

The competitive ability of male-derived sperm has useful implications for the study of sterile mutants in *C. elegans*. To be competitive, the males themselves must successfully mate and transfer sperm to the hermaphrodite while their sperm must have successfully completed *in vivo* spermiogenesis, become motile, and migrated to the spermatheca. Sperm competition assays are therefore valuable tools for the characterization of sterile mutations in *C. elegans*, but these assumptions should be confirmed via other direct means (such as the sperm tracking methods described above). In a typical experiment, morphologically marked L4 hermaphrodites (e.g. *dpy-5*) are crossed with unmarked mutant males. The numbers of outcrossed (Non-Dpy) and selfed (Dpy) progeny are counted, and the percentage self-progeny is recorded (Chatterjee *et al.*, 2005; Geldziler *et al.*, 2005; Kroft *et al.*, 2005; Singson *et al.*, 1999; Xu and Sternberg, 2003).

In addition to unmated hermaphrodite and wild-type male controls, only healthy and age-matched worms should be used as male mating behavior and efficiency may be affected by age and overall health. To minimize variations in the microenvironment, all crosses should be done concurrently using the same food batch and incubator.

---

---

---

## IX. Oogenesis and Oocyte Maturation

The earliest stages of oocyte and sperm development are indistinguishable in *C. elegans* hermaphrodites; a common pool of germ cells gives rise to both gamete types (Fig. 1). During oogenesis, the most proximal pachytene nuclei are induced to exit their arrest by a localized MAPK-mediated signal (Church *et al.*, 1995). They then undergo programmed cell death or begin the late stage of oocyte growth and differentiation (Gumienny *et al.*, 1999). This developmental switch is partly regulated by GLD-1, which localizes to the pachytene region of the germ line where it actively represses the expression of late-stage oocyte differentiation markers (Lee and Schedl, 2001). As nonapoptotic presumptive oocytes then pass through the loop region, they progress from diplotene to diakinesis. During this phase, they rapidly enlarge while remaining connected to a progressively narrowing rachis. In wild-type hermaphrodites only the most proximal oocytes are fully cellularized, and only the most proximal oocyte undergoes oocyte maturation.

A variety of cell cycle and differentiation markers are available to check the developmental progression of proximal oocytes for those egg-sterile mutants that may have defects in late-stage oocyte differentiation. In the early stages of this post-loop differentiation process, wild-type oocytes begin to express two proteins whose levels continuously increase during oocyte development – OMA-1 (Detwiler *et al.*, 2001; Lin, 2003) and the yolk protein receptor RME-2 (Grant and Hirsh, 1999). The uptake of intestinally synthesized yolk proteins can also be visualized using the vitellogenin GFP strain YP170::GFP (Grant and Hirsh, 1999).

As oocytes progress closer to the spermatheca, they can be distinguished by their temporal response to the MSP maturation signal provided by sperm in the spermatheca (McCarter *et al.*, 1999; Miller *et al.*, 2003). Approximately every 23 min, the most proximal oocyte in the gonad undergoes meiotic maturation in response to MSP via VAB-1, an Eph receptor protein tyrosine kinase (McCarter *et al.*, 1999; Miller *et al.*, 2003). This progressive response can be monitored using antibodies against activated MAPK (MAPK-YT) and Ser10 phosphohistone H3 (Hsu *et al.*, 2000; Miller *et al.*, 2001).

Maturation promoting factor (MPF) is necessary for allowing the oocyte to progress from mitotic division to metaphase I of the first meiosis (Burrows *et al.*, 2006; Schmitt and Nebreda, 2002). MPF is a complex of a cyclin-dependent kinase (Cdk1) and cyclin B and the activity of this complex is inhibited by WEE-1.3 (Burrows *et al.*, 2006; Doree and Hunt, 2002). MPF promotes a progression of events beginning as early as in the -3 oocyte (the third oocyte proximal to the spermatheca). One of the earliest events is the phosphorylation of serine 10 on histone H3 (Burrows *et al.*, 2006). The subsequent sequence of events occur primarily in the -1 oocytes and can be visualized using DIC microscopy. These include the disappearance of the nucleolus, migration of the nucleus to the distal side of the oocyte, nuclear envelope breakdown, and rounding of the oocyte (McCarter *et al.*, 1999). Other events including entry into metaphase I of meiosis and formation of the meiosis I spindle can be assessed in strains that include expression constructs with GFP:His2B and GFP:tubulin to visualize the chromatin and microtubule structures (McNally *et al.*, 2006; Tenenhaus *et al.*, 2001). The anaphase-promoting complex subsequently promotes the rotation of the meiotic spindle and the metaphase to anaphase transition (McNally and McNally, 2005).

---

---

---

## X. Assessing Fertilization and Egg Activation in Egg-Sterile Mutants

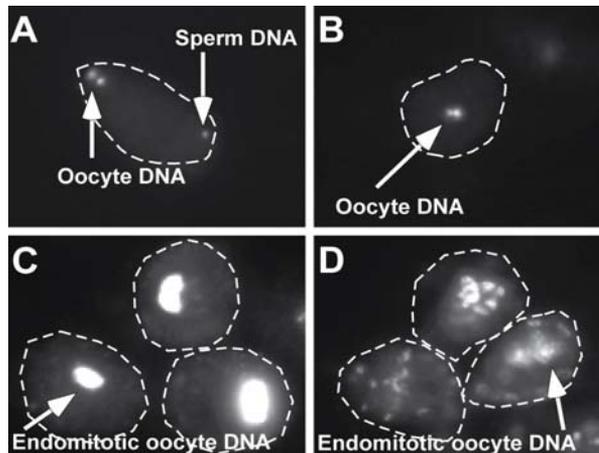
The oocyte enters the spermatheca in metaphase I of the first meiotic division and meiotic resumption and fertilization occur concurrently (McNally and McNally, 2005; Ward and Carrel, 1979). As the oocyte enters the spermatheca, the meiosis I spindle assembles into a pentagonal array of chromosomes. Next, the DNA is translocated to the cortex, the spindle then rotates perpendicular to the cortex, and the chromosomes begin to separate (Albertson and Thomson, 1993; McNally and McNally, 2005). If fertilization occurs, the cell will proceed through anaphase I and half of the chromosomes will be deposited into the first polar body (McNally and McNally, 2005). The oocyte chromosomes subsequently undergo the second meiotic division and a second polar body is extruded. Finally, pronuclei form around both the maternal and paternal chromatin (McNally and McNally, 2005; Sadler and Shakes, 2000).

In addition to triggering meiotic resumption, the newly fertilized egg undergoes a number of changes that are necessary for proper egg activation and embryo

development. These include the activation and/or degradation of selected maternal mRNAs and proteins, release of cortical granules (CGs), secretion of a chitinous eggshell, and the mounting a membrane block to polyspermy to prevent fertilization by a second sperm (Horner and Wolfner, 2008; Marcello and Singson, 2010; Singson *et al.*, 2008; Stitzel *et al.*, 2007). If all of these processes are coordinated and completed properly, the embryo will develop as it passes through the uterus and is eventually laid at approximately the 30-cell stage (Ringstad and Horvitz, 2008; Singson *et al.*, 2008).

In mutagenesis or RNAi screens, egg-sterile animals (*egg*) with defects in either sperm–egg fusion or egg activation are identified in the same way as *spe* mutants. Animals are screened for mutants that lay “eggs” that either possess weak, osmotically sensitive eggshells (Fig. 2E) or lack eggshells altogether (Fig. 2D). To distinguish egg-activation defect mutants from fertilization defective mutants, DAPI staining can be used to score young meiotic stage “eggs” within the spermatheca or uterus for the presence of sperm chromatin (sperm entry) (Fig. 7).

If mutant oocytes remain unfertilized, they will lack both a sperm chromatin mass and meiotic polar bodies (McNally and McNally, 2005) (Fig. 7B). In unfertilized oocytes, the maternal chromosomes initiate the meiotic divisions and reach anaphase I but the resulting anaphase chromosome masses subsequently decondense to form two distinct pronuclei. These unfertilized oocytes fail to form polar bodies or attempt the second meiotic division (McNally and McNally, 2005). During the subsequent rounds of endomitotic cell cycling they form a single, large, polyploid DNA mass (Chatterjee *et al.*, 2005; Doniach and Hodgkin, 1984; Miller *et al.*, 2003) (Fig. 7C).



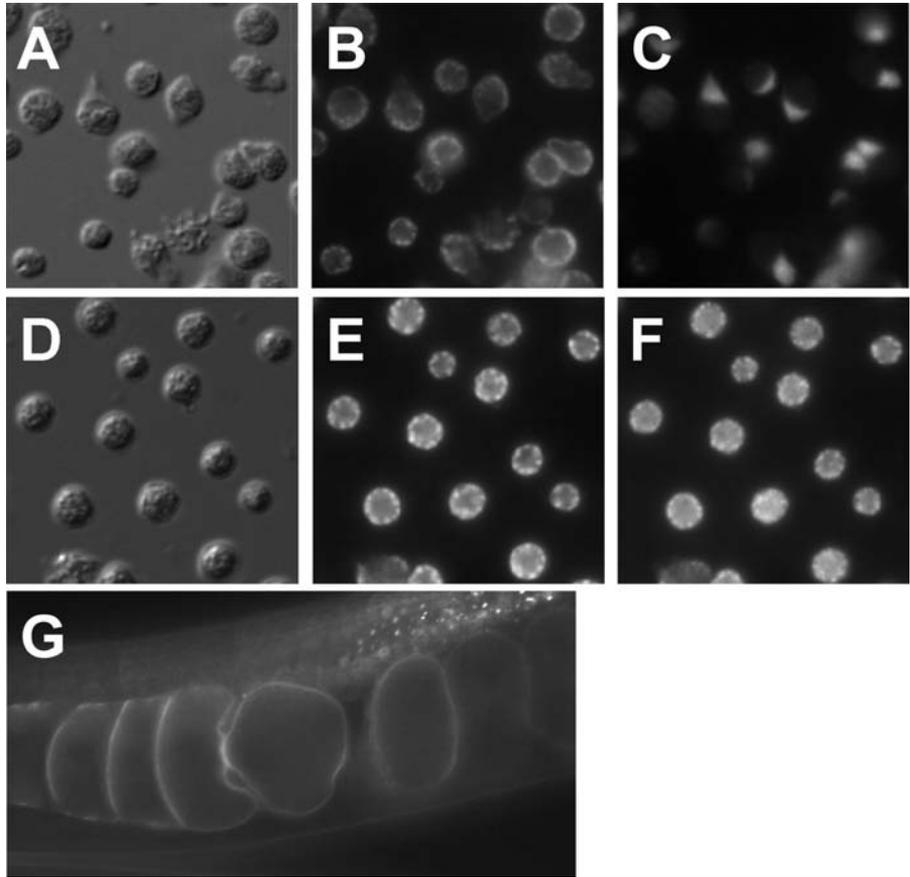
**Fig. 7** DAPI-stained dissected oocytes. (A) Newly fertilized oocyte with both oocyte DNA and a visible sperm chromatin mass. (B) Unfertilized oocytes lacking a sperm chromatin mass. (C) Older endomitotic unfertilized oocyte. (D) Older endomitotic egg-activation mutant oocyte.

In contrast, newly fertilized oocytes from maternal or paternal egg-activation mutants are expected to contain a visible sperm chromatin mass (Kadandale *et al.*, 2005b; Maruyama *et al.*, 2007; Parry *et al.*, 2009) (Fig. 7A). Further, in egg-activation mutants, a distinct endomitotic phenotype is seen. Rather than having a single large polyploidy DNA mass, many smaller DNA masses are formed (Fig. 7D). This is likely due to the action of sperm-contributed centrosomes separating newly replicated DNA with each endomitotic cycle. This difference in DNA morphology can be used to help determine whether sperm entry has occurred.

Postfertilization defects in meiosis or subsequent cell divisions can be observed in fixed embryos by staining with combinations of DAPI, the M-phase marker anti-phosphohistone H3 (serine 10), and anti- $\beta$ -tubulin DM1A antibodies (Golden, 2000). These defects can also be seen in live embryos by crossing two strains containing either *pie-1::Histone 2B::GFP* (Praitis *et al.*, 2001) or *pie-1::GFP:: $\beta$ -tubulin* (Strome *et al.*, 2001).

EGG-1 and EGG-2 are two proteins in the plasma membrane of the oocyte which are necessary for fertilization (Kadandale *et al.*, 2005b). These genes have proven to be extremely useful diagnostic tools for analyzing egg-sterile mutants with defects in fertilization (Kadandale *et al.*, 2005b). Both genes encode type II transmembrane molecules and their extracellular domains include eight low-density lipoprotein (LDL)-receptor-repeats, which are known to function as receptors for a variety of ligands and mediate multiple cellular responses (Kadandale *et al.*, 2005b; Nykjaer and Willnow, 2002). Sperm are able to make contact with *egg-1/2* deficient oocytes but are unable to fuse with them (Kadandale *et al.*, 2005b). The localization of EGG-1 and EGG-2 can be visualized using GFP-tagged versions of the proteins. GFP:EGG-1 (Fig. 8G) and EGG-2:GFP are localized to the surface of developing oocytes but become undetectable after fertilization, presumably as a result of endocytosis (Kadandale *et al.*, 2005b).

Oocyte maturation, fertilization, and the completion of meiosis are overlapping and dependent processes that are coordinated by a number of important regulatory factors (Marcello and Singson, 2010). Defects in any step can have deleterious effects on subsequent steps and ultimately result in improper embryogenesis. To understand how a specific gene or protein of interest functions in these processes, mutants with defects in specific stages of meiosis or related GFP marked strains can be used. For example, to analyze the relationship between cell cycle progression and egg activation, the activity of the anaphase-promoting complex (APC/C) was blocked using either temperature-sensitive mutants or RNAi. In oocytes depleted of the APC/C subunit *mat-1*, fertilization occurs, but the fertilized oocytes remain locked in a meiotic metaphase I state and subsequent events of the oocyte-to-embryo transition do not occur (Golden *et al.*, 2000). In other studies, cyclin B:GFP levels have been used to indicate meiotic progression. Degradation of maternally supplied cyclin B by the APC/C was shown to drive the transition from metaphase I to anaphase I (Davis *et al.*, 2002; Furuta *et al.*, 2000; Golden *et al.*, 2000; Rahman and Kipreos, 2010); however, in the absence of fertilization, oocytes progress to anaphase I but the drop in their cyclin B levels proved to be incomplete (McNally and



**Fig. 8** Examples of sperm and oocyte protein localization. (A–C) Localization of 1CB4 and SPE-9. (A) Nomarski DIC micrograph of mature spermatozoa. (B) Localization of 1CB4 to MOs. (C) Localization of SPE-9 to pseudopods. (D–F) Colocalization of 1CB4 and SPE-38 in spermatids. (D) Nomarski DIC micrograph of spermatids. (E) Localization of 1CB4 to MOs. (F) Localization of SPE-38 to MOs. Note the identical distribution of staining in panels E and F. (G) GFP fluorescence of a GFP:EGG-1 fusion protein in oocytes.

McNally, 2005). Destruction of the remaining cyclin B was found to require fertilization and passage through the second meiotic division and more specifically, a second E3 ubiquitin complex, Cul-2/Zyg-11, which allows progression to anaphase II (Liu *et al.*, 2004; Sonnevile and Gonczy, 2004). The use of *mat-1*-depleted animals and cyclin B:GFP will help place the functioning of many genes or proteins of interest in the context of meiotic progression. The coordination of the cell cycle, fertilization, and egg activation are still poorly understood. The concurrent and dependent processes of meiotic resumption and fertilization and their subsequent

events likely result in a multitude of branching cellular pathways (Marcello and Singson, 2010; Singson *et al.*, 2008).

Egg-activation mutants are likely to fall into different classes, and appropriate GFP strains and disruption of gene function by RNAi or mutation can be used to analyze genes or proteins of interest in egg activation. Sperm entry promotes the establishment of embryonic polarity and triggers egg activation (Goldstein and Hird, 1996; Sadler and Shakes, 2000). One example of a paternal egg-activation mutant is *spe-11*; when *spe-11* sperm fertilize wild-type oocytes, the oocytes fail to either complete meiosis or secrete an eggshell, but instead form multinucleate one-cell embryos (Browning and Strome, 1996; Hill *et al.*, 1989). Other egg-activation mutants may arise from mutations in maternally required genes that regulate and coordinate egg-activation events (Cheng *et al.*, 2009; Maruyama *et al.*, 2007; Parry *et al.*, 2009; Stitzel *et al.*, 2007). Such genes would include regulatory proteins at the oocyte cortex that couple fertilization with cell-cycle progression through the second meiotic division and the events of egg activation (Govindan and Greenstein, 2007; Maruyama *et al.*, 2007; Parry *et al.*, 2009; Parry and Singson, 2011). For example, EGG-3, EGG-4, and EGG-5 are pseudophosphatases located in the oocyte cortex that coordinate fertilization with cell-cycle resumption by regulating the activity of minibrain kinase-2 (MBK-2) (Cheng *et al.*, 2009; Maruyama *et al.*, 2007; Parry *et al.*, 2009; Stitzel *et al.*, 2007; Stitzel and Seydoux, 2007). MBK-2 helps regulate early embryonic development after fertilization by phosphorylating maternal substrates for degradation. These MBK-2 targets include the katanin subunit MEI-1, the RNA/TAF-4 binding proteins OMA-1 and OMA-2, and the polarity factors MEX-5 and MEX-6 (Detwiler *et al.*, 2001; Guven-Ozkan *et al.*, 2008; Nishi and Lin, 2005; Pang *et al.*, 2004; Pellettieri *et al.*, 2003; Quintin *et al.*, 2003). Appropriate GFP strains can be used to assess whether other egg-activation mutants exhibit defects in the localization, sorting, or degradation of any of these known components.

---

---

---

## XI. Eggshell Production

In response to fertilization and proper egg activation, a multilayered eggshell is formed around the developing embryo to allow for completion of meiosis, polar body extrusion, embryo polarity, and formation of an osmotic barrier. The developing oocyte is covered by a thin vitelline layer that can be observed by electron microscopy or staining with *Malcura pomifera* agglutinin (MPA) or *Griffonia simplicifolia* lectin I (GSL I). At the time of fertilization this vitelline layer begins to separate from the plasma membrane and becomes the outermost layer of the mature eggshell (Bembenek *et al.*, 2007; Rappleye *et al.*, 1999). The second (middle) layer is formed when the oocyte membrane protein CHS-1 catalyses UDP-N-acetylglucosamine polymerization to produce chitin, the material that gives eggshells their mechanical strength. Chitin deposition can be assayed by staining with a rhodamine-conjugated chitin-binding probe. The proteolipid inner

layer is formed immediately before the first zygotic cell division and provides a permeability and osmotic barrier. The fidelity of this osmotic/permeability barrier can be determined by staining with lipophilic fluorescent plasma membrane dye FM 4-64, which can only penetrate and stain embryos that have not yet formed an intact barrier or eggshell.

Unfertilized oocytes and activation-defective eggs are sensitive to osmotic strength due to the lack of a protective eggshell and must be handled with special care. Intact and dissected hermaphrodites should at a minimum be handled in osmotic egg buffers (Edgar, 1995). Otherwise, unfertilized oocytes may burst under osmotic pressure and even intact embryos may swell within a weak eggshell and cause an embryo to be mistakenly scored as cytokinesis defective. In some cases, oocytes are so fragile that dissections from the uterus are impossible. These oocytes must be examined carefully in intact animals (Parry *et al.*, 2009).

CG exocytosis is necessary for proper eggshell formation (Bembenek *et al.*, 2007). CGs transport chondroitin proteoglycans to the extracellular space surrounding the embryo, and can be detected by wheat germ agglutinin (WGA) or the Golgi marker UGTP-1 (Bembenek *et al.*, 2007). CAV-1 is prominent but nonessential component of CGs (Sato *et al.*, 2008). CG exocytosis, which occurs during anaphase I, does not require fertilization. However, it does require a variety of cell-cycle components including the APC/C and separase (*sep-1*) as well as the small GTPase RAB-11 and the target-SNARE SYN-4 (Bembenek *et al.*, 2007). Additional exocytotic events must be required for proper eggshell formation; however, the exact nature of these events remains unclear (Bembenek *et al.*, 2007).

Although CG exocytosis is known to contribute to the membrane-based polyspermy block in other organisms, the connection between CG exocytosis and polyspermy in *C. elegans* is not well understood (Wessel *et al.*, 2001). In *C. elegans*, polyspermy has been observed after the depletion of *chs-1*, *gna-2*, or *egg-4/5* (Parry *et al.*, 2009; Johnston *et al.*, 2010). *gna-2* encodes a GLD-regulated glucosamine-6-P N acetyltransferase that supplies UDP-N-acetyl glucosamine for chitin biosynthesis (Johnston *et al.*, 2006). The deposition of chitin is independent of CG exocytosis and the presence of chitin seems to play a role in the membrane-based polyspermy block (Johnston *et al.*, 2010; Parry *et al.*, 2009; Sato *et al.*, 2006, 2008). Potential defects in the polyspermy block can be assessed by DAPI staining intact hermaphrodites or dissected embryos and looking for evidence of multiple sperm chromatin masses within recently fertilized, meiotic-stage oocytes (Johnston *et al.*, 2010; Parry *et al.*, 2009). However, in intact or poorly dissected animals, it can be difficult to distinguish nuclei within the embryo from those in the surrounding periphery (Johnston *et al.*, 2010; Parry *et al.*, 2009). When staining dissected embryos that do not have eggshells, a grouping of embryos can lack clear demarcation of each embryo boundary. The use of GFP:PH (Plextrin Homology) construct can be used to help visualize the plasma membrane (Audhya *et al.*, 2005) and minimize this problem.

---

---

## XII. Analysis of Fertilization-Specific Gene Products

Antibody-based approaches are useful for determining the subcellular localization of fertilization-specific gene products. The standard approach is to generate polyclonal anti-peptide antibodies against two or more regions of the candidate fertility protein since not all regions will be useful for generating antibodies (Chatterjee *et al.*, 2005; Parry *et al.*, 2009; Zannoni *et al.*, 2003). For these fertility proteins, polyclonal, anti-peptide antibodies have two distinct advantages over monoclonal antibodies: (1) they do not require the isolation of pure sperm and oocytes, (2) polyclonal antibodies typically have higher binding efficiencies, which is useful when attempting to detect cell-surface proteins that are typically expressed at low levels. It is also important always to prescreen the animals used for antibody production since many express unrelated nematode antibodies stemming from previous nematode infections.

It is useful to understand the localization pattern of SPE proteins in both spermatids and spermatozoa. Sperm squashes are used for staining spermatids, and spermatozoa preparations can be obtained either from mated hermaphrodites or from males dissected in non-protease-based sperm activators. Samples should be fixed separately in both paraformaldehyde and cold methanol as membrane proteins can differ in their response to each. *C. elegans* spermatids are small and possess only four major organelles (a central chromatin mass, an associated inactive centriole, numerous mitochondria, and multiple MOs) simplifying the analysis of localization patterns. MO localization may be confirmed using the monoclonal antibody 1CB4 (Okamoto and Thomson, 1985) as was done for the proteins SPE-9 and SPE-38 (Chatterjee *et al.*, 2005; Zannoni *et al.*, 2003) (Fig. 8A–F).

Live cell staining (with antibodies added before fixation) can also be used to assess protein localization on the external surface of the plasma membrane or in fused MOs (Chatterjee *et al.*, 2005). *fer-1* mutants can be used to confirm an initial restriction to the MO since the *fer-1* spermatozoa are specifically defective in MO fusion despite their ability to form a small motile pseudopod (Achanzar and Ward, 1997; Chatterjee *et al.*, 2005; Xu and Sternberg, 2003).

To date, all fertilization-defective SPE proteins have at least partially localized to the pseudopods of spermatozoa, now presumed to be the point of oocyte–sperm contact (Chatterjee *et al.*, 2005; Xu and Sternberg, 2003; Zannoni *et al.*, 2003). In spermatids, some fertilization-defective SPE proteins such as SPE-9 localize to the plasma membrane (Zannoni *et al.*, 2003) while others, such as SPE-38 and TRP-3/SPE-41, localize to the unfused MOs of spermatids. Any patterns of localization common to multiple members of the SPE-9 class remain to be determined (Fig. 8) (Chatterjee *et al.*, 2005).

Fusion proteins (e.g., GFP fusions) can be convenient tools for determining a protein's distribution and dynamics in fixed or live cells. However, the expression of oocyte and sperm fusion proteins is often repressed in the germ line (Kelly *et al.*, 1997; Putiri *et al.*, 2004; Seydoux and Schedl, 2001). Oocyte fusion protein

expression (Fig. 8G) can be enhanced using complex arrays (Kelly *et al.*, 1997), integrated low copy number microparticle bombardment transformation strategies (Praitis *et al.*, 2001), or Mos1 transposon single copy insertion (MosSCI) based methods (Robert and Bessereau, 2010). Reliable expression of sperm fusion proteins has not yet been achieved. For a more in-depth discussion of protein localization studies see the chapter by Hutter.

---

---

---

### XIII. Future Prospects/Issues

This chapter attempts to provide a useful overview of how *C. elegans* can be used as a model system for addressing questions of fertility. A clear understanding of any biological process as complex as fertilization is impossible without a complete inventory of its cellular and molecular components; consequently, each new gene identified adds significantly to our knowledge and insight into fertilization mechanisms.

Although much has already been elucidated via *C. elegans* fertility research, many fundamental questions remain unanswered. For example, how are the events surrounding fertilization (e.g., meiotic maturation, ovulation, fertilization, sperm migratory behavior, and the nature of signaling events) coordinated? What is the mechanism of the block to polyspermy? How are the events of fertilization and egg activation related to early development and patterning of the embryo?

The fundamental question of what happens to sperm and oocyte fertility proteins during the physical joining of the gametes is being actively studied in our laboratories. The investigation of this deceptively simple question has proven challenging since there are at most only two recently fertilized oocytes present in any hermaphrodite at one time, and only those in the very earliest stage of postfertilization meiosis are informative (A. Richmond and D. Shakes, unpublished data). Although difficult, such studies are feasible and will be ultimately useful in analyzing the interactions between sperm and oocyte fertility proteins.

Exciting progress continues to be made and the field is still developing new experimental tools. Calcium imaging, for instance, has been applied to *C. elegans* sperm and oocytes (Samuel *et al.*, 2001; Xu and Sternberg, 2003). Using Calcium Green-1 dextran as an indicator, an increase in cytoplasmic calcium is observed at the same time as fertilization, but the trigger for the calcium release and its functional consequence are unknown (Samuel *et al.*, 2001). Voltage-sensitive reagents or other physiological approaches could be used to test whether the fast block to polyspermy is dependent on depolarization of the oocyte plasma membrane (Yanagimachi, 1994) as in other organisms. New transgenic approaches are helping to overcome the germ line expression repression that currently hampers research in this area (Robert and Bessereau, 2010). *In vitro* fertilization systems in *C. elegans* could be useful for analyzing the function of newly discovered molecules; however,

the important physical interactions between gametes and the somatic germ line suggest that this approach may not be feasible.

Insights gained from the study of fertilization in *C. elegans* will also increase our understanding of diverse reproductive strategies and those mechanisms relevant to molecular evolution and speciation. The recent sequencing projects of other nematode species (Bird *et al.*, 2005; Ghedin *et al.*, 2007; Mitreva *et al.*, 2005) and baseline *spe* gene phenotypic data (Geldziler *et al.*, 2006) should enable this work to progress quickly.

*C. elegans* remains an extremely useful organism with which to study the nature of fertilization. An integrated approach to the analysis of fertilization that combines molecular genetic and cell biological worm techniques with the more traditional antibody-based and biochemical methods will continue to further our understanding of this most fascinating and fundamental process.

### Acknowledgments

The authors thank Lisa Caiafa and members of the Singson Laboratory for helpful discussions, critical comments and advice. The authors also specifically wish to thank Pavan Kadandale, Indrani Chatterjee, and Rika Maruyama for assistance with figures, Penny Sadler for sharing unpublished spermatogenesis insights, and Alissa Richmond for immunostaining micrographs and unpublished sperm–oocyte fusion insights. The Singson Laboratory is supported by a grant from the NIH (R01 HD054681). The Shakes laboratory is supported by a grant from the NIH (R15 GM060359-02) to Diane C. Shakes.

### References

- Achanzar, W. E., and Ward, S. (1997). A nematode gene required for sperm vesicle fusion. *J. Cell Sci.* **110**, 1073–1081.
- Albertson, D. G., and Thomson, J. N. (1993). Segregation of holocentric chromosomes at meiosis in the nematode, *Caenorhabditis elegans*. *Chromosome Res.* **1**(1), 15–26.
- Aroian, R. V., and Field, C., *et al.* (1997). Isolation of actin-associated proteins from *Caenorhabditis elegans* oocytes and their localization in the early embryo. *EMBO J.* **16**(7), 1541–1549.
- Artal-Sanz, M., and Tsang, W. Y., *et al.* (2003). The mitochondrial prohibitin complex is essential for embryonic viability and germline function in *Caenorhabditis elegans*. *J. Biol. Chem.* **278**(34), 32091–32099.
- Audhya, A., and Hyndman, F., *et al.* (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*. *J. Cell Biol.* **171**(2), 267–279.
- Bembenek, J. N., and Richie, C. T., *et al.* (2007). Cortical granule exocytosis in *C. elegans* is regulated by cell cycle components including separase. *Development* **134**(21), 3837–3848.
- Bird, D. M., and Blaxter, M. L., *et al.* (2005). A white paper on nematode comparative genomics. *J. Nematol.* **37**(4), 408–416.
- Browning, H., and Strome, S. (1996). A sperm-supplied factor required for embryogenesis in *C. elegans*. *Development* **122**, 391–404.
- Burrows, A. E., and Scurman, B. K., *et al.* (2006). The *C. elegans* Myt1 ortholog is required for the proper timing of oocyte maturation. *Development* **133**(4), 697–709.

- Chatterjee, I., and Richmond, A., *et al.* (2005). The *Caenorhabditis elegans spe-38* gene encodes a novel four-pass integral membrane protein required for sperm function at fertilization. *Development* **132**, 2795–2808.
- Cheng, K. C., and Klancer, R., *et al.* (2009). Regulation of MBK-2/DYRK by CDK-1 and the pseudophosphatases EGG-4 and EGG-5 during the oocyte-to-embryo transition. *Cell* **139**(3), 560–572.
- Church, D. L., and Guan, K. L., *et al.* (1995). Three genes of the MAP kinase cascade, *mek-2*, *mpk-1/sur-1* and *let-60 ras*, are required for meiotic cell cycle progression in *Caenorhabditis elegans*. *Development* **121**(8), 2525–2535.
- Davis, E. S., and Wille, L., *et al.* (2002). Multiple subunits of the *Caenorhabditis elegans* anaphase-promoting complex are required for chromosome segregation during meiosis I. *Genetics* **160**(2), 805–813.
- Detwiler, M. R., and Reuben, M., *et al.* (2001). Two zinc finger proteins, OMA-1 and OMA-2, are redundantly required for oocyte maturation in *C. elegans*. *Dev. Cell* **1**(2), 187–199.
- Doniach, T., and Hodgkin, J. (1984). A sex-determining gene, *fem-1*, required for both male and hermaphrodite development in *Caenorhabditis elegans*. *Dev. Biol.* **106**(1), 223–235.
- Doree, M., and Hunt, T. (2002). From Cdc2 to Cdk1: When did the cell cycle kinase join its cyclin partner? *J. Cell Sci.* **115**(Pt 12), 2461–2464.
- Edgar, L. G. (1995). Blastomere culture and analysis. *Methods Cell Biol.* **48**, 303–321.
- Edgley, M. L., and Baillie, D. L., *et al.* (1995). Genetic balancers. *Methods Cell Biol.* **48**, 147–184.
- Edgley, M. L., Baillie, D. L., *et al.* (2006). Genetic Balancers. *WormBook* ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.7.1, <http://www.wormbook.org>.
- Ensslin, M. A., and Shur, B. D. (2003). Identification of mouse sperm SED1, a bimotif EGF repeat and discoidin-domain protein involved in sperm–egg binding. *Cell* **114**(4), 405–417.
- Fay, D. (2006). Genetic Mapping and Manipulation: Chapter 1 – Introduction and Basics. *WormBook* ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.7.1, <http://www.wormbook.org>.
- Fire, A., and Xu, S., *et al.* (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- Furuta, T., and Tuck, S., *et al.* (2000). EMB-30: an APC4 homologue required for metaphase-to-anaphase transitions during meiosis and mitosis in *Caenorhabditis elegans*. *Mol. Biol. Cell* **11**(4), 1401–1419.
- Geldziler, B., and Chatterjee, I., *et al.* (2006). A comparative study of sperm morphology, cytology and activation in *Caenorhabditis elegans*, *Caenorhabditis remanei* and *Caenorhabditis briggsae*. *Dev. Genes Evol.* **216**(4), 198–208.
- Geldziler, B., and Chatterjee, I., *et al.* (2005). The genetic and molecular analysis of *spe-19*, a gene required for sperm activation in *Caenorhabditis elegans*. *Dev. Biol.* **283**(2), 424–436.
- Geldziler, B., and Kadandale, P., *et al.* (2004). Molecular genetic approaches to studying fertilization in model systems. *Reproduction* **127**(4), 409–416.
- Ghedini, E., and Wang, S., *et al.* (2007). Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* **317**(5845), 1756–1760.
- Golden, A. (2000). Cytoplasmic flow and the establishment of polarity in *C. elegans* 1-cell embryos. *Curr. Opin. Genet. Dev.* **10**(4), 414–420.
- Golden, A., and Sadler, P. L., *et al.* (2000). Metaphase to anaphase (mat) transition-defective mutants in *Caenorhabditis elegans*. *J. Cell Biol.* **151**(7), 1469–1482.
- Goldstein, B., and Hird, S. N. (1996). Specification of the anteroposterior axis in *Caenorhabditis elegans*. *Development* **122**, 1467–1474.
- Govindan, J. A., and Greenstein, D. (2007). Embryogenesis: Anchors away!. *Curr. Biol.* **17**(20), R890–R892.
- Grant, B., and Hirsh, D. (1999). Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. *Mol. Biol. Cell* **10**(12), 4311–4326.
- Gumienny, T. L., and Lambie, E., *et al.* (1999). Genetic control of programmed cell death in the *Caenorhabditis elegans* hermaphrodite germline. *Development* **126**(5), 1011–1022.

- Güven-Ozkan, T., and Nishi, Y., *et al.* (2008). Global transcriptional repression in *C. elegans* germline precursors by regulated sequestration of TAF-4. *Cell* **135**(1), 149–160.
- Hill, D. P., and Shakes, D. C., *et al.* (1989). A sperm-supplied product essential for initiation of normal embryogenesis in *Caenorhabditis elegans* is encoded by the paternal-effect embryonic-lethal gene, *spe-11*. *Dev. Biol.* **136**(1), 154–166 [published erratum appears in *Dev Biol* 1990 May;139(1):230].
- Hill, K.L., and L'Hernault, S. W. (2001). Analyses of reproductive interactions that occur after hetero-specific matings within the genus *Caenorhabditis*. *Dev. Biol.* **232**(1), 105–114.
- Hodgkin, J. (1997). Appendix 1 Genetics. In “*C. elegans* II,” (D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, eds.), pp. 881–1048. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Hodgkin, J. (1999). Conventional Genetics. *C. elegans: A Practical Approach*. I. Hope, pp. 245–270. Oxford University Press, New York.
- Horner, V. L., and Wolfner, M. F. (2008). Transitioning from egg to embryo: Triggers and mechanisms of egg activation. *Dev. Dyn.* **237**(3), 527–544.
- Hsu, J. Y., and Sun, Z. W., *et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* **102**(3), 279–291.
- Iwasaki, K., and McCarter, J., *et al.* (1996). *emo-1*, a *Caenorhabditis elegans* Sec61p gamma homologue, is required for oocyte development and ovulation. *J. Cell Biol.* **134**(3), 699–714.
- Jansen, G., and Hazendonk, E., *et al.* (1997). Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* **17**(1), 119–121.
- Jaramillo-Lambert, A., and Ellefson, M., *et al.* (2007). Differential timing of S phases, X chromosome replication, and meiotic prophase in the *C. elegans* germ line. *Dev. Biol.* **308**(1), 206–221.
- Johnston, W. L., and Krizus, A., *et al.* (2006). The eggshell is required for meiotic fidelity, polar-body extrusion and polarization of the *C. elegans* embryo. *BMC Biol.* **4**, 35.
- Johnston, W. L., and Krizus, A., *et al.* (2010). Eggshell chitin and chitin-interacting proteins prevent polyspermy in *C. elegans*. *Curr. Biol.* **20**(21), 1932–1937.
- Jorgensen, E. M., and Mango, S. E. (2002). The art and design of genetic screens: *Caenorhabditis elegans*. *Nat. Rev. Genet.* **3**(5), 356–369.
- Kadandale, P., and Geldziler, B., *et al.* (2005a). Use of SNPs to determine the breakpoints of complex deficiencies, facilitating gene mapping in *Caenorhabditis elegans*. *BMC Genet.* **6**(1), 28.
- Kadandale, P., and Singson, A. (2004). Oocyte production and sperm utilization patterns in semi-fertile strains of *Caenorhabditis elegans*. *BMC Dev. Biol.* **4**(1), 3.
- Kadandale, P., and Stewart-Michaelis, A., *et al.* (2005b). The egg surface LDL receptor repeat-containing proteins EGG-1 and EGG-2 are required for fertilization in *Caenorhabditis elegans*. *Curr. Biol.* **15**(24), 2222–2229.
- Kamath, R. S., and Ahringer, J. (2003). Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**(4), 313–321.
- Kelly, W. G., and Xu, S., *et al.* (1997). Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238.
- Kimble, J., and Ward, S. (1988). Germ-Line Development and Fertilization. In “The Nematode *Caenorhabditis elegans*,” (and W. B. Wood, ed.), pp. 191–213. Cold Spring Harbor Laboratory, Cold Spring Harbor.
- Kosinski, M., and McDonald, K., *et al.* (2005). *C. elegans* sperm bud vesicles to deliver a meiotic maturation signal to distant oocytes. *Development* **132**(15), 3357–3369.
- Kroft, T. L., and Gleason, E. J., *et al.* (2005). The *spe-42* gene is required for sperm–egg interactions during *C. elegans* fertilization and encodes a sperm-specific transmembrane protein. *Dev. Biol.* **286**, 169–181.
- Kubagawa, H. M., and Watts, J. L., *et al.* (2006). Oocyte signals derived from polyunsaturated fatty acids control sperm recruitment *in vivo*. *Nat. Cell Biol.* **8**(10), 1143–1148.
- L'Hernault, S. W. (1997). Spermatogenesis. In “*C. Elegans* II,” (D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, eds.), pp. 271–294. Cold Spring Harbor Laboratory, Cold Spring Harbor.

- L'Hernault, S. W. (2006). Spermatogenesis. *WormBook: Online Review of C. elegans Biology*. T. C. e. R. Community, <http://www.wormbook.org>.
- L'Hernault, S. W., and Roberts, T. M. (1995). Cell biology of nematode sperm. *Methods Cell Biol.* **48**, 273–301.
- L'Hernault, S. W., and Shakes, D. C., *et al.* (1988). Developmental genetics of chromosome I spermatogenesis-defective mutants in the nematode *Caenorhabditis elegans*. *Genetics* **120**, 435–452.
- LaMunyon, C. W., and Ward, S. (1995). Sperm precedence in a hermaphroditic nematode *Caenorhabditis elegans* is due to competitive superiority of male sperm. *Experientia* **51**(8), 817–823.
- LaMunyon, C. W., and Ward, S. (1997). Increased competitiveness of nematode sperm bearing the male X chromosome. *Proc. Natl Acad. Sci. U. S. A.* **94**(1), 185–189.
- LaMunyon, C. W., and Ward, S. (1998). Larger sperm outcompete smaller sperm in the nematode *Caenorhabditis elegans*. *Proc. R. Soc. Ser. B* **265**, 1997–2002.
- LaMunyon, C. W., and Ward, S. (1999). Evolution of sperm size in nematodes: Sperm competition favours larger sperm. *Proc. R. Soc. Ser. B* **266**(1416), 263–267.
- Lee, M. H., and Schedl, T. (2001). Identification of *in vivo* mRNA targets of GLD-1, a maxi-KH motif containing protein required for *C. elegans* germ cell development. *Genes Dev.* **15**(18), 2408–2420.
- Lessard, C., and Pendola, J. K., *et al.* (2004). New mouse genetic models for human contraceptive development. *Cytogenet Genome Res.* **105**(2–4), 222–227.
- Lin, R. (2003). A gain-of-function mutation in *oma-1*, a *C. elegans* gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. *Dev. Biol.* **258**(1), 226–239.
- Liu, J., and Vasudevan, S., *et al.* (2004). CUL-2 and ZYG-11 promote meiotic anaphase II and the proper placement of the anterior–posterior axis in *C. elegans*. *Development* **131**(15), 3513–3525.
- Liu, L. X., and Spoerke, J. M., *et al.* (1999). High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.* **9**(9), 859–867.
- Maeda, I., and Kohara, Y., *et al.* (2001). Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.* **11**(3), 171–176.
- Marcello, M. R., and Singson, A. (2010). Fertilization and the oocyte-to-embryo transition in *C. elegans*. *BMB Rep.* **43**(6), 389–399.
- Maruyama, R., and Velarde, N. V., *et al.* (2007). EGG-3 regulates cell-surface and cortex rearrangements during egg activation in *Caenorhabditis elegans*. *Curr. Biol.* **17**(18), 1555–1560.
- McCarter, J., and Bartlett, B., *et al.* (1999). On the control of oocyte meiotic maturation and ovulation in *C. elegans*. *Dev. Biol.* **205**, 111–128.
- McNally, K., and Audhya, A., *et al.* (2006). Katanin controls mitotic and meiotic spindle length. *J. Cell Biol.* **175**(6), 881–891.
- McNally, K. L., and McNally, F. J. (2005). Fertilization initiates the transition from anaphase I to metaphase II during female meiosis in *C. elegans*. *Dev. Biol.* **282**(1), 218–230.
- Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biol.* **48**, 451–482.
- Miller, M. A. (2006). Sperm and oocyte isolation methods for biochemical and proteomic analysis. *Methods Mol. Biol.* **351**, 193–201.
- Miller, M. A., and Nguyen, V. Q., *et al.* (2001). A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation. *Science* **291**(5511), 2144–2147.
- Miller, M. A., and Ruest, P. J., *et al.* (2003). An Eph receptor sperm-sensing control mechanism for oocyte meiotic maturation in *Caenorhabditis elegans*. *Genes Dev.* **17**(2), 187–200.
- Mitreva, M., and Blaxter, M. L., *et al.* (2005). Comparative genomics of nematodes. *Trends Genet.* **21**(10), 573–581.
- Nelson, G. A., and Ward, S. (1980). Vesicle fusion, pseudopod extension and amoeboid motility are induced in nematode spermatids by the ionophore monensin. *Cell* **19**(2), 457–464.
- Nishi, Y., and Lin, R. (2005). DYRK2 and GSK-3 phosphorylate and promote the timely degradation of OMA-1, a key regulator of the oocyte-to-embryo transition in *C. elegans*. *Dev. Biol.* **288**(1), 139–149.
- Nishimura, H., and L'Hernault, S. W. (2010). Spermatogenesis-defective (*spe*) mutants of the nematode *Caenorhabditis elegans* provide clues to solve the puzzle of male germline functions during reproduction. *Dev. Dyn.* **239**(5), 1502–1514.

- Nykjaer, A., and Willnow, T. E. (2002). The low-density lipoprotein receptor gene family: A cellular Swiss army knife? *Trends Cell Biol.* **12**(6), 273–280.
- O’Connell, K. F., and Leys, C. M., *et al.* (1998). A genetic screen for temperature-sensitive cell-division mutants of *Caenorhabditis elegans*. *Genetics* **149**(3), 1303–1321.
- Okamoto, H., and Thomson, J. N. (1985). Monoclonal antibodies which distinguish certain classes of neuronal and supporting cells in the nervous tissue of the nematode *Caenorhabditis elegans*. *J. Neurosci.* **5**(3), 643–653.
- Pang, K. M., and Ishidate, T., *et al.* (2004). The minibrain kinase homolog, mbk-2, is required for spindle positioning and asymmetric cell division in early *C. elegans* embryos. *Dev. Biol.* **265**(1), 127–139.
- Parry, J. M., and Singson, A. (2011). EGG molecules couple the oocyte-to-embryo transition with cell cycle progression. *Results Probl. Cell Differ.* **53**, 135–151.
- Parry, J. M., and Velarde, N. V., *et al.* (2009). EGG-4 and EGG-5 link events of the oocyte-to-embryo transition with meiotic progression in *C. elegans*. *Curr. Biol.* **19**(20), 1752–1757.
- Pellettieri, J., and Reinke, V., *et al.* (2003). Coordinate activation of maternal protein degradation during the egg-to-embryo transition in *C. elegans*. *Dev. Cell* **5**(3), 451–462.
- Praitis, V., and Casey, E., *et al.* (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**(3), 1217–1226.
- Primakoff, P., and Myles, D. G. (2002). Penetration, adhesion, and fusion in mammalian sperm–egg interaction. *Science* **296**(5576), 2183–2185.
- Putiri, E., and Zannoni, S., *et al.* (2004). Functional domains and temperature-sensitive mutations in SPE-9, an EGF repeat-containing protein required for fertility in *Caenorhabditis elegans*. *Dev. Biol.* **272**(2), 448–459.
- Quintin, S., and Mains, P. E., *et al.* (2003). The mbk-2 kinase is required for inactivation of MEI-1/katanin in the one-cell *Caenorhabditis elegans* embryo. *EMBO Rep.* **4**(12), 1175–1181.
- Rahman, M. M., and Kipreos, E. T. (2010). The specific roles of mitotic cyclins revealed. *Cell Cycle* **9**(1), 22–23.
- Rappleye, C. A., and Paredes, A. R., *et al.* (1999). The coronin-like protein POD-1 is required for anterior–posterior axis formation and cellular architecture in the nematode *Caenorhabditis elegans*. *Genes Dev.* **13**(21), 2838–2851.
- Reinke, V., and Cutter, A. D. (2009). Germline expression influences operon organization in the *Caenorhabditis elegans* genome. *Genetics* **181**(4), 1219–1228.
- Reinke, V., and Gil, I. S., *et al.* (2004). Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. *Development* **131**(2), 311–323.
- Ringstad, N., and Horvitz, H. R. (2008). FMRFamide neuropeptides and acetylcholine synergistically inhibit egg-laying by *C. elegans*. *Nat. Neurosci.* **11**(10), 1168–1176.
- Robert, V. J., and Bessereau, J. L. (2010). Manipulating the *Caenorhabditis elegans* genome using mariner transposons. *Genetica* **138**(5), 541–549.
- Roberts, T. M., and Pavalko, F. M., *et al.* (1986). Membrane and cytoplasmic proteins are transported in the same organelle complex during nematode spermatogenesis. *J. Cell Biol.* **102**, 1787–1796.
- Sadler, P. L., and Shakes, D. C. (2000). Anucleate *Caenorhabditis elegans* sperm can crawl, fertilize oocytes and direct anterior–posterior polarization of the 1-cell embryo. *Development* **127**(2), 355–366.
- Samuel, A. D., and Murthy, V. N., *et al.* (2001). Calcium dynamics during fertilization in *C. elegans*. *BMC Dev. Biol.* **1**, 8.
- Sato, K., and Sato, M., *et al.* (2006). Dynamic regulation of caveolin-1 trafficking in the germ line and embryo of *Caenorhabditis elegans*. *Mol. Biol. Cell* **17**(7), 3085–3094.
- Sato, M., and Grant, B. D., *et al.* (2008). Rab11 is required for synchronous secretion of chondroitin proteoglycans after fertilization in *Caenorhabditis elegans*. *J. Cell Sci.* **121**(Pt 19), 3177–3186.
- Sawada, H., and Tanaka, E., *et al.* (2004). Self/nonself recognition in ascidian fertilization: Vitelline coat protein HrVC70 is a candidate allorecognition molecule. *Proc. Natl Acad. Sci. U. S. A.* **101**(44), 15615–15620.
- Schmitt, A., and Nebreda, A. R. (2002). Signalling pathways in oocyte meiotic maturation. *J. Cell Sci.* **115** (Pt 12), 2457–2459.

- Seydoux, G., and Schedl, T. (2001). The germline in *C. elegans*: Origins, proliferation, and silencing. *Int. Rev. Cytol.* **203**, 139–185.
- Shakes, D. C., and Ward, S. (1989). Initiation of spermiogenesis in *C. elegans*: A pharmacological and genetic analysis. *Dev. Biol.* **134**, 189–200.
- Shakes, D. C., and Wu, J. C., *et al.* (2009). Spermatogenesis-specific features of the meiotic program in *Caenorhabditis elegans*. *PLoS Genet.* **5**(8), e1000611.
- Singaravelu, G., and Chatterjee, I., *et al.* (2011). Isolation and in vitro activation of *Caenorhabditis elegans* sperm. *J. Vis. Exp.* **47**.
- Singson, A. (2001). Every sperm is sacred: Fertilization in *Caenorhabditis elegans*. *Dev. Biol.* **230**(2), 101–109.
- Singson, A. (2006). Sperm activation: Time and tide wait for no sperm. *Curr. Biol.* **16**(5), R160–R162.
- Singson, A., and Hang, J. S., *et al.* (2008). Genes required for the common miracle of fertilization in *Caenorhabditis elegans*. *Int. J. Dev. Biol.* **52**(5–6), 647–656.
- Singson, A., and Hill, K. L., *et al.* (1999). Sperm competition in the absence of fertilization in *Caenorhabditis elegans*. *Genetics* **152**(1), 201–208.
- Singson, A., and Mercer, K. B., *et al.* (1998). The *C. elegans spe-9* gene encodes a sperm transmembrane protein that contains EGF-like repeats and is required for fertilization. *Cell* **93**, 71–79.
- Singson, A., and Zannoni, S., *et al.* (2001). Molecules that function in the steps of fertilization. *Cytokine Growth Factor Rev.* **12**(4), 299–304.
- Sonneville, R., and Gonczy, P. (2004). *Zyg-11* and *cul-2* regulate progression through meiosis II and polarity establishment in *C. elegans*. *Development* **131**(15), 3527–3543.
- Stanfield, G. M., and Villeneuve, A. M. (2006). Regulation of sperm activation by SWM-1 is required for reproductive success of *C. elegans* males. *Curr. Biol.*
- Stitzel, M. L., and Cheng, K. C., *et al.* (2007). Regulation of MBK-2/Dyrk kinase by dynamic cortical anchoring during the oocyte-to-zygote transition. *Curr. Biol.* **17**(18), 1545–1554.
- Stitzel, M. L., and Seydoux, G. (2007). Regulation of the oocyte-to-zygote transition. *Science* **316**(5823), 407–408.
- Strome, S., and Powers, J., *et al.* (2001). Spindle dynamics and the role of gamma-tubulin in early *Caenorhabditis elegans* embryos. *Mol. Biol. Cell* **12**(6), 1751–1764.
- Sumiyoshi, E., and Sugimoto, A., *et al.* (2002). Protein phosphatase 4 is required for centrosome maturation in mitosis and sperm meiosis in *C. elegans*. *J. Cell Sci.* **115**(Pt 7), 1403–1410.
- Tenenhaus, C., and Subramaniam, K., *et al.* (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of *Caenorhabditis elegans*. *Genes Dev.* **15**(8), 1031–1040.
- Timmons, L., and Court, D. L., *et al.* (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**(1–2), 103–112.
- Timmons, L., and Fire, A. (1998). Specific interference by ingested dsRNA. *Nature* **395**(6705), 854.
- Wakimoto, B. T., and Lindsley, D. L., *et al.* (2004). Toward a comprehensive genetic analysis of male fertility in *Drosophila melanogaster*. *Genetics* **167**(1), 207–216.
- Ward, S., and Argon, Y., *et al.* (1981). Sperm morphogenesis in wild-type and fertilization-defective mutants of *Caenorhabditis elegans*. *J. Cell Biol.* **91**(1), 26–44.
- Ward, S., and Carrel, J. S. (1979). Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **73**, 304–321.
- Wassarman, P. M. (1999). Mammalian fertilization: Molecular aspects of gamete adhesion, exocytosis, and fusion. *Cell* **96**, 175–183.
- Wessel, G. M., and Brooks, J. M., *et al.* (2001). The biology of cortical granules. *Int. Rev. Cytol.* **209**, 117–206.
- Wu, T. F., and Nera, B., *et al.* (2010). Elucidating gene regulatory mechanisms for sperm function through the integration of classical and systems approaches in *C. elegans*. *Syst. Biol. Reprod. Med.* **56**(3), 222–235.
- Xu, X. Z., and Sternberg, P. W. (2003). A *C. elegans* sperm TRP protein required for sperm–egg interactions during fertilization. *Cell* **114**(3), 285–297.

- Yamamoto, I., and Kosinski, M. E., *et al.* (2006). Start me up: Cell signaling and the journey from oocyte to embryo in *C. elegans*. *Dev. Dyn.* **235**(3), 571–585.
- Yanagimachi, R. (1994). Mammalian Fertilization. In “The Physiology of Reproduction,” (E. Knobil, and J. D. Neill, eds.), pp. 189–317. Raven Press, New York.
- Zannoni, S., and L’Hernault, S. W., *et al.* (2003). Dynamic localization of SPE-9 in sperm: A protein required for sperm–oocyte interactions in *Caenorhabditis elegans*. *BMC Dev. Biol.* **3**(1), 10.

---

---

## CHAPTER 14

# Imaging Embryonic Morphogenesis in *C. elegans*

**Jeff Hardin**

Department of Zoology, University of Wisconsin-Madison, Madison, Wisconsin, USA

---

Abstract

- I. Introduction
- II. Methods
  - A. Mounting Embryos for Imaging Morphogenesis
  - B. 4D Nomarski Imaging of Morphogenesis
  - C. Fluorescence Imaging of Morphogenesis
  - D. Correlative Fluorescence and TEM During Morphogenesis (see Sims and Hardin, 2007 for more details)
- III. Summary
- Acknowledgments
- References

---

---

### Abstract

The *Caenorhabditis elegans* embryo is well suited to morphogenetic analysis via modern microscopy, due to its short generation time, transparency, invariant lineage, and the ability to generate transgenic embryos expressing various fluorescent proteins. This chapter provides an overview of microscopy techniques for imaging embryonic morphogenesis, including making agar mounts, capturing four-dimensional (4D) data using Nomarski microscopy, imaging of actin in embryos, factors important for optimizing 4D fluorescence microscopy, and recent techniques that leverage fluorescence microscopy for intracellular imaging of cellular components during morphogenesis.

---

---

## I. Introduction

*Caenorhabditis elegans* has obvious advantages that make it well suited for analyzing morphogenesis of living embryos. Its organizational simplicity, transparency, and essentially invariant development enabled the determination of the complete embryonic cell lineage (Sulston *et al.*, 1983). Such invariant development allows the assessment of mutant phenotypes at the level of single cells in *C. elegans*. The wild-type embryonic lineage was originally determined by direct observation using Nomarski microscopy. This was a very slow process, since only one or two of the >500 total embryonic cells could be followed per embryo. More recently, the use of histone::GFP and other technologies, along with automated analysis, has streamlined lineaging even further. These developments have been discussed elsewhere (Murray *et al.*, 2008; see the chapter by Cowan and Chisholm, this volume). The analysis of postmitotic movements of cells in embryos has likewise benefitted from technological advances. The advent of four-dimensional (4D) microscopy made analysis of morphogenesis much more practical by using computer-controlled equipment to record development of embryos in three dimensions over time. Simultaneous software advances made analysis of 4D movies practical (Thomas *et al.*, 1996).

The fundamental concepts of microscopy that apply to any context in *C. elegans* are covered elsewhere in this volume (see the chapter by Maddox and Maddox, this volume). In this chapter, we focus on the uses of modern microscopy specifically for imaging later morphogenesis in *C. elegans* embryos, after many embryonic cells have undergone their terminal divisions. We describe preparation of standard agar mounts and other approaches for immobilizing embryos and treating embryos prior to performing routine 4D microscopy, discuss simple methods for capturing 4D movies, and discuss various probes for imaging fluorescently tagged cells or structures in living embryos. We also describe one way of performing correlative fluorescence and transmission electron microscopy (F-TEM) during embryonic morphogenesis.

Since the most dramatic movements occur among hypodermal cells during this period of embryonic development, much of this chapter focuses on techniques that are particularly useful for analysis of hypodermal cells. Because the terms “hypodermal” and “epidermal” are used interchangeably, this chapter will use the latter for better consistency with standard usage in other organisms (see Chisholm and Hardin, 2005 for discussion).

---

---

## II. Methods

### A. Mounting Embryos for Imaging Morphogenesis

#### 1. Agar Mount (Modified from Heid and Hardin, 2000)

The agar mount is a simple way to prepare *C. elegans* embryos for microscopy. The basic technique is presented elsewhere in this volume (see the chapter by

Maddox and Maddox). Here we provide additional details regarding this key technique as it relates to analyzing morphogenesis. Agar mounts have several key advantages for analyzing morphogenesis. First, the mount slightly compresses the embryo, holding it in place. Second, such compression produces a consistent orientation convenient for imaging many aspects of embryonic morphogenesis. As the processes of morphogenesis proceed, either the dorsal or ventral surface of the embryo will be against the coverslip. After ventral enclosure is complete, the embryo then turns on its side, such that every embryo will be positioned with either its right or left side facing the coverslip. For many morphogenetic events, especially those involving the embryonic epidermis, such mounts are very useful (see Chisholm and Hardin, 2005 for details of the basic movements associated with epidermal morphogenesis). For some events, other orientations of the embryo may be preferable, and for these purposes, other mounting techniques may be used (see below).

### Imaging setup

For assembling the mount, a standard stereomicroscope is required. To identify early embryos (1–4 cell), a total zoom of 80× or greater is recommended. We have typically used either a Wild MZ5 microscope with 20× oculars or Leica MZ12.5 microscope with 16× oculars.

### Materials

#### i. Reagents:

Agar (5% w/v)

M9 buffer:

3 g  $\text{KH}_2\text{PO}_4$

6 g  $\text{Na}_2\text{HPO}_4$

5 g NaCl

1 mL 1 M  $\text{MgSO}_4$

1 L  $\text{H}_2\text{O}$

Valap:

Equal parts by volume of vaseline, lanolin, and paraffin.

Heat thoroughly until melted and mix.

#### ii. Equipment:

Calibrated glass pipettes (50  $\mu\text{L}$ )

Coverslips (18 × 18 mm, No. 1)

Eyelash brush (eyelash glued to end of round toothpick)

Mouth pipette 15" aspirator tube assembly

Microscope slides (25 × 75 × 1 mm)

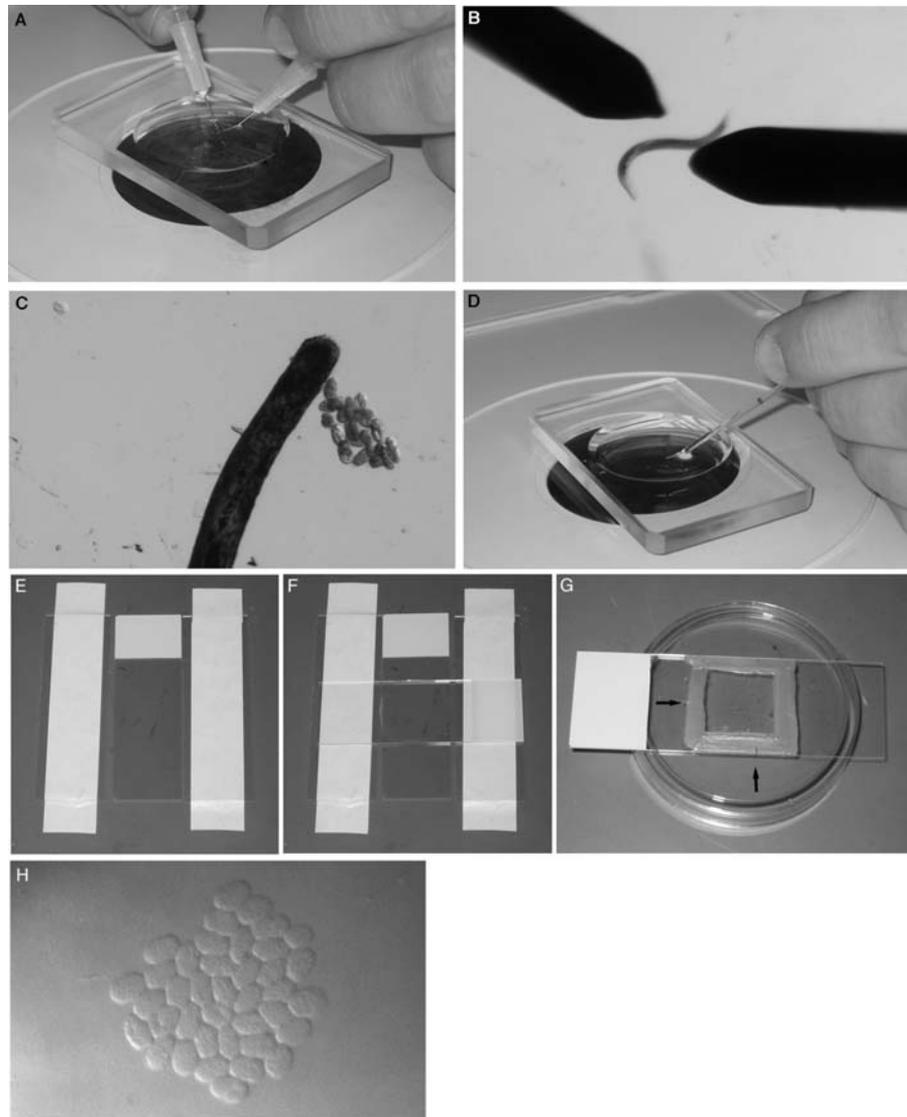
Platinum wire pick: 2.5 cm of 30 gauge platinum or 90% platinum/10% iridium wire inserted into a 6" Pasteur pipette and heated in a flame until the glass melts around the wire. Flat-end hobby pliers or a small tack hammer can be used to flatten the end of the pick.

Single-depression microscope slide (3 mm)

Syringes (1 cc) with 27 gauge  $\times \frac{1}{2}$ " needles

## Method

- i. Use platinum wire pick to move approximately five gravid *C. elegans* hermaphrodites from culture dish to single-depression microscope slide containing M9 buffer. The number of hermaphrodites needed will depend on several factors, including the number of embryos of an appropriate age desired and the gravidity of the worms being used.
- ii. Holding one syringe and needle in each hand, place one on either side of a hermaphrodite and draw flat sides of tips of needles across each other to cut the worm in half transversely. The embryos will be released from the halves of the hermaphrodite. Use eyelash brush to carefully prod halves to expel any remaining embryos. It is important to cut as close to the vulva as possible to release newly fertilized embryos in the uterus. This step can also be conducted by cutting the worm in half with a #15 curved blade scalpel (Fig. 1).
- iii. Sort embryos using eyelash brush and brush together into group of approximately ten embryos. Embryos will tend to stick slightly to each other when grouped. If one desires a certain stage of embryogenesis, it is at this point that embryo stage should be assessed and sorted appropriately. Two-cell stage embryos are the easiest developmental stage to collect.
- iv. Using colored laboratory label tape, tape two microscope slides parallel and one slide width apart on the laboratory bench. Place a third slide between the two taped slides. Using a 6" Pasteur pipette, place three to four drops of molten 5% agar onto the middle slide. Immediately lay fourth slide perpendicular to other three slides over agar and press over taped slides to flatten agar before it cools.
- v. Once agar has set up, use a razor blade to trim excess agar from edges of the slides. Solidified agar in the slide "sandwich" can be left assembled until embryos are ready to be added to the pad. Carefully slide the untaped slides apart so agar pad is left in center of one slide.
- vi. Heat glass 50  $\mu$ L pipette in flame. Once glass is soft and fluid, remove from flame and quickly pull apart ends. Break two ends apart to create a pipette with a tapered end with a diameter of approximately 40  $\mu$ m. Place pipette in mouth pipette aspirator.
- vii. When ready, carefully slide the untaped slides apart so agar pad is left in the center of one slide. Using mouth pipette, transfer the grouping of embryos (from step iii) and approximately 20  $\mu$ L of M9 to the corner of agar pad on the microscope slide.



**Fig. 1** Making a standard agar mount. (A, B) Gravid hermaphrodites are cut in half with  $27 \times \frac{1}{2}$ " needles. (C) At a higher magnification, embryos are sorted and grouped using an eyelash. (D) Embryos and M9 buffer are transferred using a mouth pipette. (E) Three slides are placed on the bench and the outer two are taped down to the bench. (F) A drop of molten 5% agar is placed onto the middle slide. A fourth slide is then placed perpendicular to the three original slides. The top slide is compressed over the taped slides. (G) The finished slide is sealed with VALAP. Using a toothpick to make hash marks in the VALAP (arrows) aids in finding the grouping of embryos on the compound microscope. (H) A low-magnification view of embryos within a completed mount. Embryos are grouped tightly together (image courtesy of T. Loveless).

- viii. Brush embryos out of M9 into the center of slide using eyelash. Position embryos in a single layer side-by-side. We find that especially in the case of embryos in which only a percentage show a phenotype of interest (e.g., homozygous mutant progeny from heterozygous mutant mothers, weak RNAi, etc.), that a large contiguous grouping of embryos is useful (Fig. 1H).
- ix. Set the edge of a coverslip at the side of the agar pad opposite the M9 and slowly drop so that the coverslip lands on the embryos prior to the M9. Use a Kimwipe to wick excess buffer from edges of coverslip and wick air bubbles from under coverslip.
- x. Trim excess agar from edges of coverslip using a razor blade. Seal edges of the coverslip with melted VALAP using a paintbrush.

### Troubleshooting

- i. Problem: Embryos fail to develop.  
Solution: One-cell embryos are especially vulnerable to mechanical stress and are challenging to mount without killing. If studying a later stage of development, the likelihood of embryos surviving is markedly increased if two-cell or later-stage embryos are used to make the mount. Groupings larger than 15–20 embryos often display increased lethality due to oxygen starvation. By keeping groupings of embryos to less than 20 embryos, oxygen starvation should not be a problem.
- ii. Problem: The agar pad dries on the slide before it can be used.  
Solution: Make the pad immediately before use. Stereomicroscopes with light sources mounted under the stage have the potential to heat the stage after long use, which can quickly dry agar pads. Using a stereomicroscope with an external bulb or a cool temperature bulb will reduce this problem.
- iii. Problem: When coverslip is placed on slide, all the embryos wash to the edge of the coverslip.  
Solution: Too much M9 buffer is used and the M9 buffer is hitting the embryos before the coverslip can land on them and hold them in the agar.
- iv. Problem: The slide has air bubbles under the coverslip.  
Solution: Use more M9 buffer. This will allow M9 buffer to completely wash under coverslip. However, too much M9 buffer will cause embryos to wash away (see previous Problem).

### Discussion

Mounting *C. elegans* embryos on agar mounts provides a stable, long-term environment for microscopic analysis of development. The slight compression from the coverslip will result in embryos reproducibly positioned with either the left or right side facing toward the objective lens. During later stages of embryogenesis embryos turn such that left-hand views become dorsal views and right-side views become ventral views. Embryos on agar mounts will survive and hatch from the eggshell on

the mount. Embryos prepared with an agar mount are amenable to both light microscopy (with differential interference contrast optics) or confocal microscopy. Preparing *C. elegans* embryos on an agar mount is a simple technique that can easily be mastered and is regularly done by undergraduates. It provides a consistent embryonic orientation and environment that is suitable for long-term microscopy of *C. elegans* embryos.

## 2. Other Mounting Methods

For many morphogenetic events, agar mounts are convenient because they produce uniform orientation of developing embryos. However, there may be times when more randomized orientations are desired. Examples include imaging of the anterior of the embryo during head enclosure, or events during gastrulation when an *en face* view is desired. For these cases, other mounts are more useful. We discuss two types here. Because protocols for producing these mounts are published elsewhere, we only briefly mention them here.

a. **Simple poly-L-lysine mount** (see Mohler and Isaacson, 2010 for further details)  
The simplest approach is to mouth pipette embryos in random orientations onto a poly-L-lysine coated coverslip supported by grease feet above a microscope slide.

- i. Spread a small volume of a 1 mg/mL stock of poly-L-lysine onto coverslips. Allow the coverslips to air-dry for >1 h. Poly-L-lysine is typically applied from a premixed stock solution in distilled water (Sigma). Frozen stocks can be aliquoted and stored at  $-20^{\circ}\text{C}$  indefinitely. Avoid refreezing.
- ii. Cut gravid hermaphrodites at the vulva in M9. Mouth pipette embryos in a small volume (approx. 3  $\mu\text{L}$ ) of water onto precoated coverslips.
- iii. Pipette a ring of silicon oil around the drop, and four dots of silicon vacuum grease (Dow Corning) to the corners of the coverslip. It is typically convenient to insert vacuum grease into a 10 mL plastic syringe without needle, from which it can then be extruded. The grease “feet” provide a spacer that allows a microscope slide to be affixed to the coverslip.
- iv. Inverted a slide over the coverslip to form the mount. Press gently to allow fluid to contact the slide.

b. **Liquid mount with bead spacer** (see Murray *et al.*, 2006 for further details)  
As an alternative to the simple poly-L-lysine mount method, polymer beads are added to the medium to serve as spacers between the coverslip and the slide to prevent the embryo from being excessively compressed. If the bead diameter is  $<25\ \mu\text{m}$  (e.g., 20  $\mu\text{m}$ ), then this will result in slight compression of the embryos and results similar to the standard Agar mount. If larger diameter beads are used (e.g., 30  $\mu\text{m}$ ), then poly-L-lysine should be used as above.

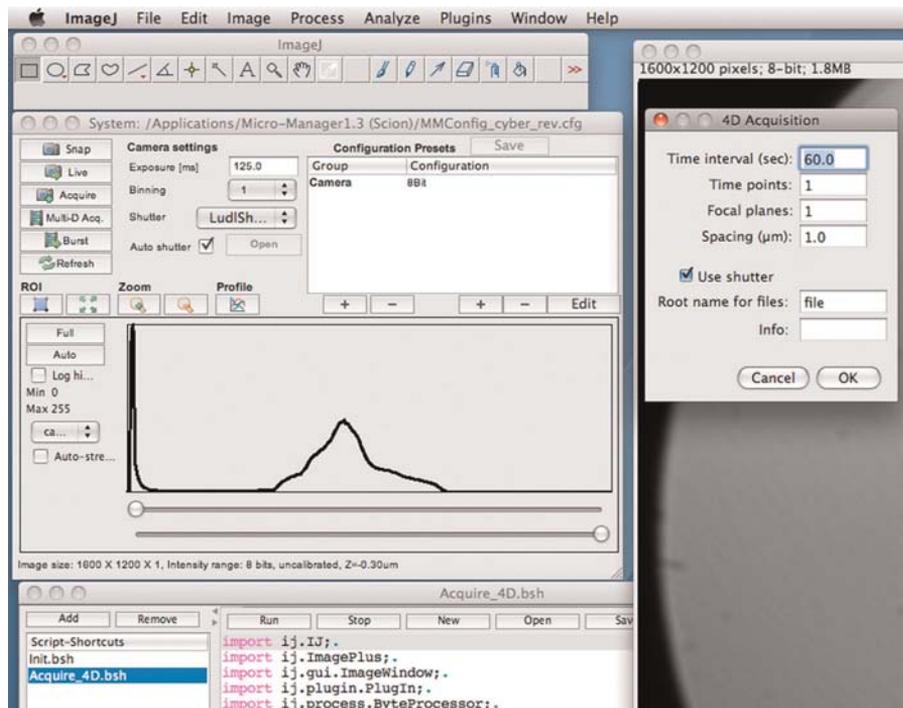
- i. Prepare a 1:30 dilution of 20  $\mu\text{m}$  polymer beads (Polysciences, Inc., Warrington, PA; catalog #18329-5) mixed with M9 in a microfuge tube.

- ii. The bead mixture should then be resuspended with a pipette tip prior to use and added to the coverslip as above. Embryos can then be pipetted into the drop of bead slurry.
- iii. No grease feet are needed using this method. Instead, seal the edges as described above for the standard agar mount, using Valap.

## B. 4D Nomarski Imaging of Morphogenesis

### 1. Introduction

By its very nature, developmental biology requires thinking in four dimensions. Not only do embryos change dramatically over time, as the seemingly featureless single-celled zygote is transformed into an embryo with recognizable body axes and organ systems, but this remarkable transformation occurs in three spatial dimensions. The coordinated changes that occur within the developing embryo include carefully orchestrated signaling events, changes in gene expression, and morphogenetic movements, that is, regulated cell divisions and cell movements that sculpt the basic body plan as a recognizable organism emerges.



**Fig. 2** 4D acquisition script running within MicroManager. A screen shot of the Micro-Manager Beanshell script for 4D acquisition running under Mac OS X 10.5. The script is available at <http://worms.zoology.wisc.edu/4d/4d.html>. (For color version of this figure, the reader is referred to the web version of this book.)

## 2. Acquisition of 4D Nomarski Images During Morphogenesis

### Introduction

Acquisition of stacks of images throughout the thickness of the embryo over time is a crucial method for identifying the positions and contacts between cells. Such “four-dimensional” (4D) microscopy is a routine tool in laboratories that study early *C. elegans* development.

### Imaging Setup

- i. *Microscopy/camera hardware*: This protocol assumes a basic high numerical aperture (NA) microscope equipped with oil immersion objectives and, optionally, an oilable, high NA condenser from any of the major microscope manufacturers. We typically acquire 4D movies using a 60–63 $\times$ , 1.4–1.45 N.A. PlanApo objective. Older Newvicon video cameras, coupled to the video port on the microscope, are adequate for many applications, especially if they are equipped with a zoomable video lens attachment, such as those sold by Nikon Corporation. In this case, an AG-5 digitizing board (Scion Corporation) or similar video frame grabber can be used to digitize the video signal. Alternatively, modern cameras are almost exclusively CCD cameras, and have much higher spatial resolution than older video formats. We have successfully used cameras from Scion Corporation and QImaging Corporation. The mounting hardware for such devices differs depending on the microscope being used.
- ii. *Cooled environment*: We have found that *C. elegans* embryos can be imaged for long periods of time if the ambient temperature is reduced to approximately 20 °C.
- iii. *Z-axis controller/shutter/serial port*: A variety of Z-axis controllers are available from commercial sources (e.g., Prior, Ludl, ASI). In addition, a shutter to block the transmitted light path between time points is strongly encouraged to minimize exposure of embryos to light and heat. A number of shutters are available. We have used shutters from Ludl and Vincent Associates.
- iv. *Software*: Many commercial software packages can be used to acquire 4D footage. If an inexpensive alternative is desired, I have written free software as an alternative. Several options are available:
  - (a) Legacy acquisition plugins for ImageJ: These plugins are available free of charge at the following URL: <http://worms.zoology.wisc.edu/research/4d/4d.html>. Full documentation of the plugins and detailed instructions for installation of ImageJ and Quicktime for Java are available at the same URL. Image J can be obtained at <http://rsbweb.nih.gov/ij/>
  - (b) Micro-Manager: The public domain program Micro-Manager, which is built on top of ImageJ, supports a variety of CCD cameras, Z motors, and shutters, and, with some effort, novices can extend its functionality using the Beanshell scripting language or through the Java plugin architecture supported by ImageJ. Micro-Manager has a major advantage over other free

alternatives, in that it is being continuously updated to support new hardware. The Micro-Manager program can be obtained at the following URL: <http://www.micro-manager.org/>. The remainder of this section will describe the basic use a Beanshell script written within the Micro-Manager environment as an example of how to acquire 4D data.

## Methods

This protocol describes the use of a custom script within Micro-Manager's Beanshell scripting language. The script is helpful to reduce the number of shutter open/close events during 4D acquisition. Alternatively, we have used the standard Micro-Manager package successfully to acquire 4D footage, but this cannot be customized, and does not permit custom file naming conventions and directory structure. The script is available at <http://worms.zoology.wisc.edu/4d/4d.html>

### i. *Setting up a 4D acquisition sequence*

- a. Turn on the Z-axis and shutter control boxes and the CCD camera. Turn on the light switch on the microscope. Find a group of embryos using the 10× objective, prior to oiling the coverslip.
- b. If a high NA condenser is present, place a drop of oil on the condenser (for upright microscopes) or the bottom of the slide (inverted microscopes). Carefully position the condenser so that it contacts the oil and spreads it uniformly between the condenser and microscope slide.
- c. Focus the condenser. The simplest method for achieving good condenser focus is to stop down the condenser using the iris diaphragm, closing it almost completely. Then the height of the condenser can be adjusted at high magnification until the octagonal outline of the diaphragm is in focus. When done, open the condenser.
- d. Once embryos have been located at 10× and the condenser has been focused, swing the 10× objective out of the way and add a drop of immersion oil to the coverslip (upright microscope) or the 60× objective lens (inverted microscope). We find that Type DF oil works well.
- e. Carefully slide the 60 or 100× objective into place (it should just clear the sealant on the slide, as long as it is not too thick). Make sure the correct condenser setting is selected to match the lens.
- f. Refocus on the embryos, and refocus the condenser.
- g. Open Micro-Manager. Use the “Live” button in the main Micro-Manager Studio window to display an image from the camera. If the “Autoshutter” option is not checked, click the “Open” button to open the shutter. Otherwise it should open when the “Live” button is clicked. Optimize the positioning of embryos in the field using the stage controls on the microscope, and/or by rotating the CCD camera gently by hand (if the mount supports this). Optimize the Nomarski optics through a combination of the following:

- Center the condenser by closing it and moving the octagon to the center of the field of view. Reopen the diaphragm to encompass the entire field of view.
  - Adjust the light level. High-quality Nomarski optics requires a substantial amount of light. Optimal settings must be empirically determined.
  - Adjust the exposure time, gain, and other settings on the CCD camera within Micro-Manager a final time if needed.
- h. Invoke the Micro-Manager 4D acquisition script (Fig. 2). This protocol presupposes that a “favorite” has been created previously using the script window in Micro-Manager. This window is invoked using the “Tools -> Script Panel” menu command in Micro-Manager. When the script window appears, select “Acquire\_4D.bsh” from the list of favorites. Make sure that the cursor is blinking within the code of this script. Then click “Run.” Enter the desired parameters for time interval, number of time points, number of focal planes, and distance between focal planes. Enter the root name for the images that will be collected. [Note: because most operating systems limit the total length of a file’s name to 32 characters, the root name should be kept short]. If a shutter is being used, makes sure that the “Use shutter” option is selected.
- i. Click “OK.” The parameters that have been entered will be displayed. If these are acceptable, click “OK.” When prompted for a location to which to save images, make a new directory that will contain the images from the 4D sequence. Within the newly created directory, we recommend making two additional directories: (a) one called “working” and (b) one called “terminal.” The latter is useful for acquiring a final Z stack of the terminal embryos. Typical settings for a long overnight movie are the following:
- Number of time points: usually 200–300 for an overnight movie  
Time interval (s): usually 120–180  
Number of shutters: 1  
Number of focal planes: 20  
Distance between focal planes: 1  $\mu\text{m}$   
Root name: “working,” or a short name of choice  
Information for movie: Enter any pertinent information.
- j. Once a directory is specified, the computer should start acquiring images. Status updates will be displayed in the ImageJ main window. To abort, click the “Stop” button in the “Script Panel” window.
- k. When the movie is finished, we recommend collecting a terminal image stack. To do so, keep the field of view the same. Collect a second movie, specifying “1” as the number of time points. Save this movie in the “terminal” folder created previously.
- l. To view the movie, there are several options available:
- (i) Raw 4D datasets: These can be viewed in one of several ways, including
    - (a) importing the sequences as a “Virtual 5D Stack” or Hyperstack within

ImageJ, using the “Virtual 5D Stack” plugin or ImageJ itself, available on the ImageJ web site, or (b) using the “Browse4D” plugin available at <http://worms.zoology.wisc.edu/4d/4d.html>.

- (ii) Compressed movies: Movies can also be compressed to save disk space and viewed using QuickTime and the “QT4D Writer” and “QT4D Player” plugins available at the same URL. The advantage of postacquisition compression is that the storage requirements are greatly reduced, and, if QuickTime is used, the entire dataset does not need to be imported into RAM or imported as a virtual stack. Unfortunately, Quicktime for Java has been deprecated by Apple Computer, Inc. If QuickTime compression is being used, we typically save the movie using the same root name as the raw files, with the word “movie” appended in the same directory created for the experiment. Although many compression algorithms are available, we typically use “Photo/JPEG,” compression, “gray scale,” and “Medium” quality. This approach can compress movies 30 fold. To play compressed movies, use the “QT4D Player” plugin. Select the desired movie. A graphical interface with clickable buttons or the arrow keys on the numeric keyboard can be used to navigate through movies. The up and down arrow keys can be used to scroll up and down through the focal planes; the left and right arrow keys scroll backwards and forwards in time. Once the movie has been successfully compressed and its quality verified, for routine purposes it is now fine to delete the original files.

### Troubleshooting

- i. Problem: No light appears to be reaching the camera.  
Solution: Make sure the slider that diverts light from the microscope to the camera port is in the proper position, and that the power supply to the camera is on. If the shutter has an external toggle switch, make sure that it is in the correct position. If the exposure time is set to too low a value, increase the exposure time using the controls in the Main Micro-Manager window.
- ii. Problem: The plane of focus drifts systematically over time.  
Solution: This often occurs in the first few minutes after making an agar mount. For this reason, it is advisable to check the focus several times during the first 15–20 min of acquisition. To reset the focus, open the shutter and use the coarse focus on the microscope to refocus on the top focal plane.
- iii. Problem: Temperature variation in the room results in inconsistent time course of development or variable phenotypes.  
Solution: For best results, filming should take place in a room held at constant temperature, approximately 20 °C. Make sure the air conditioner is on and that the door remains closed.

- iv. Problem: After several hundred time points Micro-Manager reports an error from which it cannot recover.

Solution: Some users have reported errors under Micro-Manager when using USB-to-serial port adapters. This is known issue with version of Micro-Manager prior to 1.4. Using a PCI-based serial port card appears to alleviate this problem. Alternatively, acquire several shorter movies. We have successfully used this script under Micro-Manager for acquiring up to 150 time points with 25 focal planes/time point. We have had success acquiring long movies with Micro-Manager 1.46, Mac OS X 10.7, and a KeySpan USB-to-serial adapter.

## Discussion

This procedure will result in the production of 4D datasets in the form of a series of consecutively named TIFF files that can be read by many different programs, including ImageJ, especially when supplemented with appropriate plugins. The reduced costs of such a system make this basic system feasible for teaching laboratories and research laboratories within limited funds. While we have described the use of such a setup for imaging *C. elegans* embryos, this apparatus is well suited to acquiring images of any transparent specimen.

## 3. Introducing Pharmacological Agents During 4D Acquisition

(Adapted from Williams–Masson *et al.*, 1997)

### Introduction

This procedure assumes a fairly standard laser ablation setup. Many *C. elegans* laboratories use a nitrogen laser to pump a tunable dye laser routed through the epifluorescence light path of the microscope. The dye cuvette typically contains Coumarin 440 dye, which can be obtained from Sigma or other suppliers, and is reconstituted in methanol at a concentration of 5 mM in methanol (see the protocol in Walston and Hardin, 2010 for further details). A common version of this setup is sold by Photonic Instruments, a division of Andor Corp., under the trade name Micropoint.

### Methods

#### 1. Creating the mount

- i. Cut gravid hermaphrodites at the vulva in water. Mouth pipette embryos in a small volume (approx. 3  $\mu$ L) of water onto coverslips precoated with 1 mg/mL poly-L-lysine. Poly-L-lysine is typically applied from a premixed stock solution in distilled water (Sigma).
- ii. Allow embryos to settle for 30 s and then treat for 2 min with 100  $\mu$ g of FITC-conjugated poly-L-lysine (Sigma). FITC-poly-L-lysine enhances the

ability of a standard ablation laser tuned to 440 nm using Coumarin dye to perforate the eggshell.

- iii. Rinse embryos three times with embryonic growth medium (EGM; for a detailed recipe, see Shelton and Bowerman, 1996) + 3  $\mu\text{g}/\text{mL}$  Nile Blue A (Sigma) + 1–2  $\mu\text{g}/\text{mL}$  cytochalasin D or nocodazole. The addition of Nile blue allows the assessment of perforation of the eggshell; permeabilized embryos will take up the dye into granules in gut cells.
- iv. Cover with a 30  $\mu\text{L}$  drop of EGM plus drug. Stock solutions of 2 mg/mL cytochalasin D or nocodazole (Sigma) in DMSO can be stored at 4 °C.
- v. Pipette a ring of silicon oil around the drop, and four dots of silicon vacuum grease (Dow Corning) to the corners of the coverslip. It is typically convenient to insert vacuum grease into a 10 mL plastic syringe without needle, from which it can then be extruded. The grease “feet” provide a removable spacer that allows a microscope slide to be affixed to the coverslip.
- vi. Invert a slide over the coverslip to form the mount. Press gently to allow fluid to contact the slide.

## 2. Perforating the eggshell

- i. Determine the position of the ablation beam by moving to a region of the mount away from the embryos. Focus on the coverslip, and crack the coverslip using the laser. In the case of the standard manual Micropoint laser, a sliding neutral density grating can be used to attenuate the beam strength. Beam amplitude should be just sufficient to crack the coverslip. If more power is desired later, then the grating can be adjusted accordingly.
- ii. Select an embryo of the desired developmental stage. Find a region of the embryo where there is a space between the cells and the eggshell. Embryos are typically oriented within the eggshell such that there is a larger space between the anterior end of the embryo and the eggshell than in other regions of the embryo, making this region convenient for laser irradiation.
- iii. Position the embryo so that the eggshell of this region is over where the ablation laser will hit.
- iv. Using a foot pedal or push button, pulse the laser. Usually only one hit is necessary, but sometimes more pulses are required.
- v. It is often possible to tell that the eggshell is perforated because the embryo will move slightly toward the perforation site. However, we have found that sufficiently small holes will not induce this response, yet dye penetration will nevertheless be observed.
- vi. If needed, additional perforations can be induced to increase the rate of penetration of the compound of interest. However, we have found that the number of viable embryos obtained in these cases goes down markedly.

- vii. These mounts can be imaged using 4D microscopy. We have also found it possible to remove the coverslip and process embryos for phalloidin staining (see above; for examples, see references).
- viii. Embryos can be scored after 4 h for blue gut granules, which indicate that sufficient permeabilization was achieved for Nile Blue A penetration.

## Discussion

Performing long-term 4D filming after perforation of the eggshell is difficult. Many embryos show abnormalities in subsequent development. Extensive negative controls (i.e., perforation of the eggshell in the presence of carrier, such as DMSO, alone) are therefore highly advisable. If sufficient precautions are taken, however, it is possible to perform pharmacological inhibition followed by 4D filming, as we have shown in several circumstances (Thomas-Virnig *et al.*, 2004; Williams-Masson *et al.*, 1997).

## C. Fluorescence Imaging of Morphogenesis

### 1. Introduction: Imaging Modalities for 4D Fluorescence Imaging

Nomarski microscopy, while a daily workhorse for imaging morphogenesis and performing basic phenotyping, is limited. Refractile elements in the cytoplasm of embryonic cells, combined with the inherent curvature of the embryo, limits the resolution of the standard Nomarski microscope. In addition, the epidermis is exceedingly thin (less than 0.5  $\mu\text{m}$  in some cases), making it difficult to resolve. Fluorescence imaging of specific structures in embryos, combined with confocal or multiphoton microscopy, overcomes these challenges. The chapter by Maddox and Maddox in this volume covers basic modalities of fluorescence microscopy. Here we discuss several useful strategies for visualizing cells and subcellular structures during morphogenesis, an alternative phalloidin staining procedure, and a simple strategy for analyzing fluorescence recovery after photobleaching (FRAP) data.

Given the wide array of genetically encoded fluorescent probes available (see below), 4D datasets of fluorescent specimens acquired using confocal, multiphoton, or widefield deconvolution techniques have several advantages over 4D datasets acquired using transmitted light optics, such as Nomarski microscopy. First, such techniques permit much more refined optical sectioning of the specimen with little contribution by out-of-focus information. Secondly, it is much easier to understand the distribution of the fluorescent signal from a 3D reconstruction of a sample than a 3D-stack of DIC images (for an attempt at the latter, see Heid *et al.*, 2002). For thin specimens imaged within 5  $\mu\text{m}$  of the coverslip such as *C. elegans* embryos, oil immersion optics and a high NA lens (NA = 1.4–1.45) are typically the best choice.

Although the use of fluorescent probes present several key advantages, it also presents several challenges for 4D imaging of morphogenesis. Because *C. elegans*

embryos are not flat, multiple focal planes must be acquired at each successive time point, requiring repeated exposure of the region of interest (ROI) to excitation photons that compromise embryo viability. In addition, the photobleaching that accompanies repeated observation in 4D experiments complicates image quantification. If photobleaching is not too severe, it is sometimes possible to normalize the signal from bleached specimens using several algorithms (e.g., see the ImageJ plugin that is part of the McMaster Biophotonics suite of routines: <http://www.macbiophotonics.ca/imagej/>).

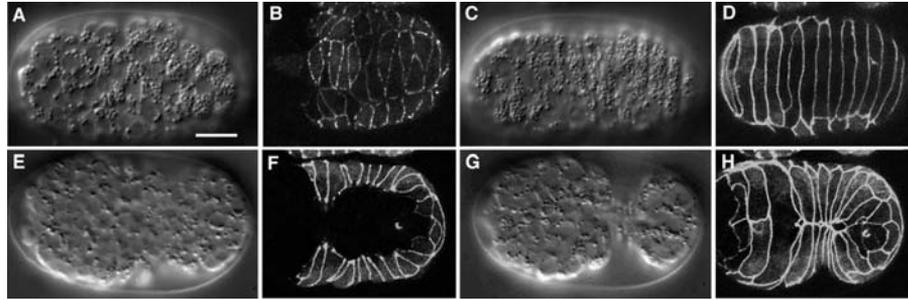
Several imaging modalities have been used frequently for analyzing morphogenesis. For the physical basis of each, readers should consult the chapter by Maddox and Maddox in this volume. Here we will focus on issues related to filming morphogenesis.

### Laser scanning confocal microscopy (LSCM)

In some cases LSCM suffices. Depending on the density of fluorophore labeling and the nature of the tagged moiety, many experiments have been performed using rather unremarkable equipment, such as a Bio-Rad 1024 or similar older laser scanning confocal devices. Interested readers are urged to consult older reviews describing methods of confocal imaging in *C. elegans* (e.g., Mohler, 1999). LSCM is typically the mode of choice for performing single focal plane photobleaching experiments (see below). For long-term 4D acquisition in which many focal planes are acquired, LSCM quickly leads to arrest of embryos. The time required to induce arrest varies, depending on the ambient temperature and the probe being imaged. However, in general LSCM is not well suited to long-term viability of embryos filmed over many hours.

### Multiphoton excitation laser scanning microscopy (MPLSM)

Multiphoton laser-scanning microscopy has several potential advantages over LSCM for live imaging of embryos. MPLSM excites fluorescence using a series of short, high-energy pulses of near-infrared photons from a mode-locked laser. MPLSM has a key advantage for live embryo imaging experiments in *C. elegans*. In a two-photon microscope the probability of excitation varies as the inverse fourth power of the distance from the focal plane. Photons are thus only absorbed in a very small volume centered on the plane of focus, eliminating photobleaching and photo-damage caused by excitation of fluorophores above and below the plane of focus. The resulting improvements in viability can be quite dramatic. In our laboratory, *C. elegans* embryos expressing a GFP-tagged junctional protein survive for 30–90 min when imaged using a Bio-Rad 1024 CLSM at low power (10%; J. Hardin, unpublished) but the same embryos can be imaged for many hours using MPLSM (Raich *et al.*, 1999). An example is shown in Fig. 3 (see Köppen *et al.*, 2001 for more details).



**Fig. 3** Comparison of Nomarski 4D imaging (A, C, E, G) with the use of DLG-1::GFP as a junctional marker (B, D, F, H). (A, B) Early dorsal intercalation of epidermal cells. (C, D) Intercalation complete. (E, F) Early ventral enclosure. (G, H) Ventral enclosure complete. Bar = 5  $\mu\text{m}$ . (Adapted from Chisholm and Hardin, 2005).

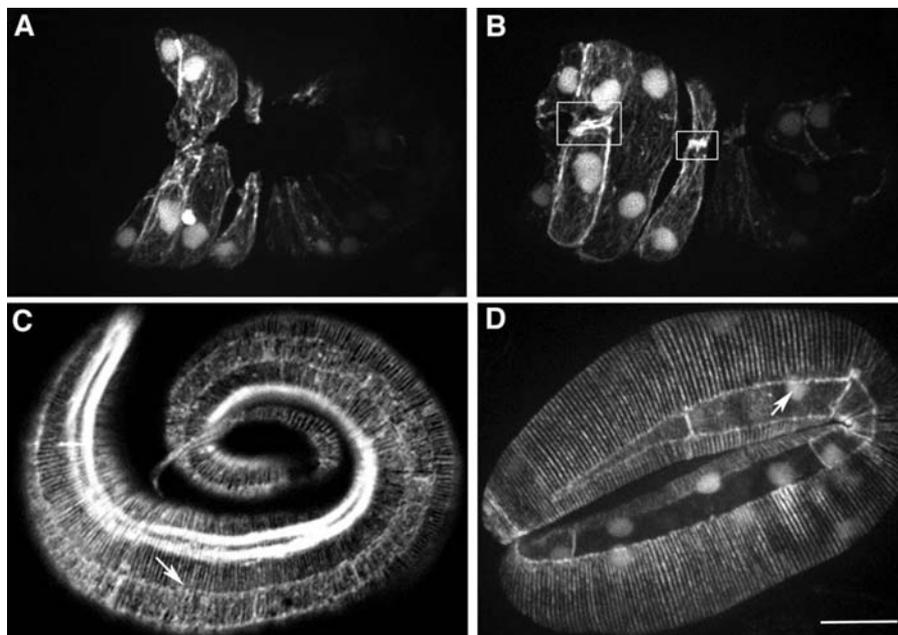
Although MPLSM can be superior to CLSM for many applications, there are several issues that can make MPLSM less than optimal. First, the typical MPLSM device is expensive, placing it out of reach of most individual laboratories. In contrast, individual labs can often afford disk-scanning confocal microscopes, a factor that is particularly important for live embryo studies, which often monopolize microscope time. Second, for fluorophores that emit in the red portion of the visible spectrum, the wavelengths needed to generate a two-photon event are longer than those produced by the Ti:Sapphire lasers commonly used in commercial MPLSM devices. For such probes, Nd:YLF lasers, which emit at 1047 nm, are very effective (Mohler and White, 1998), but may not be readily available.

### Disk-scanning confocal microscopy

For the developmental biologist, spinning disk systems based on Yokagawa scan-heads are an inexpensive alternative that provides many of the benefits of more elaborate technologies, such as MPLSM. Because disk-scanning systems use an off-the-shelf focus motor, CCD, filter wheel, and shutter components, commercial imaging packages or freeware packages such as Micro-Manager can be used to drive data acquisition. In our laboratory, disk-scanning technology has largely replaced both CLSM and multiphoton microscopy for routine 4D data acquisition during morphogenesis (for examples, see Figs. 4A,B, 5, and 6B,C). The data generated by spinning disk can also be improved via postacquisition deconvolution, making its Z resolution comparable to LSCM or MPLSM.

### Considerations for optimizing 4D data during morphogenesis

There are several thorough treatments of empirical considerations that lead to high-quality 4D fluorescence dataset (e.g., Hardin, 2006). Here, we will focus on



**Fig. 4** Imaging of F-actin during morphogenesis. (A, B) F-actin dynamics during ventral enclosure imaged in a wild-type embryo expressing *vab-10::ABD::GFP*. Images were acquired using a Perkin-Elmer UltraView LCI system, with Hamamatsu Orca II-ER camera. The boxed regions in (B) show extensive actin accumulation between two pairs of ventral epidermal cells. Elapsed time between (A) and (B) is 1380s. (From Lockwood *et al.*, 2008). (C) Phalloidin staining of an elongated embryo. Circumferential filament bundles (CFBs) are clearly visible throughout the epidermis. The arrow points to junctional actin. (From Costa *et al.*, 1998). (D) A twofold stage embryo expressing *VAB-10ABD::GFP*. CFBs are prominent; arrow points to junctional actin. Image courtesy of R. Zaidel-Bar. Bars = 5  $\mu$ m.

several key factors affecting 4D fluorescence acquisition during morphogenesis. First, as with 4D Nomarski imaging, ambient temperature must be controlled carefully. The temperature must be kept below 25 °C; for long films, a temperature closer to 20 °C is advisable. This is often not possible in shared user facilities, in which elevated temperatures suited to tissue culture work are the focus. Second, despite theoretical calculations of voxel sampling in Z stacks of fluorescent images, it is typically advisable to acquire very closely spaced optical sections if one is imaging events in the epidermis. We have found that focal planes spaced 0.5  $\mu$ m apart or less are necessary, due to the extreme thinness of the epidermis. Finally, modern lenses with newer coatings make a significant difference. For very high-resolution filming of cytoskeletal elements or thin structures, we have found that very high NA lenses are helpful. In particular, we have found that lenses designed for total internal reflection microscopy (TIRF), but without the internal optics for TIRF itself, provide

excellent results, such as recent vintage Nikon NA 1.45 oil immersion “TIRF” lenses.

Ultimately, the choice of a particular imaging modality is largely based on empirical issues, including the type of probe being imaged. We now turn to types of probes useful for analyzing morphogenesis.

## 2. Probes for Visualizing Morphogenesis

### Introduction

The discovery and widespread use of variants of the green fluorescent protein (GFP) as a tag for visualizing gene expression and protein localization within living organisms has revolutionized live embryo imaging, including in *C. elegans*. Many of the probes we use routinely for analyzing morphogenetic events are genetically encoded. However, several other techniques have proven useful for studying morphogenesis in *C. elegans*.

### Junctions

For studying morphogenetic movements in embryos, junction-localized FPs are extremely useful. Our laboratory used junction-localized FPs to study epithelial sheet movement in *C. elegans* early on, and these tools have become standard among *C. elegans* researchers. These include AJM-1 (Mohler *et al.*, 1998), DLG-1 (Köppen *et al.*, 2001), and HMP-1/ $\alpha$ -catenin (Raich *et al.*, 1999). Others include JAC-1/p120ctn (Pettitt *et al.*, 2003) and HMR-1/cadherin (Achilleos *et al.*, 2010). The use of epithelial junctional markers is particularly useful for following cellular movements at single-cell resolution during events such as dorsal intercalation, ventral enclosure, and the early steps of elongation. The use of such markers provides much clearer views of morphogenetic movements than can often be achieved with Nomarski microscopy (Fig. 3). We have also found that it is possible to create transgenic lines expressing more than one FP-tagged junctional marker, such as *dlg-1::dsRed* and *hmp-1::gfp* (Zaidel-Bar *et al.*, 2010). In some cases, however, especially if two junctional proteins physically interact, coexpression can lead to artifactual aggregation of proteins (C. Lockwood and J. Hardin, unpublished).

### Actin and tubulin

#### i. Actin

We have employed two basic methods to visualize the F-actin cytoskeleton in embryos during morphogenesis: (i) phalloidin staining of fixed specimens and (ii) imaging of the F-actin binding fragment of the spectraplaklin, VAB-10, fused

to GFP. Since phalloidin staining of F-actin containing structures is such an important technique for studying epidermal morphogenesis, we provide this alternative procedure to that described by Maddox and Maddox in Chapter XX. An example of an embryo stained with phalloidin to reveal actin-rich structures is shown in Fig. 4C.

(a) *Phalloidin staining for analyzing morphogenesis*

**Method**

1. Make solutions and coat ring slides with 25  $\mu\text{L}$  poly-L-lysine.
2. Obtain eggs from adults by bleaching. Make sure to sweep embryos off the plate using a glass pipette.
  - a. Hint: It is sometimes advantageous to bleach embryos for a slightly longer period of time to further weaken the eggshell. This can be done by allowing the embryos to remain in the bleach solution for another minute after the worm carcasses have all been dissolved.
3. Wash embryos at least  $2\times$  with ddH<sub>2</sub>O. We find they stick to the slide better when using water, not M9.
4. Add embryos to ring slides that have been precoated with poly-L-lysine. Let sit 5–10 min to give embryos time to adhere.
  - a. Hint: It helps the embryos to adhere when the slide is rinsed with ddH<sub>2</sub>O before adding the embryos.
5. Remove the water from the ring slide. The bulk of the liquid can be poured off into a liquid waste container, if care is taken to avoid mixing embryos from one ring to another. Use a Kimwipe to gently remove the remaining liquid.
6. Add 45–60  $\mu\text{L}$  of the fix/permeabilization solution to each ring.
  - a. Incubate the slides for 20 min in a humid chamber at room temperature. If there are many embryos, it is advisable to incubate 2–5 min longer.
  - b. A simple humid chamber can be made by placing wet paper towels in the bottom of a Tupperware container.
7. Wash slides  $2\times 5$  min with  $1\times$  PBS by adding 70  $\mu\text{L}$  of PBS to each ring per wash. Remove liquid after each wash as above.
8. Add 60  $\mu\text{L}$  phalloidin solution to each ring.
  - a. Incubate 1–2 h in a humid chamber at room temperature in the dark (we usually incubate for 90 min). Alternatively, incubations can be carried out at 4 °C overnight in the dark, which can reduce background.
9. Wash slides  $2\times 5$  min with  $1\times$  PBS while rotating.
10. Dry off slide using a Kimwipe. Mount by inverting an 18 mm square coverslip with 8  $\mu\text{L}$  SlowFade over each sample. Seal edges with nail polish.

**Reagents****Bleach Solutions**

1. Bleach solution:
  - a. 0.4 mL bleach
  - b. 0.4 mL 10N KOH
  - c. 3.2 mL water
2. Fix/permeabilization solution:
  - a. 200  $\mu$ L 10% PFA
  - b. 10  $\mu$ L 10% Triton X-100 (Sigma) or 5  $\mu$ L 10 mg/mL lysolecithin in chloroform (Sigma). Lysolecithin is much gentler than Triton, and will preserve fine structure, but also results in less overall extraction, leading to higher background signal in some cases. For preserving fine filaments in the epidermis, lysolecithin is preferable.
  - c. 48  $\mu$ L 0.5 M PIPES (pH = 6.8)
  - d. 25  $\mu$ L 0.5 M HEPES (pH = 6.8)
  - e. 1  $\mu$ L 1 M MgCl<sub>2</sub>
  - f. 10  $\mu$ L 0.5M EGTA
  - g. 196  $\mu$ L ddH<sub>2</sub>O
3. Phalloidin solution:
  - a. 6  $\mu$ L 6.6  $\mu$ M Alexa-488
  - b. 114  $\mu$ L 1  $\times$  PBS

**Other solutions:**

1. 10% PFA
  - a. 1 g paraformaldehyde
  - b. 30  $\mu$ L 5N NaOH
  - c. 10 mL 60 mM PIPES (1.2 mL 0.5M PIPES + 8.8 mL ddH<sub>2</sub>O)

Incubate ~30 min in 65 °C water bath until PFA is dissolved (note: do not exceed 65 °C!)

2. 10% Triton
  - a. 1 mL Triton X-100
  - b. 9 mL ddH<sub>2</sub>O

Place on a nutator platform to mix thoroughly

3. 0.5 M PIPES
4. 0.5 M HEPES
5. 1 M MgCl<sub>2</sub>
  - a. Dissolve 203.3 g of MgCl<sub>2</sub>·6 H<sub>2</sub>O in 800 mL of dH<sub>2</sub>O. Adjust the volume to 1 L with dH<sub>2</sub>O. Dispense into aliquots and sterilize by autoclaving or filtering. Keep aliquots closed when they are not being used.

## 6. 0.5M EGTA

- a. Add 190.175 g ethylene glycol bis(beta-aminoethyl ether) *N,N,N',N'*-tetraacetic acid (EGTA) to 1 L of dH<sub>2</sub>O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH. Dispense into aliquots and sterilize by either autoclaving or filtering.

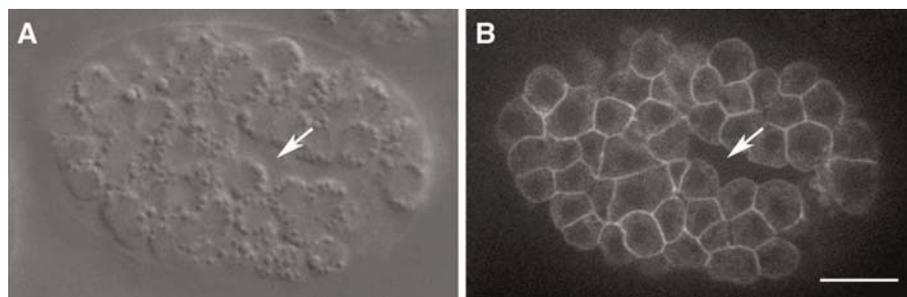
## 7. 1× PBS

- a. 137 mM NaCl
- b. 2.7 mM KCl
- c. 10 mM Na<sub>2</sub>HPO<sub>4</sub>
- d. 2 mM KH<sub>2</sub>PO<sub>4</sub>

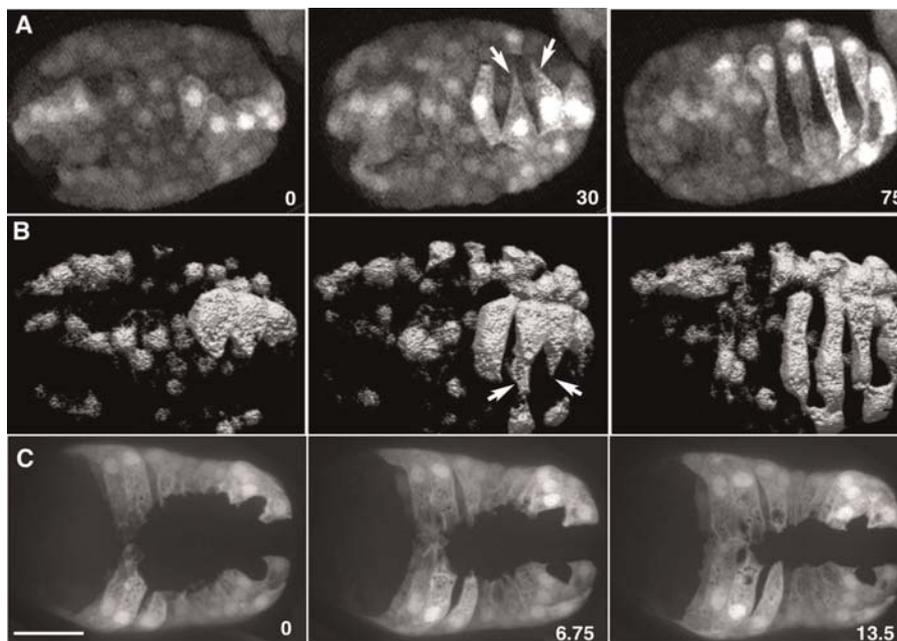
Dissolve 8 g NaCl, 0.2 g KCl, 1.44 g Na<sub>2</sub>HPO<sub>4</sub>, and 0.24 g KH<sub>2</sub>PO<sub>4</sub> in 800 mL dH<sub>2</sub>O. Adjust the pH to 7.4 with HCl. Add dH<sub>2</sub>O to 1 L. Autoclave to sterilize.

- (b) *VAB-10ABD::GFP and other GFPs*: Phalloidin staining is indispensable for studying the F-actin cytoskeleton during morphogenesis. However, it is not suitable for capturing rapid, dynamic movements involving the actin cytoskeleton. GFP-tagged versions of monomeric subunits of actin, such as *act-2* (Willis *et al.*, 2006) can be useful for some purposes, but the GFP likely affects polymerization dynamics. An alternative approach for *in vivo* analysis in living embryo exploits F-actin-binding proteins. Two that have been used during early development are moesin::GFP (Velarde *et al.*, 2007) and Lifeact::GFP (Pohl and Bao, 2010). For the study of morphogenesis, the spectraplakins VAB-10, which contains both actin- and tubulin-binding motifs, was exploited by the Labouesse laboratory to produce a construct containing the actin-binding domain (ABD) of VAB-10 tagged with GFP or mCherry (Gally *et al.*, 2009). We have used this construct to image filopodial dynamics during ventral enclosure (Fig. 4A,B), as well as circumferential filaments network and junctional actin during elongation (Fig. 4C, D). Use of the VAB-10 constructs requires care; we have found that excessive imaging can lead to lethality in strains containing *vab-10ABD* transgenes. For short periods of time, however, these constructs are invaluable for imaging rapid actin dynamics. In addition, the original *vab-10ABD* strains are mosaically expressed. For some applications, this is actually an advantage, but for others it can be somewhat problematic.
- ii Tubulin: Numerous antibodies can be used to visualize tubulin during morphogenesis, using standard freeze-cracking (Miller and Shakes, 1995). Alternatively, there are several GFP constructs that have been successfully used to visualize epidermal cells. A particularly useful set of such constructs have recently been published that use the *lbp-1* promoter to drive expression of either tubulin subunits or plus-end tracking proteins predominantly in epidermal cells. These include *Plbp-1::gfp::β-tubulin* and *Plbp-1::ebp-1/EB 1::gfp* (Fridolfsson and Starr, 2010). These

- should be especially useful for analyzing microtubule dynamics during morphogenesis.
- iii Cell cortex/membranes: Several markers have been particularly useful for visualizing the cell cortex or membrane. These include a *ced-10/Rac::gfp* translational fusion (Liu *et al.*, 2007) and various PH-domain constructs fused to mCherry or GFP (e.g., Audhya *et al.*, 2005). An example of the use of *ced-10::gfp* is shown in Fig. 5. The interested reader is encouraged to consult the primary references for further details. An alternative technology for visualizing membranes involves the use of lipophilic dyes, such as the red dye FM4-64, which has been used to visualize cell fusion in the dorsal epidermis. Since the eggshell is impermeable to such dyes, they must be introduced via laser perforation of the eggshell as described above. In the case of FM4-64, its excitation wavelength requires the use of fairly specialized lasers if multiphoton excitation is used (Mohler *et al.*, 1998).
  - iv Cytosolic markers: To developmental biologists, fusing the coding region of EGFP to the regulatory DNA associated with a gene of interest (i.e., GFP “reporter constructs”) is often used to assess the tissue-specific and temporal patterns of transcriptional activation of a gene. Such information provides valuable information about how the expression of a gene is regulated. However, such transcriptional reporters can also be invaluable for live embryo imaging for several reasons. First, such reporter constructs result in the expression of GFP in the cytosol; because GFP is fairly small, these reporters are capable of percolating into small volumes within the cytoplasm, including fine protrusions extended by cells as they migrate. Second, the highly specific pattern of expression of some genes allows either many or a very small number of cells to be visualized against a dark background,



**Fig. 5** CED-10/Rac::GFP allows imaging of cell cortices during morphogenesis. The same embryo expressing *ced-10::gfp* (Liu *et al.*, 2007) was imaged using fluorescence and Nomarski microscopy. The arrows indicate the closing ventral gastrulation cleft. Images were acquired using a Perkin-Elmer UltraView LCI system, with Hamamatsu Orca II-ER camera. Images courtesy of R. Zaidel-Bar. Bar = 5  $\mu$ m.



**Fig. 6** Cytosolic markers allow imaging of cell motility and cell movement in embryos. (A, B) Frames from 4D movies of dorsal intercalation in embryos expressing *lbp-1p::gfp*, which is expressed in a subset of dorsal epidermal cells. (A) An embryo imaged using two-photon excitation microscopy. *z*-stacks were subsequently projected using a maximum intensity procedure. (From Heid *et al.*, 2001). (B) A similar embryo imaged using a Perkin-Elmer UltraView LCI system (with Yokagawa CSU10 scanhead; images courtesy of T. Walston). The dataset was subsequently subjected to surface rendering using Volocity software. Fine protrusions are visible in both cases. It is clear in (B) that the protrusions are wedge-shaped in the *z*-dimension. (C) Frames from a 4D movie of ventral enclosure in an embryo expressing a *Pdlg-1::gfp* reporter. Elapsed time in minutes is shown. *z*-stacks were acquired at 50 s intervals, 20 focal planes/stack; acquisition time/image, 300 ms. Fine details of protrusions are visible against a dark background using this particular transcriptional reporter. (Images courtesy of M. Sheffield.) Bar = 10  $\mu$ m.

dramatically improving the effective contrast of the specimen being imaged. Third, imaging cytosolic GFP reporters typically does not cause as much photodamage as with GFP translational fusions. We have frequently used two different markers to image events in the epidermis: *Plbp-1::gfp* (Heid *et al.*, 2001) and *Pdlg-1::gfp* (Sheffield *et al.*, 2007). An example of the former is shown in Fig. 6A–B, and the latter in Fig. 6C. An extremely useful adjunct to the use of cytosolic markers is the use of a voxel rendering program, such as Volocity (Perkin-Elmer). Fig. 6B shows the results this procedure in the case of intercalation of dorsal epidermal cells. The result is a striking 3D view of cells as they intercalate.

### 3. FRAP

#### Introduction

In addition to imaging specific structures within embryos during morphogenesis, fluorescence can also be used to study the dynamics of redistribution of molecules in embryos using FRAP. GFP can be used very successfully in studies involving FRAP (Lippincott-Schwartz *et al.*, 2003), which is easily done via repeated scanning of a selected ROI in the LSCM. FRAP has been used in several contexts in *C. elegans*, particularly in one-celled zygotes (e.g., Labbe *et al.*, 2004).

#### Methods

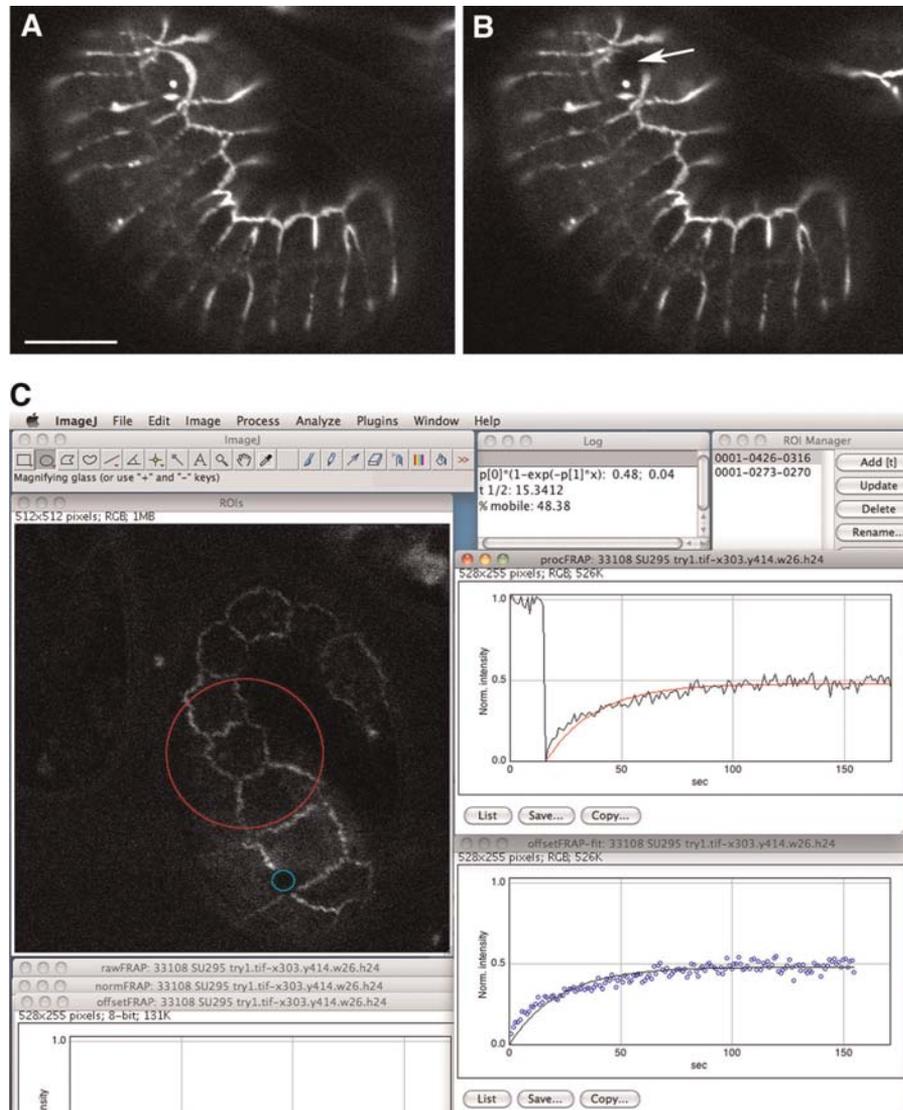
##### 1. *Acquiring FRAP data*

The acquisition of the raw data to be analyzed in a FRAP is highly dependent on the apparatus and acquisition software, so it will not be discussed here in detail. We have been able to successfully photobleach single junctional domains in embryos undergoing morphogenesis. In our case, these experiments have been performed using an Olympus Fluoview 1000 with SIM scanner for rapid photobleaching with simultaneous imaging (Fig. 7A,B). In general, the conditions that allow for successful imaging via laser scanning confocal microscopy pertain. We have found that single focal planes can be imaged in embryos in agar mounts quite successfully. We have also found several specific considerations to be important in performing FRAP during morphogenesis:

- a. During morphogenesis, embryos may move/twitch slightly. The bleach zone ROI must be chosen such that the bleach zone does not move out of the ROI. For FRAP of junctional GFPs, this can be accomplished by altering the size and/or aspect ratio of the ROI. Note that as a result, FRAP is practically limited to early phases of elongation; after this time, embryos twitch too much to make reliable measurements.
- b. Embryos in which FRAP is performed at temperatures exceeding 27 °C may show erratic results. This is likely due to overall sickness of embryos at elevated temperatures. If possible, filming should occur at  $\leq 25$  °C.

##### 2. *Analyzing FRAP data: Simplified FRAP analysis using ImageJ*

Once obtained, analysis of the recovery kinetics of a bleached region in a FRAP experiment proceeds along similar lines. There are many options for analyzing data from FRAP experiments. The technique we describe here is a free, semi-automated solution suitable for use by undergraduates, and relies on plugins we have adapted from original Java routines written by Tony Collins, McMaster University (Collins, 2007). This plugin is available for download from <http://worms.zoology.wisc.edu/research/4d/4d.html>. To use, install the plugin by copying to a convenient directory within the Plugins directory in ImageJ. For more information, see the ImageJ web site (<http://rsbweb.nih.gov/ij/>).



**Fig. 7** Fluorescence recovery after photobleaching (FRAP) analysis of junctional protein mobility. (A) An embryo expressing *hmr-1::gfp* immediately prior to bleaching (image courtesy of J. Keegan). (B) The same embryo after bleaching. The arrow indicates the region that was bleached. Bar = 5  $\mu\text{m}$ . (C) Screen capture of the output from the ImageJ-based FRAP profiler plugin. The example shown is analysis of mobility of *jac-1::gfp* (raw images courtesy of A. Lynch). (For color version of this figure, the reader is referred to the web version of this book.)

### 3. *Importing data from the confocal microscope*

- a. *Invoke the LOCI Bio-Formats Importer plugin within ImageJ:* Many standard installations of ImageJ, or the package installer based variant of ImageJ, Fiji (<http://pacific.mpi-cbg.de/wiki/index.php/Fiji>) may have Bio-Formats Importer installed, or it can be obtained from <http://www.loci.wisc.edu/software/bio-formats>. The Bio-Formats plugin can read many data formats, including most raw image formats generated by standard confocal microscopes. We have tested this code using datasets from an Olympus Fluoview 1000 (.oif format files), but the basic approach should work for other types of datasets.
- b. *Import the dataset as an image stack into ImageJ:* After opening the Bio-Formats plugin, select a file within the directory containing the dataset, and import the data. Our setup uses a SIM scanner for photobleaching. The frame in which the bleach is applied often has high background. This frame can be deleted. The Bio-Formats plugin also often imports a few extraneous files into the stack that should be deleted before proceeding.

### 4. *Selecting ROIs for FRAP analysis*

- a. Open the ROI Manager in ImageJ: Once the stack has been opened and cleaned up, open the ROI Manager [ImageJ→Analyze→Tools→ROI Manager].
- b. Define bleach zone ROI: Navigate to the frame immediately after the bleach pulse was applied. Select the elliptical marquee tool in ImageJ from the main ImageJ window, and drag out a ROI that encompasses the bleach region. Once selected, add it to the ROI Manager list by typing [t].
- c. Define control ROI: A second ROI must be selected which serves as a control for the overall bleaching of the specimen during postbleach filming. This ROI can be a large region of the embryo. For FRAP of junctions, we typically select a large, but comparable, region containing the unbleached junctions. Once selected, add it to the ROI Manager list by typing [t].
- d. Select the bleach zone ROI in the ROI Manager list to reselect it.

### 5. *Output of the plugin*

1. Run the FRAP Profiler plugin [ImageJ→Plugins→ (location where the plugin was installed)]. A dialog will appear with pull-down menus for selecting various options. These include:
  - i. Single or double exponential curve fitting. Choose the latter if it appears, or there is physical reason to suspect, a two-step recovery process.
  - ii. The real-world time interval between successive images in the movie.
  - iii. The option to produce an image showing the bleach ROI and the control ROI overlaid on the first postbleach image.

2. The FRAP Profiler plugin generates several windows (Fig. 7C):
  - i. A log file with the parameters of the curve fit (half-life, immobile fraction, and constants). For a single exponential curve fit, the form of the equation with constants is  $p[0]*(1-e^{-p[1]*x})$ .
  - ii. If the option was selected, a window containing an image showing the two ROIs.
  - iii. Graphs of (i) the raw FRAP data; (ii) normalized FRAP data; and (iii) curve fits of the normalized data, including curves with the first postbleach time point set as the origin (i.e.,  $t = 0$ ). By clicking the “List . . .” button, the data in the latter can be exported to a spreadsheet or curve-fitting program for further analysis.

#### **D. Correlative Fluorescence and TEM During Morphogenesis (see Sims and Hardin, 2007 for more details)**

##### 1. Introduction

This final section describes high-pressure freezing (HPF) techniques for correlative light and electron microscopy on the same sample, starting with embryos mounted for 4D microscopy. Fluorescence information from a whole mount can be displayed as a color overlay on transmission electron microscopy (TEM) images to generate what we have termed fluorescence-integrated TEM (F-TEM) images. An alternative for postembedding correlative TEM is described elsewhere (Sims and Hardin, 2007). The method we describe here uses a thin two-part agarose pad to immobilize live *C. elegans* embryos for LSCM, HPF, and TEM. Pre-embedding F-TEM images display fluorescent information collected from a whole mount of live embryos onto all thin sections collected from that sample. For typical uses in our laboratory, this method relies on creating a strain with a rescuing array containing a GFP tagged protein of interest, often rescuing a lethal mutation (e.g., *ajm-1*; Köppen *et al.*, 2001). Embryos that have lost the rescuing transgene arrest. Alternatively, transgenes can be used as simple markers to identify the genotype of embryos. Because *C. elegans* embryos are difficult to fix with conventional chemical fixation, HPF is the method of choice. The agarose mount we describe for embryos is used for ultrastructural analysis, using freeze substitution with 1% osmium and 0.1% uranyl acetate. The agarose pad is made of a thin base of high-strength agarose. The thin base layer provides the strength and toughness to keep the mount intact. An agarose pad composed only of low-melting agarose would not hold the pad together well enough to allow transfer to a HPF specimen carrier. Additionally, it is necessary for the top agarose layer to be very thin to allow imaging in a confocal microscope. The method described here is specifically designed to be used at the conclusion of a 4D experiment. An alternative correlative procedure has been described elsewhere (Kolotuev *et al.*, 2010).

## 2. Materials

### General Materials

1. Glass slides, glass coverslip (22-mm square), and standard implements for preparing embryos for agar mounts (see the first section of this chapter).
2. High-pressure freezing specimen carriers.
3. 1-hexadecene. 1-hexadecene is used to coat the top specimen carrier to promote release of carriers after freezing and as filler in the bottom carrier as described by McDonald (1999)
4. High-pressure freezer (we used the BAL-TEC HPM 010).
5. Petri dishes to make a humid chamber (95 and 60-mm diam. dishes).
6. 20-mL scintillation vials for resin infiltration.
7. 3-Aminopropyltriethoxy-silane, 3-APTS (Sigma, A-3648).
8. Agarose (1% gel strength of 1000 g/cm<sup>2</sup> or greater for base; Invitrogen, cal. no. 15510-027). The strength of agarose can be tested empirically by making a thin pad of agarose over a slide. If the entire pad can be lifted off the slide without tearing, using a razor blade under one corner of the pad, it is strong enough to be used as a base for the correlative pad. We start with a 5% solution of agarose and dilute to between 4% and 5% with additional 0.1M HEPES as needed to form the base pad.
9. Low melting temperature agarose (Sigma, A-9539) to immobilize embryos.
10. 2 mL polypropylene vials with screw cap lids for freeze substitution.
11. Styrofoam box, dry ice, rotary shaker in 4 °C cold room. Freeze substitution using a Styrofoam box has been described by McDonald (1999). We fit the aluminum block into a tight-fitting piece of Styrofoam, which holds the block in place as the dry ice sublimates. We do not monitor the temperature during freeze substitution.
12. Disposable polyethylene pipettes (Fisher, cat no. 12-711-7). Any 3 mL or smaller disposable pipette with its own bulb should suffice. The use of disposable pipettes for dispensing fixatives avoids contaminating more expensive pipettes.
13. Microtiter plate shaker and smaller Styrofoam box in -20 °C freezer.
14. Single-edged razor blade. A single-edged razor blade is used to cut out a small piece of agarose containing the embryos to fit within the 2 mm specimen holder. The edge of a razor blade or a fine-tipped weighing spatula can also be used to transfer to the specimen carrier.
15. Sharpened tooth picks.
16. Rain-X (UnelkoCorp., Scottsdale, AZ) to coat slides for flat embedding. Two standard 1 × 3-inch microscope slides are coated with Rain-X or Teflon release agent using a cotton tipped applicator or Kimwipe. Coat each slide 3× and buff clear with a Kimwipe. This coating prevents the epoxy resin from gluing the slides together.
17. Clear acetate tape (Scotch brand or similar) for use as a spacer.

18. Richardson's stain for thick sections. To prepare, make: A. 1% methylene blue in 1% borax (w/v in dH<sub>2</sub>O) and B. 1% Azure II in dH<sub>2</sub>O. Mix equal volumes of A and B and apply to sections with a syringe equipped with a syringe filter to remove precipitates.

### Media

1. Use 0.1 HEPES (*N*-Hydroxyethylpiperazine-*N'*-2-ethanesulfonate) buffer to prepare agarose. A stock solution of 1M HEPES buffer is prepared by the addition of a solution of HEPES acid to a solution of HEPES base (sodium salt): 13.01 g of HEPES sodium salt is added to 50 mL of dH<sub>2</sub>O.
2. The freeze substitution medium is 1% osmium tetroxide with 0.1% uranyl acetate in acetone (McDonald, 1999). Briefly, to prepare 25 mL of 1% osmium with 0.1% uranyl acetate, cool 24 mL of EM-grade acetone in a disposable 50 mL polypropylene tube on crushed dry ice. If pure OsO<sub>4</sub> crystals are not consolidated in the bottom of the vial, freeze the unopened vial of solid osmium in liquid nitrogen. Osmium crystals will fall to the bottom of the ampule. Add 1 to 2 mL of the cold acetone to the ampule, mix, and add back to the 50 mL tube of acetone on dry ice. Repeat until all of the osmium is dissolved in 25 mL of acetone. Add the UA in methanol (0.025 g UA in 1 mL of methanol) to the acetone, keep cold on dry ice. Add 1 mL of freeze substitution mix (1% osmium tetroxide with 0.1% uranyl acetate in acetone) to each 2 mL substitution vial and freeze in liquid nitrogen.

### 3. Methods

- a. *An Agarose Mount for Live Embryos*
  - i. A thin (high-strength) agarose pad is formed over a standard glass microscope slide. Agarose is dissolved in 0.1M HEPES (neutral pH) buffer to a final concentration of 4% to 5%. The thickness of the pad can be controlled by adding a single layer of cellulose tape over two slides on either side of the slide to be coated. Add 100  $\mu$ L of melted agarose to the top of the center slide and compress the hot agarose to the thickness of a layer of Scotch tape (approximately 60  $\mu$ m) with a fourth slide resting on the two adjacent tape-covered slides. Allow the agarose to solidify before sliding the top slide off. Place the slide with the agarose pad in a humid chamber. A humid chamber can be made by placing a 25-mm diameter Petri dish inside a 95-mm Petri dish and adding water to cover the bottom of the large dish.
  - ii. *C. elegans* embryos are obtained by cutting open gravid hermaphrodites in a watch glass filled with distilled water as described for standard agar mounts. Only one group of approximately 10 embryos should be used. This grouping typically fills the field of view when using a 60–63 $\times$  objective. Larger groups of embryos are impractical and difficult to navigate in the TEM.
  - iii. Place 70  $\mu$ L of 5% low melting temperature agarose dissolved in 0.1M HEPES along one edge of the agar pad. Quickly position a glass coverslip

over the low-melt agarose to spread it out before it solidifies. Spacers can be used on either side of the slide to obtain the correct thickness. Use the edge of a slide resting on two spacer slides on either side of the slide containing embryos to apply pressure to the coverslip and compress the agarose to a uniform thickness. Ideal mount thickness is 100  $\mu\text{m}$ , which is the thickness of the smallest HPF specimen carrier configuration. To obtain agarose of the right consistency, heat the low melting temperature agarose to boiling in a glass test tube. The agarose is most easily dispensed by cutting 5–10 cm off the end of a yellow pipette tip and preheating the tip by rotating in the hot agarose. Fill the tip with 70  $\mu\text{L}$  of hot agarose and then dispense the agarose in a line along the edge side of the slide, over the high-strength agarose. The actual volume of agarose dispensed on the pad is less than 70  $\mu\text{L}$ , as about half (or more) of the agarose stays in the tip. Use the 70  $\mu\text{L}$  as a starting point and adjust as needed. This small volume cools rapidly, so the coverslip must be quickly placed over the low melting temperature agarose and gently pressed down to spread a thin layer of agarose around the embryos. A microscope slide turned on edge can be used to apply pressure to the coverslip on either side of the embryos but not directly over the embryos. The top layer of agarose is thinner than the high melting temperature agarose used to make the base. If the mount is too thick, embryos will be beyond the focal depth of the confocal microscope, making it impossible to acquire a fluorescent image. Note that covering embryos with hot low melting temperature agarose does not work with embryos that have been bleached, because they are too fragile.

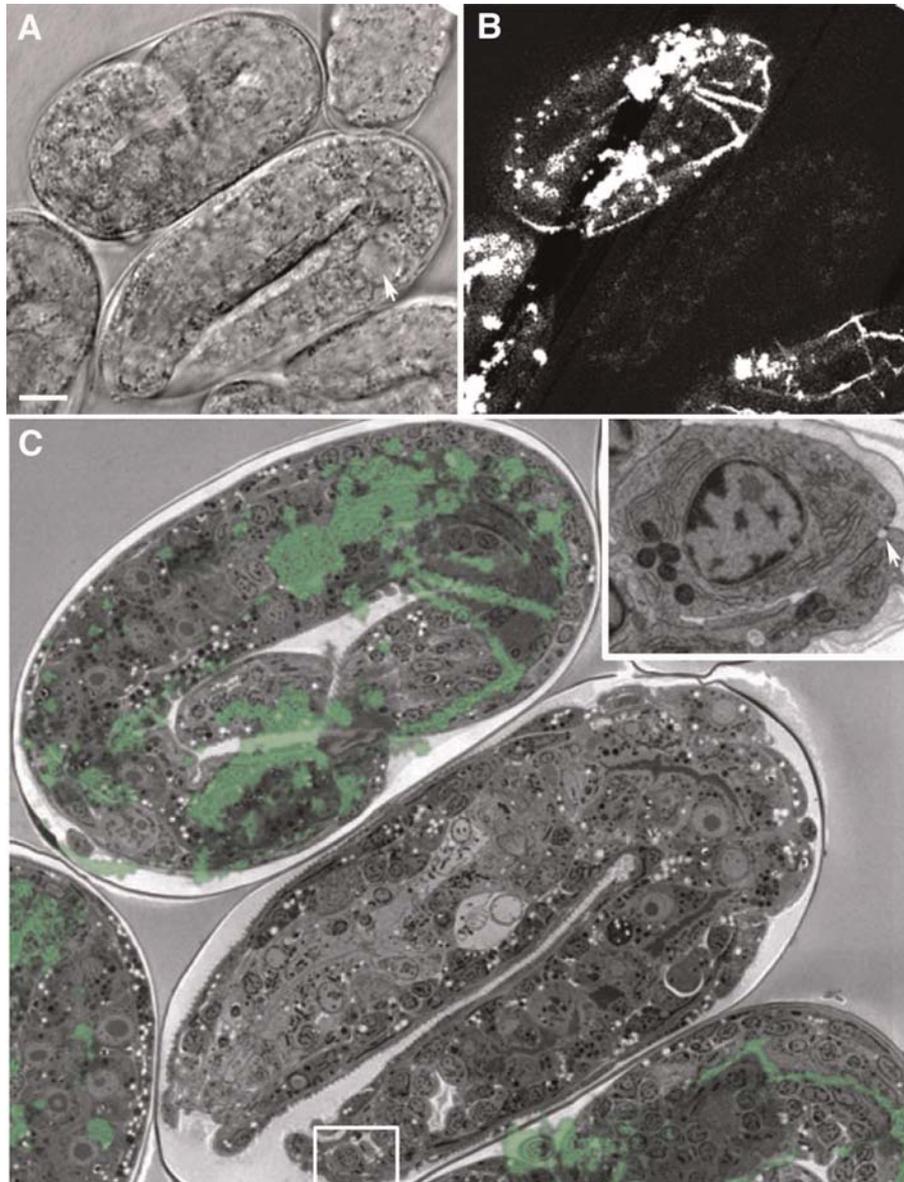
- iv. Seal the edges of the slide with Valap to prevent dehydration.
  - v. Acquire a focal series through the group of embryos at 1  $\mu\text{m}$  intervals using the appropriate excitation wavelength (488 nm for GFP). A transmitted light image can be acquired simultaneously with each fluorescent image of GFP expression. In our experiments, embryos must attain a minimum age to be able to identify “mutant” from rescued embryos. The transmitted light image can be used to confirm the developmental stage of embryos.
- b. HPF and Freeze Substitution
- i. Configure HPF specimen carriers to provide a 100- or 200- $\mu\text{m}$  deep well. To allow access to the agarose-immobilized embryos, scrape the Valap off the slide (which has already been imaged in a confocal microscope), and push the coverslip horizontally off the agarose pad using a single-edged razor blade. Cut out a small square of the agarose including the embedded embryos (maximum size is 2 mm diameter) and transfer this small pad to a bottom specimen carrier using a razor blade. The agarose pad can be pushed into a specimen carrier using a sharpened toothpick that has been coated with 1-hexadecene to keep the toothpick from adhering to the agarose pad. Loading tiny samples into specimen carriers becomes a battle with surface tension. Adhering carriers to the tops of Petri dishes with double stick tape holds carriers in place during the loading process. Double

stick tape is available in two types: “permanent” and a less aggressive “removable” (Scotch 667 from 3M). The less aggressive tape holds carriers without having to fight to remove them for loading in the freezing holder, just prior to freezing. The specimen carrier can be filled with bacteria and 1-hexadecene.

- ii. HPF is preformed in a BAL-TEC HPM 010. Readers unfamiliar with HPF and general issues related to successful specimen preparation should consult McDonald (1999), including filling of specimen carriers, removal of the specimen holder, and separation of specimen carriers.
- iii. The use of a brass bottom and aluminum top (or vice versa) allows for quick identification of the bottom carrier, which is transferred, while under LN<sub>2</sub>, to a polypropylene freeze substitution vial filled with frozen 1% osmium tetroxide and 0.1% uranyl acetate in acetone. The vial is capped and transferred to an aluminum block also cooled in liquid nitrogen.
- iv. The empty holes in the aluminum block are filled with liquid nitrogen and the block is wrapped in aluminum foil and packed with crushed dry ice in a Styrofoam box taped to a rotary shaker. The box is shaken at 100 rpm for 3 to 4 days at  $-80^{\circ}\text{C}$ .
- v. Shake vials on a microtiter plate shaker at 100 rpm for 2 to 3 days at  $-20^{\circ}\text{C}$ .
- vi. Warm to  $4^{\circ}\text{C}$  overnight, transfer to room temperature, and rinse with three to four changes of dry acetone. Agarose blocks can usually be identified by the presence of the embryos.

c. *Epoxy Infiltration and Polymerization for Agarose-Embedded Embryos*

- i. Transfer freeze substituted specimens into 20-mL scintillation vials (wash and oven dry) containing 30% Epon in acetone (EM grade) rotating on a rotary mixer for 4 h to overnight at room temperature.
- ii. 50% and 75% Epon in acetone for 2 h each at room temperature.
- iii. Three changes of 100% Epon for 1 h each at  $50-60^{\circ}\text{C}$ .
- iv. Transfer the resin-infiltrated agarose pad to a Rain-X or Teflon coated slide with some fresh resin. Place two thicknesses of Parafilm on both ends of the slide as a spacer and place a second coated (Rain-X) slide over the Epon-infiltrated agarose.
- v. Polymerize on a flat surface in a  $60^{\circ}\text{C}$  oven for 24 to 48 h.
- vi. Remove the resin from between the slides, rough up one side by scraping a razor blade across the surface, and mount/glue on a blank Epon block for sectioning.
- vii. One side of the now polymerized agarose pad containing embryos is roughed up to remove residual Rain-X or Teflon release agent and to increase the surface area for adhesion to an Epoxy blank. The small sample piece can be attached to the blank using glue or additional Epoxy resin followed by polymerization in a  $60^{\circ}\text{C}$  oven. This orients the embryos parallel with the cutting plane to obtain a similar orientation in the TEM as the view already obtained by LSCM.



**Fig. 8** (A) A transmitted light image of live *ajm-1(ok160);ajm-1::gfp* embryos embedded in agar. An embryo lacking *ajm-1* function has developed a characteristic vacuole (arrow). (B) A maximum intensity projection of AJM-1::GFP expression in the same embryos. Transmitted light and fluorescent images were obtained simultaneously on a Bio-Rad 1024 confocal using 488 nm excitation. Embryos that have lost the transgene can be easily distinguished from those that express it. (C) A TEM image of the same embryos after HPF and embedding in Epon. The fluorescence signal can be superimposed on the low-magnification transmission electron micrograph. The inset shows the small area within the box. The arrow points to epithelial cell membrane separations associated with the loss of *ajm-1* function. Bar = 5  $\mu$ m. (For color version of this figure, the reader is referred to the web version of this book.)

viii. Cut 0.5- $\mu\text{m}$  sections and place on 3-APTS coated coverslips. To coat 12-mm round coverslip with 3-APTS: clean 12-mm round coverslips in detergent solution. Rinse  $10\times$  in dH<sub>2</sub>O. Dehydrate using three changes in absolute ethanol. Coat slides in 2% solution of 3-APTS in dry acetone for several minutes. Rinse twice in dH<sub>2</sub>O. Spread out coverslips on clean filter paper and dry at room temperature. Stain thick sections with Richardson's stain. Rinse coverslips gently after staining with distilled water. Cut thin and/or semithin sections after embryos are detected in thick sections by light microscopy.

d. *TEM*

Collect thin sections on formvar-coated slot grids to provide an unobstructed view of an entire section. Stain thin and ultrathin sections for 10–20 min in 1% aqueous uranyl acetate, followed by 3 min in Reynolds' lead before viewing in a TEM. Thinner sections may require longer staining times. Large montage images can be collected manually and "stitched together" in Photoshop (Adobe) or collected and montaged automatically with analySISTM or similar software. When comparing different embryos within the same mount, it is helpful to make a map using a transmitted light image of the embedded group of embryos. Number the "mutant" and rescued embryos to keep track of higher magnification images. Because a TEM image may be a mirror image of the transmitted light image, flip the image in Photoshop with the same numbering scheme to be prepared for either orientation.

e. *Overlaying Correlative Microscopic Images*

A single transmitted light image is typically chosen from the multiple focal planes acquired based on features of interest that were in focus. The fluorescent image is a brightest point  $z$  projection (ImageJ) of all fluorescent images acquired. Because embryos are alive at the time of image capture, and hence may be moving, the fluorescent image(s) may be blurred. The detection of GFP confirms the presence of the rescuing DNA and definitively identifies the genotype of embryos. Confocal fluorescent and transmitted light images are manipulated with ImageJ and Adobe Photoshop. LSCM images are prepared using ImageJ. Some of the light (LSCM) and EM images collected may be mirror images of each other; they can be corrected in Photoshop. LSCM and TEM images are aligned starting with the TEM image, which usually contains more pixels. Fig. 8 demonstrates the results of this correlative method.



### III. Summary

*C. elegans* embryos are a powerful model system for imaging detailed cell movements, intercellular dynamics, and the overall shape changes that occur during morphogenesis. With the widespread use of genetically encoded fluorescent markers and the ability to perform correlative electron microscopy on transgenic embryos, the virtues that originally led to the selection of *C. elegans* as a model organism have been expanded. In the future, other extensions of live imaging, such as *in vivo*

fluorescence resonance energy transfer (FRET) and super resolution microscopy, will be added to the routine repertoire of developmental biologists using *C. elegans* to study morphogenesis.

## Acknowledgments

The author is grateful to T. Walston (Truman State University) and Paul Heid for permission to adapt a previously published protocol for agar mounts and Nomarski microscopy, E.A. Cox (SUNY Geneseo), A. Lynch, and S. Maiden for permission to adapt their protocols for phalloidin staining, and P. Sims for permission to adapt previously described protocols for F-TEM. M. Sheffield, T. Loveless, and R. Zaidel-Bar kindly provided additional images. This work was supported by NIH grant GM58038.

## References

- Achilleos, A., Wehman, A. M., and Nance, J. (2010). PAR-3 mediates the initial clustering and apical localization of junction and polarity proteins during *C. elegans* intestinal epithelial cell polarization. *Development* **137**, 1833–1842.
- Audhya, A., Hyndman, F., McLeod, I. X., Maddox, A. S., Yates 3rd, J. R., Desai, A., and Oegema, K. (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*. *J. Cell. Biol.* **171**, 267–279.
- Chisholm, A.D., Hardin, J. (2005). Epidermal morphogenesis. In WormBook, ed. The *C. elegans* Research Community, <http://www.wormbook.org>, pp. 1–22.
- Collins, T. J. (2007). Image J for microscopy. *Biotechniques* **43**, 25–30.
- Costa, M., Raich, W., Agbunag, C., Leung, B., Hardin, J., and Priess, J. R. (1998). A putative catenin-cadherin system mediates morphogenesis of the *Caenorhabditis elegans* embryo. *J Cell Biol* **141**, 297–308.
- Fridolfsson, H. N., and Starr, D. A. (2010). Kinesin-1 and dynein at the nuclear envelope mediate the bidirectional migrations of nuclei. *J. Cell. Biol.* **191**, 115–128.
- Gally, C., Wissler, F., Zahreddine, H., Quintin, S., Landmann, F., and Labouesse, M. (2009). Myosin II regulation during *C. elegans* embryonic elongation: LET-502/ROCK, MRCK-1 and PAK-1, three kinases with different roles. *Development* **136**, 3109–3119.
- Hardin, J. (2006). Confocal and multi-photon imaging of living embryos. In “Handbook of Biological Confocal Microscopy,” (and J. Pawley, ed.), pp. 746–768. Plenum, New York.
- Heid, P. J., and Hardin, J. (2000). Cell lineage analysis. Videomicroscopy techniques. *Methods Mol. Biol.* **135**, 323–330.
- Heid, P. J., Raich, W. B., Smith, R., Mohler, W. A., Simokat, K., Gendreau, S. B., Rothman, J. H., and Hardin, J. (2001). The zinc finger protein DIE-1 is required for late events during epithelial cell rearrangement in *C. elegans*. *Developmental Biology* **236**, 165–180.
- Heid, P. J., Voss, E., and Soll, D. R. (2002). 3D-DIASemb: A computer-assisted system for reconstructing and motion analyzing in 4D every cell and nucleus in a developing embryo. *Dev. Biol.* **245**, 329–347.
- Kolotuev, I., Schwab, Y., and Labouesse, M. (2010). A precise and rapid mapping protocol for correlative light and electron microscopy of small invertebrate organisms. *Biol. Cell.* **102**, 121–132.
- Köppen, M., Simske, J. S., Sims, P. A., Firestein, B. L., Hall, D. H., Radice, A. D., Rongo, C., and Hardin, J. D. (2001). Cooperative regulation of AJM-1 controls junctional integrity in *Caenorhabditis elegans* epithelia. *Nat. Cell. Biol.* **3**, 983–991.
- Labbe, J. C., McCarthy, E. K., and Goldstein, B. (2004). The forces that position a mitotic spindle asymmetrically are tethered until after the time of spindle assembly. *J. Cell. Biol.* **167**, 245–256.
- Lippincott-Schwartz, J., Altan-Bonnet, N., and Patterson, G. H. (2003). Photobleaching and photoactivation: Following protein dynamics in living cells. *Nat. Cell. Biol. Suppl.* S7–S14.
- Liu, Z., Nukazuka, A., and Takagi, S. (2007). Improved method for visualizing cells revealed dynamic morphological changes of ventral neuroblasts during ventral cleft closure of *Caenorhabditis elegans*. *Dev. Growth Differ.* **49**, 49–59.

- Lockwood, C., Zaidel-Bar, R., and Hardin, J. (2008). The *C. elegans* zonula occludens ortholog cooperates with the cadherin complex to recruit actin during morphogenesis. *Curr. Biol.* **18**, 1333–1337.
- McDonald, K. (1999). High-pressure freezing for preservation of high resolution fine structure and antigenicity for immunolabeling. *Methods Mol. Biol.* **117**, 77–97.
- Miller, D. M., and Shakes, D. C. (1995). Immunofluorescence microscopy. *Methods Cell. Biol.* **48**, 365–394.
- Mohler, W., and White, J. (1998). Multiphoton laser scanning microscopy for four-dimensional analysis of *Caenorhabditis elegans* embryonic development. *Opt. Express* **3**, 325–331.
- Mohler, W. A. (1999). Visual reality: Using computer reconstruction and animation to magnify the microscopist's perception. *Mol. Biol. Cell.* **10**, 3061–3065.
- Mohler, W. A., and Isaacson, A. B. (2010). Suspended embryo mount for imaging *Caenorhabditis elegans*. *Cold Spring Harb. Protoc.* [pdb prot5388](https://doi.org/10.1101/2010.09.01.35388).
- Mohler, W. A., Simske, J. S., Williams-Masson, E. M., Hardin, J. D., and White, J. G. (1998). Dynamics and ultrastructure of developmental cell fusions in the *Caenorhabditis elegans* hypodermis. *Curr. Biol.* **8**, 1087–1090.
- Murray, J. I., Bao, Z., Boyle, T. J., Boeck, M. E., Mericle, B. L., Nicholas, T. J., Zhao, Z., Sandel, M. J., and Waterston, R. H. (2008). Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat. Methods* **5**, 703–709.
- Murray, J. I., Bao, Z., Boyle, T. J., and Waterston, R. H. (2006). The lineaging of fluorescently-labeled *Caenorhabditis elegans* embryos with StarryNite and AceTree. *Nat. Protoc.* **1**, 1468–1476.
- Pettitt, J., Cox, E. A., Broadbent, I. D., Flett, A., and Hardin, J. (2003). The *Caenorhabditis elegans* p120 catenin homologue, JAC-1, modulates cadherin-catenin function during epidermal morphogenesis. *J. Cell Biol.* **162**, 15–22.
- Pohl, C., and Bao, Z. (2010). Chiral forces organize left–right patterning in *C. elegans* by uncoupling midline and anteroposterior axis. *Developmental Cell.* **19**, 402–412.
- Raich, W. B., Agbunag, C., and Hardin, J. (1999). Rapid epithelial-sheet sealing in the *Caenorhabditis elegans* embryo requires cadherin-dependent filopodial priming. *Curr. Biol.* **9**, 1139–1146.
- Sheffield, M., Loveless, T., Hardin, J., and Pettitt, J. (2007). *C. elegans* Enabled exhibits novel interactions with N-WASP, Abl, and cell-cell junctions. *Curr. Biol.: CB* **17**, 1791–1796.
- Shelton, C. A., and Bowerman, B. (1996). Time-dependent responses to glp-1-mediated inductions in early *C. elegans* embryos. *Development* **122**, 2043–2050.
- Sims, P. A., and Hardin, J. D. (2007). Fluorescence-integrated transmission electron microscopy images: Integrating fluorescence microscopy with transmission electron microscopy. *Methods Mol. Biol.* **369**, 291–308.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
- Thomas, C., DeVries, P., Hardin, J., and White, J. (1996). Four-dimensional imaging: Computer visualization of 3D movements in living specimens. *Science* **273**, 603–607.
- Thomas-Virnic, C. L., Sims, P. A., Simske, J. S., and Hardin, J. (2004). The inositol 1,4,5-trisphosphate receptor regulates epidermal cell migration in *Caenorhabditis elegans*. *Curr. Biol.* **14**, 1882–1887.
- Velarde, N., Gunsalus, K. C., and Piano, F. (2007). Diverse roles of actin in *C. elegans* early embryogenesis. *BMC Dev. Biol.* **7**, 142.
- Walston, T., and Hardin, J. (2010). Laser killing of blastomeres in *Caenorhabditis elegans*. *Cold Spring Harb. Protoc.* [pdb prot5543](https://doi.org/10.1101/2010.09.01.35543).
- Williams-Masson, E. M., Malik, A. N., and Hardin, J. (1997). An actin-mediated two-step mechanism is required for ventral enclosure of the *C. elegans* hypodermis. *Development* **124**, 2889–2901.
- Willis, J. H., Munro, E., Lyczak, R., and Bowerman, B. (2006). Conditional dominant mutations in the *Caenorhabditis elegans* gene *act-2* identify cytoplasmic and muscle roles for a redundant actin isoform. *Mol. Biol. Cell.* **17**, 1051–1064.
- Zaidel-Bar, R., Joyce, M. J., Lynch, A. M., Witte, K., Audhya, A., and Hardin, J. (2010). The F-BAR domain of SRGP-1 facilitates cell–cell adhesion during *C. elegans* morphogenesis. *J. Cell Biol.* **191**, 761–769.

---

---

## CHAPTER 15

# Molecular and Genetic Approaches for the Analysis of *C. elegans* Neuronal Development

**Dong Yan and Yishi Jin**

Howard Hughes Medical Institute, Neurobiology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, California, USA

---

- I. Introduction
  - A. Overview of Nervous System Anatomy
  - B. Overview of Nervous System Development
- II. Strategies for *In Vivo* Labeling of Neurons and Compartments
  - A. Constructing Transgenic Reporters to Visualize Neuronal Morphology
  - B. Labeling Neuronal Proteins and Subcellular Structures
  - C. Complex Strategy for Labeling and Manipulating Unique Neurons
  - D. Questions and Answers
- III. Other Methods for Examining the Development of Neurons
  - A. Dye-Filling
  - B. Immunocytochemistry
  - C. Electron Microscopy (EM)
- IV. Functional Dissection of Signaling Pathways Controlling Neuronal Development
  - A. Phenotypic Characterizations
  - B. Identifying Mutants in Forward Genetic Screens
  - C. Gene Expression Profiling of Neurons
  - D. Laser Surgery of Cells and Processes
  - E. Questions and Answers
- V. Examples of Neuronal Development
  - A. Neuronal Fate Specification and Differentiation
  - B. Axon Guidance and Dendrite Morphogenesis
  - C. Synapse Formation and Specificity
  - D. Maintenance of Neuronal Architecture
- VI. Outlook
  - References

---

---

## Abstract

The complete known anatomical connections of *C. elegans* nervous system have provided researchers ample opportunities to discover fundamental principles underlying neuronal development. Transgenic labeling with fluorescent proteins in neuronal cells has had an unparalleled impact on our ability to visualize cellular architecture and dynamics. In this chapter, we summarize the common methods and guidelines for dissecting the molecular and cellular mechanisms controlling nervous system formation. We end with a brief description of several applications that illustrate the use of these methods.

---

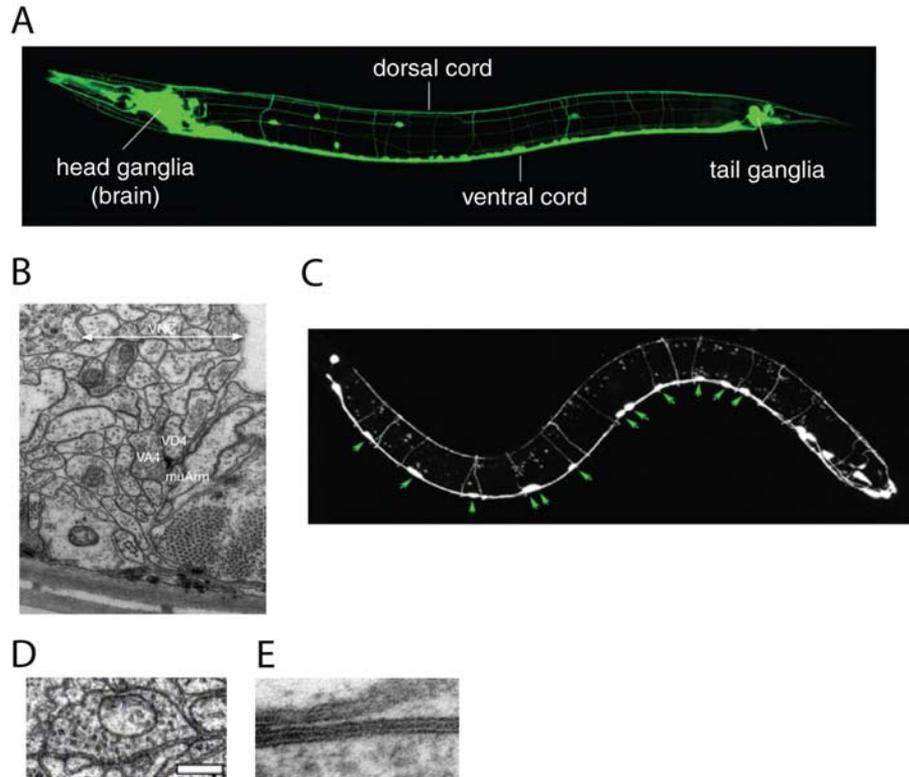
---

## I. Introduction

### A. Overview of Nervous System Anatomy

The nervous system is the most complex organ of *Caenorhabditis elegans*. An adult hermaphrodite has 302 neurons and 56 glial cells, which make up almost 40% of the somatic cells. An adult male has 383 neurons, 294 of which are common with hermaphrodites and 89 are specific to males. The structure of the hermaphrodite nervous system was described in unprecedented detail using electron microscopic serial reconstruction (White *et al.*, 1986). The structure of the male nervous system has also recently been completed by David Hall and Scott Emmons (personal communication). In hermaphrodites, 20 neurons reside in the pharynx, forming the pharyngeal nervous system, and 282 neurons are spatially organized into ganglia and nerve bundles forming the somatic nervous system (Fig. 1A). Based on morphology and position, the 302 neurons are classified into 118 types, and neurons of the same type usually exhibit two-, four-, or sixfold symmetry along the body axis (White *et al.*, 1986). The Wormatlas (<http://www.wormatlas.org>) provides accurate illustrations of individual neurons and up-to-date information on the function and molecular composition. Recent findings have now revealed functional and molecular differences between neurons in the same morphological group.

The soma of the majority of sensory neurons and interneurons are located in the ganglia in the head and tail. Each ganglion is surrounded by the basement membrane. The nerve processes of sensory neurons, interneurons, and anterior motor neurons form the major neuropil in the head, called the nerve ring, or the “brain,” where extensive information processing occurs (Fig. 1A). The ventral nerve cord (VNC) consists of a linear array of motor neuron soma and nerve processes of interneurons and motor neurons (Fig. 1A, B). The dorsal nerve cord is primarily composed of nerve processes of the motor neurons (Fig. 1A). The motor neurons provide excitatory and inhibitory input to body wall muscles. The peripheral sensory neurons include mechanosensory neurons and a few other neurons; their somas reside at various lateral positions along the body, and extend axons singly or as a small bundle (Fig. 1A). The 56 glial cells are divided into three classes: sheath (24 cells) and socket (26 cells) cells reside in the head ganglia and surround the



**Fig. 1** Overview of the *C. elegans* nervous system. (A) The *C. elegans* nervous system visualized by a pan-neuronal GFP. This image is courtesy of Dr. Harald Hutter. (B) Ultrastructure of the ventral nerve cord (VNC), shown as an electron micrograph of cross-section. This image is adapted with permission from White *et al.* (1986). (C) GFP labeling of the GABAergic motor neurons (Green arrowheads mark the ventral cord neurons); the image is adapted with permission from Knobel *et al.* (2001). (D) An electron micrograph of the ultrastructure of a chemical synapse between a motor neuron and muscle cell. This image is adapted with permission from Nakata *et al.* (2005). (E) An electron micrograph image of the ultrastructure of a gap junction between a hypodermal cell and excretory cells. This image is adapted with permission from Chuang *et al.* (2007). (For color version of this figure, the reader is referred to the web version of this book.)

dendrites of the amphid and phasmid sensory neurons, and GLR cells (6 cells) are located near the nerve ring and form a thin cylindrical sheet between the pharynx and the nerve ring. Recent studies have shown that these glial cells play important roles in the development and maintenance of the nervous system (Shaham, 2006).

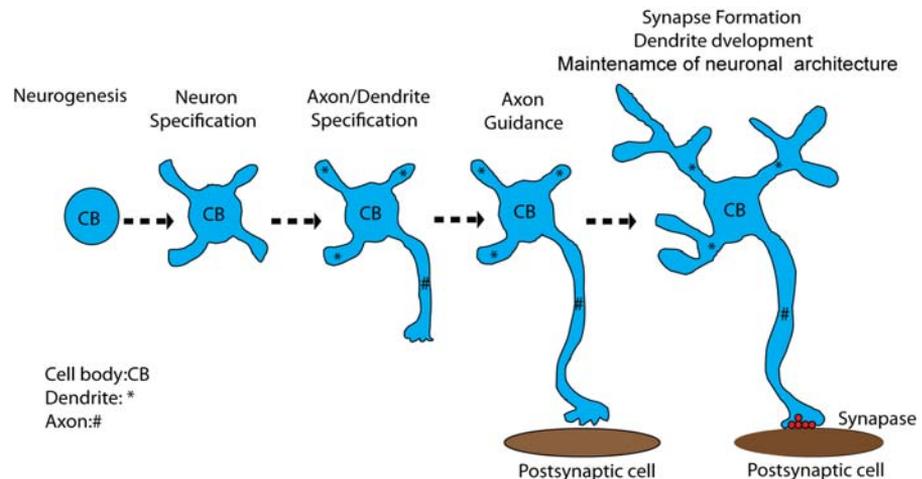
Within the nerve ring or bundle, nerve processes are positioned in a stereotypical neighborhood pattern. An example is shown in Fig. 1C, which shows the position and axonal trajectory of the GABAergic motor neurons. Chemical synapses are formed *en passant* between specific partners, either along the entire nerve or at defined regions

(Fig. 1 D). Gap junctions, or electrical synapses, are formed between selective neuron–neuron, neuron–muscle, or neuron–glial partners (Fig. 1 E). Reanalysis of the original EM micrographs have recently identified close to 7000 chemical synapses and 500 gap junctions; a web-based interactive wiring diagram is found on <http://www.wormatlas.org/neuronalwiring.html>. The cellular complexity and anatomical details make the *C. elegans* nervous system an ideal model system to study neuronal development *in vivo* at the resolution of single-cell and single-synapse.

## B. Overview of Nervous System Development

All neurons are descendants of the founder cell AB. Individual neurons are produced in an invariant lineage fashion. The identity of a neuron is determined based on the position, axon morphology, and unique features such as neurotransmitter type, dye-filling ability, and distinct molecular labels. Questions concerning nervous system development of the nervous system generally fall into several categories, reflecting the developmental sequences of a neuron (Fig. 2). The early steps of neuronal development concern neurogenesis and specification of neurons. The late steps address neuronal differentiation, which include a series of events from neuritogenesis, polarity determination, axon path finding or guidance, to synapse formation. Recently, the maintenance and plasticity of the nervous system have also emerged as a fascinating topic in the analysis of neuronal development.

Early approaches to the analysis of neuronal development relied heavily on lineage studies using Nomarski microscopy, combined with labeling methods such as dye filling and antibody immunostaining. Morphological changes in genetic



**Fig. 2** Illustration of the major steps of development from a neuronal precursor cell to a mature neuron. (For color version of this figure, the reader is referred to the web version of this book.)

mutants were correlated with behavioral defects, such as locomotion and chemotaxis. Ultrastructural examination was particularly instrumental to deciphering anatomical disruptions. These studies provided ground-breaking discoveries in the development of the nervous system, such as identifying UNC-6/Netrin axon guidance pathway, and discovering regulators of neuron specification.

The introduction of green fluorescent protein (GFP) from *Aequorea victoria* as a noninvasive *in vivo* reporter revolutionized the research in *C. elegans* (Chalfie *et al.*, 1994). Nearly all neurons and their molecular components or cellular compartments can be visualized using a plethora of transgenic fluorescent protein reporters in live animals. Combined with the awesome power of genetic analysis, *C. elegans* is a prime system for elucidating the molecular and functional logic of neuronal circuits. In this chapter, we summarize the general strategies for transgenically labeling neurons and cellular compartments, and provide guidelines for genetically dissecting genes and pathways controlling neuronal development.

---

---

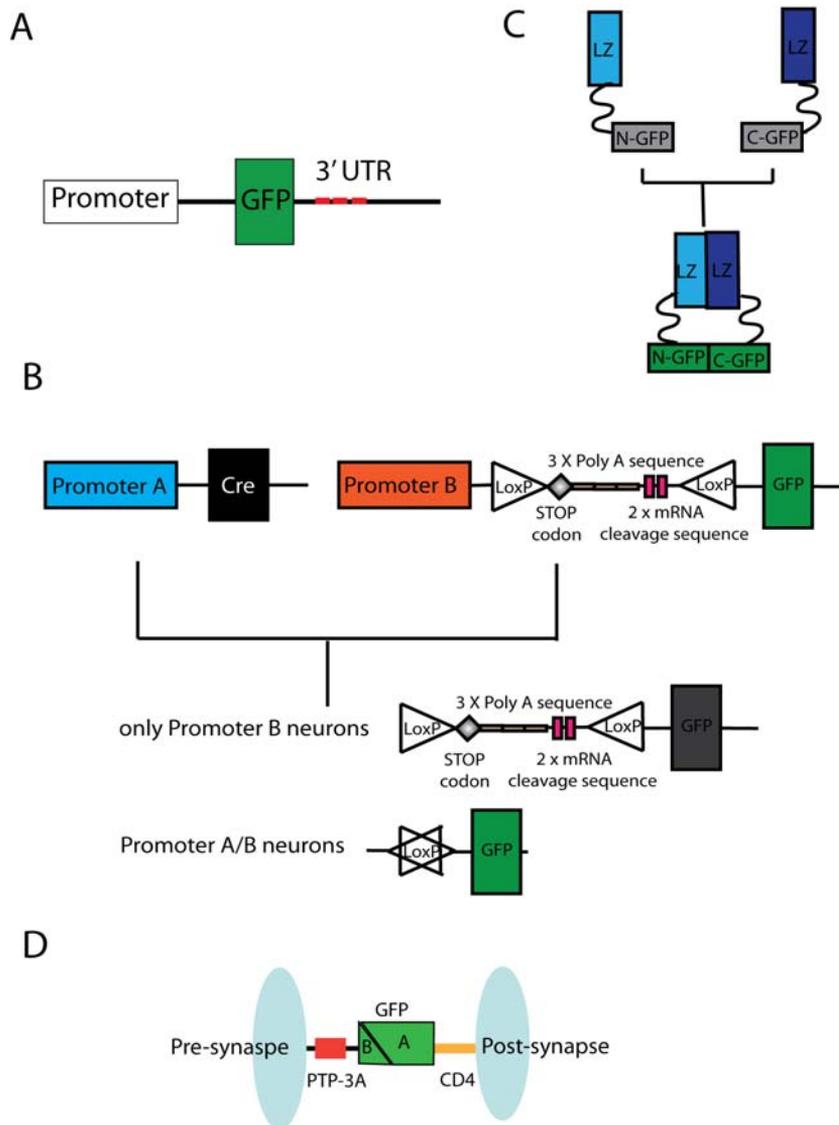
---

## II. Strategies for *In Vivo* Labeling of Neurons and Compartments

**Rationale:** The transparency of *C. elegans* and the ease of making transgenic animals offer great advantages for *in vivo* labeling of cells, tissues, and cellular components. The starting point to study the development of neurons of interest is to learn about the position, morphology, and unique features of the neurons. This entails the understanding of the cell lineage that generates the neurons and finding the promoters or regulatory DNA sequences that are selectively active in these neurons. Nearly every neuron in *C. elegans* can be labeled using promoter (see supplemental table in Chelur and Chalfie, 2007). Since the initial use of GFP, there are now over 58 fluorescent protein reporters (XFPs) that can be excited from a wide range of wavelengths and emit different colors. The most commonly used XFPs in *C. elegans* are GFP, YFP, CFP, mCherry, and more recently, tdTomato and mStrawberry. The preference for transgenic labeling is based on the question of interest, from simple promoter reporters to complex labeling of proteins and organelles. Below, we summarize the common strategies.

### A. Constructing Transgenic Reporters to Visualize Neuronal Morphology

The general design of reporter constructs began with Andy Fire's toolkit in the 1990s (Fire *et al.*, 1990). XFP coding sequences are optimized for the starting ATG and codon usage and inclusion of synthetic introns. Multiple cloning sites at 5' and 3' allow for insertion of promoters, making fusion proteins, or attaching 3' untranslated regions (3'UTR) (Fig. 3 A). Most Fire vectors contain the 3' UTR of the *unc-54* gene for stable mRNA production. These basic features have recently been transposed into the recombinase-based Gateway cloning system ([www.invitrogen.com](http://www.invitrogen.com)). Practical issues to consider when designing a reporter construct are: strength and



**Fig. 3** Overview of reporter construction strategies. (A) A simple construction scheme for labeling neurons. (B) The Cre-LoxP strategy for labeling and manipulating unique neurons; This illustration is based on Macosko *et al.* (2009). (C) The reconstituted GFP method. This illustration is based on (Zhang *et al.*, 2004). (D) The GFP reconstitution across synaptic partners (GRASP) method. This illustration is based on Feinberg *et al.* (2008). (For color version of this figure, the reader is referred to the web version of this book.)

specificity of the promoter, type of fluorescent protein reporters, choice of 3' UTR, forms of DNA for transformation (circular plasmids or linear DNA fragments), and method of generating transgenic worms. The following is a general guideline for making transgenic worms to label neurons by XFP.

(1) Reporter construction:

- (a) Choose a neuronal promoter: To label a specific neuron, one needs to find a promoter that can be strongly activated in the target neuron, but not in others. In *C. elegans*, most promoters refer to a few hundred base pairs to 3–5 kilobases of sequences 5' to the starting ATG. Extensive information regarding tissue and cell specificity of many promoters is provided in *wormbase* and *C. elegans* modENCODE (Celniker *et al.*, 2009; Gerstein *et al.*, 2010). For neuron-type specific promoters, a good place to start is the supplemental table in Chelur and Chalfie (2007).
- (b) Choose a reporter: Almost all fluorescent proteins are suitable for labeling neurons. One usually begins with GFP, as it expresses well in most cells and rarely forms aggregates. Most XFP reporters are usually stable. For analyzing dynamic regulation, several modifications help to destabilize fluorescent proteins. For example, the PEST domain is a target of the ubiquitin proteasome system (UPS), and experiments showed that adding the PEST domain to the C-terminus of GFP greatly reduces the half life of GFP (Frاند *et al.*, 2005; Li *et al.*, 1998). Similarly, tagging GFP with a ubiquitin E3 ligase Ring finger domain can destabilize GFP (Bounoutas *et al.*, 2010; Poyurovsky *et al.*, 2003)
- (c) Choose a 3' UTR: The 3'UTR plays important regulatory roles in gene expression. For stable and high-levels of expression, one of the most used 3'UTRs is from the *unc-54* myosin gene. Increasing evidence shows that important regulatory information may also be conveyed by the gene-specific 3' UTR. In general, one relies on cDNA clones to identify 3' UTRs. A recent genomic study has revealed extensive 3' UTR variants for a given gene (Mangone *et al.*, 2010), and a search in *wormbase* is always helpful when deciding which 3'UTR to include in a reporter design.

(2) Making transgenes:

- (a) Forms of DNA: Both circular plasmids and linear DNA fragments (such as amplified by PCR) work well in generating transgenic worms. In traditional germ line transformation procedure, the transgene expression levels can be controlled by the concentration of DNA (Mello *et al.*, 1991). Recent studies have suggested that using vector-free DNA fragments may improve transgene expression levels and tissue specificities (Etchberger and Hobert, 2008).
- (b) High-copy expression of transgenes by germ line transformation: The most common method to generate transgenic worms is microinjecting DNA of interest into the gonads of young adult hermaphrodites (Mello *et al.*, 1991). DNA usually forms extrachromosomal arrays that are stably transmitted, but

**Table 1**

Stable transgenic lines for labeling neurons

Alleles	Plasmids	Neurons	Linkage	Reference
<i>akIs3</i>	<i>Pnmr-1::GFP</i>	AVA, AVD, AVE, and PVC	LGX	(Zheng <i>et al.</i> , 1999)
<i>bwIs2</i>	<i>Pflp-1::GFP</i>	AVK	LGII	(Much <i>et al.</i> , 2000)
<i>evIs82</i>	<i>unc-129::gfp</i>	DA/DB neuron	LGIV	(Lim <i>et al.</i> , 1999)
<i>evIs111</i>	<i>Prgef-1::GFP</i>	Pan-neuronal	LGIII	(Pilon <i>et al.</i> , 2000)
<i>gmIs18</i>	<i>Pceh-23::GFP</i>	AIY BAG ASI ADL AWC ASE AFD ASH ASG	LGX	(Frank <i>et al.</i> , 2005)
<i>hdIs26</i>	<i>Psra-6::DsRed2Podr-1::GFP</i>	PVP	LGV	(Hutter, 2003)
<i>hmIs4</i>	<i>Pdes-2::GFP</i>	PVD		(Oren-Suissa <i>et al.</i> , 2010)
<i>jcIs1</i>	<i>Pajm-1::GFP</i>	PVD	LGIV	(Mohler <i>et al.</i> , 1998)
<i>juIs73</i>	<i>Punc-25::GFP</i>	DD VD RME	LGIII	(Hallam <i>et al.</i> , 2000)
<i>juIs76</i>	<i>Punc-25::GFP</i>	DD VD RME	LGII	(Huang <i>et al.</i> , 2002)
<i>kyIs4</i>	<i>Pceh-23::GFP</i>	AIY BAG ASI ADL AWC ASE AFD ASH ASG	LGX	(Forrester and Garriga, 1997)
<i>kyIs5</i>	<i>Pceh-23::GFP</i>	AIY BAG ASI ADL AWC ASE AFD ASH ASG	LGIV	(Zipkin <i>et al.</i> , 1997)
<i>kyIs8</i>	<i>Pceh-23::GFP</i>	AIY BAG ASI ADL AWC ASE AFD ASH ASG	LGI	(Lundquist <i>et al.</i> , 2001)
<i>kyIs104</i>	<i>Pstr-1::GFP</i>	AWB	LGX	(Sagasti <i>et al.</i> , 1999)
<i>kyIs136</i>	<i>Pstr-2::GFP</i>	AWC <sup>ON</sup>	LGX	(Troemel <i>et al.</i> , 1999)
<i>kyIs140</i>	<i>Pstr-2::GFP</i>	AWC <sup>ON</sup>	LGI	(Troemel <i>et al.</i> , 1999)
<i>kyIs170</i>	<i>Psrh-220::GFP</i>	ADL	LGI	(Chang <i>et al.</i> , 2006)
<i>kyIs179</i>	<i>Punc-86::GFP</i>	HSN neurons	LGIV	(Shen and Bargmann, 2003)
<i>kyIs258</i>	<i>Podr-1::DsRed; Pofm-1::GFP</i>	AWC <sup>ON</sup> AWC <sup>OFF</sup>	LGX	(Vanhoven <i>et al.</i> , 2006)
<i>kyIs262</i>	<i>Punc-86::myr GFP; Podr-1::dRed</i>	HSN, AWC ASI AWB ASK ASJ	LGIV	(Chang <i>et al.</i> , 2006)
<i>kyIs323</i>	<i>Pstr-2::GFP; Punc-122::GFP</i>	AWC <sup>ON</sup> AWC <sup>OFF</sup>	LGII	(Bauer Huang <i>et al.</i> , 2007)
<i>kyIs408</i>	<i>Pstr-2::dsRed2; Psrsx-3::GFP</i>	AWC <sup>ON</sup> AWC <sup>OFF</sup>	LGII	(Lesch <i>et al.</i> , 2009)
<i>mgIs18</i>	<i>Pttx-3::GFP</i>	AIY	LGIV	(Duerr <i>et al.</i> , 1999)
<i>mgIs25</i>	<i>Punc-97::GFP</i>	Touch neurons		(Hobert <i>et al.</i> , 1999)
<i>muIs32</i>	<i>Pmec-7::GFP</i>	Touch neurons	LGII	(Pujol <i>et al.</i> , 2000)
<i>nlIs1</i>	<i>Pgcy-5::GFP</i>	ASER	LGV	(Sarafi-Reinach <i>et al.</i> , 2001)
<i>nuIs1</i>	<i>Pglr-1::GFP</i>	AVG AVJ DVC PVC PVQ RIG RIS RMD RME	LGX	(Hobert <i>et al.</i> , 1999)
<i>nuIs63</i>	<i>Pceh-24::GFP</i>			(Mehta <i>et al.</i> , 2004)
<i>nuIs208</i>	<i>Punc-129::GFP</i>	Motor neurons		(Sieburth <i>et al.</i> , 2007)
<i>otIs3</i>	<i>Pgcy-7::GFP</i>	ASEL	LGV	(Tsalik and Hobert, 2003)
<i>otIs6</i>	<i>Plim-6prom::GFP</i>	ASG AWA		(Chang <i>et al.</i> , 2003)
<i>otIs7</i>	<i>Pzig-2::GFP</i>	ASI		(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs14</i>	<i>Pzig-3::GFP</i>	AIM ASI		(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs24</i>	<i>Psre-1::GFP</i>	ADL		(Sarafi-Reinach <i>et al.</i> , 2001)

(Continued)

**Table I** (Continued)

Alleles	Plasmids	Neurons	Linkage	Reference
<i>otIs33</i>	<i>Pkal-1::GFP</i>	AIY, AIZ, RID, M5, ASI, motorneurons, midbody neurons HSN, CAN, PVM, DVB, DVC, PDB	LGIV	(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs39</i>	<i>Punc-47Δ::GFP</i>	PVT	LGII	(Tsalik and Hobert, 2003)
<i>otIs45</i>	<i>Punc-119::GFP</i>	Pan-neuronal	LGIV	(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs62</i>	<i>Psra-11-3::GFP</i>	AIY AVB		(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs85</i>	<i>PF59B2.13::GFP</i>	PVT		(Aurelio <i>et al.</i> , 2003)
<i>otIs90</i>	<i>Ppin-2::GFP</i>	PVT		(Aurelio <i>et al.</i> , 2003)
<i>otIs92</i>	<i>Pflp-10::GFP</i>	AIM ASI AUA BAG BDU DVB PQR PVR URX	LGIV	(Mehta <i>et al.</i> , 2004)
<i>otIs114</i>	<i>Plim-6prom::GFP</i>	ASG AWA	LGI	(Chang <i>et al.</i> , 2006)
<i>otIs123</i>	<i>Psra-11-3::GFP</i>	AIY AVB		(Altun-Gultekin <i>et al.</i> , 2001)
<i>otIs125</i>	<i>Pflp-6::GFP</i>	ASE AFD ASG PVT	LGX	(Chang <i>et al.</i> , 2006)
<i>otIs131</i>	<i>Pgcy-7::RFP</i>	ASEL		(Chang <i>et al.</i> , 2006)
<i>otIs151</i>	<i>Pceh-36::dsRed2</i>	AWC ASE	LGIV	(Johnston and Hobert, 2003)
<i>otIs173</i>	<i>Prgef-1::dsRed2</i>	Pan-neuronal	LGIII	(Boulin <i>et al.</i> , 2006)
<i>oxIs12</i>	<i>Punc-47::GFP</i>	GABAergic neurons	LGX	(McIntire <i>et al.</i> , 1997)
<i>oyIs14</i>	<i>Psra-6::GFP</i>	PVQ PVP	LGIV	(Sarafi-Reinach <i>et al.</i> , 2001)
<i>oyIs17</i>	<i>Pgcy-8::GFP</i>	AFD	LGIV	(Sarafi-Reinach <i>et al.</i> , 2001)
<i>oyIs44</i>	<i>Pord-1::dsRed</i>	AWC	LGIV	(Lanjuin <i>et al.</i> , 2003)
<i>oyIs45</i>	<i>Pord-1::YFP</i>	AWC	LGIV	(Bacaj <i>et al.</i> , 2008)
<i>oyIs51</i>	<i>Psrh-142::YFP</i>	ADF	LGIV	(Ortiz <i>et al.</i> , 2006)
<i>rhIs4</i>	<i>Pglr-1::GFP</i>	AVA AVB AVD AVE PVC	LGIV	(Lim <i>et al.</i> , 1999)
<i>syIs63</i>	<i>Pcog-1::GFP</i>	ADL, ASE, and ASJ		(Palmer <i>et al.</i> , 2002)
<i>syIs73</i>	<i>Pcog-1prom::GFP</i>	ASE		(Chang <i>et al.</i> , 2003)
<i>uls25</i>	<i>Pmec-18::GFP</i>	Touch neurons		(Schaefer <i>et al.</i> , 2000)
<i>vtIs1</i>	<i>Pdat-1::GFP</i>	ADE PDE CEPV CEPD	LGIV	(Nass <i>et al.</i> , 2002)
<i>wyIs75</i>	<i>Punc-47l::RFP; Pexp-1::GFP</i>	DA neurons		(Poon <i>et al.</i> , 2008)
<i>zdlIs1</i>	<i>Pceh-23::GFP</i>	AIY BAG ASI ADL AWC ASE AFD ASH ASG neurons	LGIV	(Chang <i>et al.</i> , 2006)
<i>zdlIs4</i>	<i>Pmec-4::GFP</i>	Touch neurons	LGIV	(Christensen and Strange, 2001)
<i>zdlIs5</i>	<i>Pmec-4::GFP</i>	Touch neurons	LGI	(Hao <i>et al.</i> , 2001)
<i>zdlIs13</i>	<i>Ptph-1::GFP</i>	HSN	LGIV	(Clark and Chiu, 2003)
<i>zdlIs21</i>	<i>Pzag-1::GFP</i>	neurons in head and tail ganglia	LGIV	(Clark and Chiu, 2003)

have variable expression levels and mosaic patterns. For stable expression, transgenic extrachromosomal arrays are often integrated into the genome through mutagenesis by UV, X-ray, or chemical mutagens. Expression levels of transgenes may be modified following integration, the cause of which is not fully understood. Table I lists many published stable transgenic fluorescent proteins lines that label different types of neurons.

- (c) Low-copy expression of transgenes by microparticle bombardment: This method has been widely used to generate integrated transgenes, especially for gene expression in early embryos and the germ line. With this technique, DNA is bound to gold particles and shot into worms using a biolistic bombardment instrument or gene gun (Biorad). The integrated transgenes usually contain only a few copies of the DNA of interest (Praitis *et al.*, 2001).
- (d) Single-copy expression of transgenes by Mos-mediated integration: MosSci (*Mos1* mediated single-copy transgene insertion) was developed based on the demonstration that the *Drosophila Mos* transposon is active in *C. elegans* (Bessereau *et al.*, 2001). *Mos1* encodes a transposase that is not present in the genetic background of *C. elegans*. Transposition is induced by coexpressing *Mos1* transposase and the *Mos1* transposon. MosSci has become widely used for making single-copy integration of transgene. The insertion sites of the transgene are fixed by the Mos sites and can be found in Fig. 2 and Supplemental methods of Frokjaer-Jensen *et al.* (2008). A slight limitation of single-copy insertion is that the expression level of the reporter is not sufficiently high to be observed in visual inspection. Conceivably, manipulating the strength of the promoters or copy numbers of XFP tag could increase the expression levels for such single-copy insertion of transgene reporters.

## B. Labeling Neuronal Proteins and Subcellular Structures

Cytosolic free GFP labeling reveals the overall morphology of the neuron, and allows assessment of neuronal fate and differentiation. Because neurons are polarized cells, an area of intense study in neuronal development is how polarity of neurons is determined, and how a compartment is established. Most proteins can be tagged with XFP while maintaining their physiological functions. Additional modifications can make tagged proteins to be further useful for revealing the dynamics of subcellular compartments. The ease of observing tagged proteins in live worms has made *C. elegans* a prime model organism to investigate the molecular pathways regulating the subcellular structure formation.

- (1) Labeling proteins: fluorescent proteins are usually added to the N- or C-terminus, or occasionally the middle of the protein of interest. The principle for selecting the tagging position is that fusion proteins should retain similar functions to endogenous proteins. Besides regular XFP as tags, photoactivatable or photoconvertible fluorescent proteins can also be used as tags to observe dynamics of proteins (such as: transport, translation, and degradation) (Matsuda *et al.*, 2008; Patterson and Lippincott-Schwartz, 2002; Yampolsky *et al.*, 2008).
- (2) Labeling subcellular structures: Subcellular structures such as the Golgi, the nucleus, and synapses have unique protein and lipid constituents. For example, the hallmark of synapses is that synaptic vesicles are clustered near active zones at presynaptic terminals, and receptors are concentrated at the postsynaptic region (Jin, 2005; Zhai and Bellen, 2004). The initial success of labeling

presynaptic terminals using SNB-1:: GFP (Nonet *et al.*, 1998) has led to the generation of multiple labeling of different synaptic domains using fluorescent fusion proteins. Table II lists many well-characterized transgenic lines that label presynaptic terminals and postsynaptic sites of different neurons.

### C. Complex Strategy for Labeling and Manipulating Unique Neurons

Finding promoters that are activated in multiple neurons is usually straight forward, while it remains challenging to obtain promoters that are uniquely specific for

**Table II**

Transgenic lines for labeling neurons synapses and compartments

Transgenic worms	Plasmids	Labeled neurons	Linkage	Reference
<i>hpls3</i>	<i>Punc-25::GFP::SYD-2</i>	DD VD neurons presynaptic active zones	LGX	(Yeh <i>et al.</i> , 2005)
<i>jsIs1</i>	<i>Psnb-1::snb-1::GFP</i>	P neuronal presynaptic terminals		(Nonet, 1999)
<i>jsIs37</i>	<i>Pmec-7::snb-1::GFP</i>	Touch neurons presynaptic terminals	LGV	(Nonet, 1999)
<i>jsIs42</i>	<i>Punc-4::snb-1::GFP</i>	SAB, DA, I5, VA, AVF VC neurons presynaptic terminals		(Nonet, 1999)
<i>julS1</i>	<i>Punc-25::snb-1::GFP</i>	DD and VD neurons presynaptic terminals	LGIV	(Hallam and Jin, 1998)
<i>kyIs105</i>	<i>Pstr-3::snb-1::GFP</i>	ASI neurons presynaptic terminals	LGV	(Crump <i>et al.</i> , 2001)
<i>kyIs235</i>	<i>Punc-86::snb-1::YFO</i>	HSN neurons synapses	LGV	(Shen and Bargmann, 2003)
<i>kyIs439</i>	<i>Punc-4::lin-10::dsRed</i>			
<i>kyIs442</i>	<i>Podr-3::GFP::unc-2, Podr-3::mCherry::rab-3</i>	AWC neuron presynaptic terminal and active zones		(Saheki and Bargmann, 2009)
<i>kyIs442</i>	<i>Podr-3::GFP::unc-2, Podr-3::mCherry::rab-3</i>	AWC neuron presynaptic terminal and active zones		(Saheki and Bargmann, 2009)
<i>kyIs479</i>	<i>Punc-25::GFP::UNC-2, Punc-25::mCherry::rab-3</i>	DD and VD neuron presynaptic terminals and active zones		(Saheki and Bargmann, 2009)
<i>nuIs24</i>	<i>Pglr-1::GLR-1::GFP</i>	AVG AVJ DVC PVC PVQ RIG RIS RMD RME postsynaptic region		(Rongo <i>et al.</i> , 1998)
<i>nuIs152</i>	<i>Punc-129::snb-1::GFP</i>	Motor neurons presynaptic terminals		(Sieburth <i>et al.</i> , 2005)
<i>nuIs125</i>	<i>Pglr-1::snb-1::GFP</i>	AVG AVJ DVC PVC PVQ RIG RIS RMD RME presynaptic terminals		(Juo and Kaplan, 2004)
<i>nuIs283</i>	<i>Punc-25::RFP-rab-3;Pmyo-3::GFP GABAAR</i>	DD VD neurons synapses		(Vashlishan <i>et al.</i> , 2008)
<i>oxIs22</i>	<i>Pmyo-3::GFP GABAAR</i>	Motor neurons postsynaptic region		(Richmond <i>et al.</i> , 1999)
<i>wyIs45</i>	<i>Pttx3::GFP::rab3</i>	AIY neurons presynaptic terminals		(Colon-Ramos <i>et al.</i> , 2007)
<i>wyIs85</i>	<i>Pitr-1 pB::GFP::rab-3</i>	DA 9 neurons presynaptic terminals		(Klassen and Shen, 2007)
<i>wyIs92</i>	<i>Pmig-13::snb-1::YFP</i>	DA 9 neurons presynaptic terminals		(Klassen and Shen, 2007)
<i>wyIs109</i>	<i>Pmig-3::CFP::rab-3; Pmig-13::snb-1::YFP;Pmig-13::sng-1::mCherry</i>	DA-9 neurons presynaptic terminal synapses		(Poon <i>et al.</i> , 2008)

a single cell. Multiple methods have recently been developed for labeling individual neurons using two or more promoters with overlapping cell-type expression.

- (1) The Cre-loxP recombination system: Cre-LoxP recombination refers to recombining a specific sequence of DNA with the help of an enzyme called Cre recombinase, which is widely used in mammalian systems to generate knockout or transgenic animals. Recently, the Cre-loxP recombination method has also been used successfully in *C. elegans* to label a single neuron called RMG, which has no specific promoter available (Macosko *et al.*, 2009). The general strategy is shown in Fig. 3 B; the key is to find two promoters that exhibit overlapping activities in the neurons of interest. Two transgenic worms are generated: One expresses Cre under promoter A (*Promoter A::Cre*) and the other expresses lox-STOP codon-lox::GFP under promoter B (*Promoter B::Lox-Stop-Lox::GFP*). In the *Promoter B::Lox-STOP-Lox::GFP* transgene, one stop codon, three poly-A sequence and two mRNA cleavage sequences are inserted between the two Lox sites to ensure that GFP can not be expressed in the single-transgenic worms. In the double transgenic worms (*Promoter A::Cre; Promoter B::Lox-STOP-Lox::GFP*), Cre can mediate recombination between the two Lox sites to eliminate the transcriptional stop sequence, and allow GFP expression (Fig. 3 B).
- (2) Reconstituted GFP: Another way to label specific neurons utilizes the properties of reconstituted fluorescent proteins (Zhang *et al.*, 2004). An early version of split GFP uses two fragments (GFP(1–157) and GFP(158–238)) which can restore fluorescence when they are reconstituted by the leucine zipper domain dimerization (Hu *et al.*, 2002). As shown in Fig. 3 C, the N terminus GFP fragment(1–157) is fused to a leucine zipper domain, modified from the leucine zipper domains of mammalian Fos and Jun proteins, in C terminal, and the C terminus GFP fragment (158–238) is fused to a leucine zipper (LZ) domain in N terminal (Zhang *et al.*, 2004). Neither N-GFP(1–157)-LZ nor LZ-C-GFP(158–238) alone can produce fluorescence, but when N-GFP-LZ and LZ-C-GFP are coexpressed in one cell, GFP can be reconstituted through the dimerization of LZ domains (Fig. 3 C) (Zhang *et al.*, 2004). Using this reconstituted fluorescent protein method, specific types of neurons can be labeled by two different promoters driving N-GFP-LZ and LZ-C-GFP (Zhang *et al.*, 2004). This method is also referred as “bimolecular fluorescence complementation (BiFC),” which allows for detecting protein interactions *in vivo* by fusing N-GFP(1–157) and C-GFP(158–238) to different proteins (Hu *et al.*, 2002, 2005).

Recently, a variation of this split-GFP based combinatorial labeling has been developed for visualizing synapses between specific pair of neurons, a method named GRASP for GFP Reconstitution across Synaptic Partners (Feinberg *et al.*, 2008). The GRASP method utilizes the superfolder split-GFP, which has a different split module: GFP-A (aa 1–214) and GFP-B

(aa 215–230) (Pedelacq *et al.*, 2006). Fragment A is fused to the extracellular juxtamembrane position of a surface protein, and expressed in postsynaptic cells. Fragment B is inserted after an artificial signal peptide followed by the coding sequence of a presynaptic terminal member protein PTP-3A, and expressed in the presynaptic cells (Feinberg *et al.*, 2008). If the two membrane proteins are localized to close proximity, such as synaptic junctions, the split GFP fragments A and B are close enough to generate fluorescence, marking the synaptic region (Fig. 3 D). The GRASP method can also be adapted to label any types of cellular junctions.

#### D. Questions and Answers

- How does one specifically label neurons of interest?  
If there is not a specific promoter to label neurons of interest, systematic promoter bashing is necessary to define regulatory sequences for labeling specific neurons.
- How can one be sure about labeling specificity?  
Ectopic or improper expression is an inherent problem in transgene formation. A caution is that one should always confirm the cell-type-specificity of a reporter by analyzing multiple lines. In some cases, if the tagging of protein of interest does not allow verification of cell type, one may consider the strategy to include an operon-reporter, SL2-XFP, in the same construct (Macosko *et al.*, 2009).
- How does one get stable and good expression of transgenes?  
Some extrachromosomal transgenes can be silenced in succession of passage. The cause is largely unknown, though repetitive elements are suspected to be a likely trigger. To overcome this problem, integrated transgenes (multiple copies or single copy) are a good option. One additional suggestion is to include complex DNA, such as random genomic DNA, in making transgenes.
- How does one minimize transgenic labeling artifacts?  
One common concern for using transgenes is that transgenic expression could alter the normal development or function of a cell. To minimize such effects, it is advised to perform functional assays (e.g. rescue the null allele phenotypes) to confirm that the transgenes are functional in a manner similar to the wild-type proteins. One should also test multiple transgenic lines generated from different DNA concentrations. Another possible artifact for using fluorescent protein reporters is from the extremely stable fluorescent proteins. For example gene A is only expressed in the nervous system at early larval stages, but not in adult. However, when using standard GFP as a reporter to observe the expression pattern of A gene, you may see that the reporter of gene A is expressed in adult stages due to the persistent expression of GFP. Use of an unstable GFP as a reporter is an option in this case.

---

---

### III. Other Methods for Examining the Development of Neurons

**Rationale:** Although transgenic fluorescent protein labeling is now universally used to study neuronal development, other methods remain extremely valuable for revealing particular neuronal features and for validating the effects on endogenous proteins and native cells.

#### A. Dye-Filling

*C. elegans* head amphid and tail phasmid sensory neurons can be labeled by lipophilic fluorescent dyes, such as fluorescein isothiocyanate (FITC), 1,1-dioctadecyl-3,3,3,3-tetramethylindocarbocyanine perchlorate (DiI), 3,3'-dilinoleoyloxacarbocyanine perchlorate (DiO), and 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiD). The mechanism of dye uptake seems to involve the exposed endings of sensory cilia and correlates with some aspects of neuronal function. FITC can be used to label the ADF, ASH, ASI, ASJ, ASK, and ADL cells of the amphid sensilla, and the PHA and PHB neurons of the phasmid sensillum. DiI, DiO, and DiD can be used to visualize the ASI, ADL, ASK, AWB, ASH, and ASJ amphid neurons, as well as the PHA and PHB phasmid neurons. Detailed protocols for dye-filling experiments can be found in Shai Shaham's review ("Methods in Cell Biology") of the *wormbook* ([http://www.wormbook.org/chapters/www\\_intro-methodscellbiology/intromethodscellbiology.html](http://www.wormbook.org/chapters/www_intro-methodscellbiology/intromethodscellbiology.html)).

#### B. Immunocytochemistry

Immunostaining using antibodies against various proteins remains invaluable for analyzing subcellular structures and protein localization *in vivo*. Numerous antibodies of neuronal proteins such as UNC-17/ChAT, synaptotagmin, and UNC-10/Rim, mechanosensory microtubules, or neurotransmitters such as serotonin and GABA produce highly specific and consistent patterns (Table III). A very useful monoclonal antibody toolkit for nervous system analysis was made by the Nonet lab (Table III) (Hadwiger *et al.*, 2010). The major hurdles in *C. elegans* antibody staining are antigen fixation and permeating the eggshell of embryos and cuticle of larvae and adults. The best staining comes from practice and trying out a number of fixation procedures. Three commonly used procedures are paraformaldehyde-fixation by Finney and Ruvkun (1990), Bouin fixation by Mike Nonet (Nonet *et al.*, 1997), and freeze-crack by Janet Duerr (Duerr *et al.*, 1999). For detailed protocols on antibody staining, see the *wormbook* chapter "Immunohistochemistry" by Janet S. Duerr ([http://wormbook.org/chapters/www\\_immunohistochemistry/immunohistochemistry.html](http://wormbook.org/chapters/www_immunohistochemistry/immunohistochemistry.html)).

**Table III**Antibody toolkit for *C. elegans*

Markers for	Antibodies against	Reference
Synaptic vesicles	Synaptobrevin (SNB-1)	(Hadwiger <i>et al.</i> , 2010)
Synaptic active zones	Rim (UNC-10)	(Hadwiger <i>et al.</i> , 2010)
Synaptic active zones	SYD-2	(Zhen and Jin, 1999)
Synaptic active zones	UNC-13	(Charlie <i>et al.</i> , 2006)
Synaptic periaactive zones	RPM-1	(Abrams <i>et al.</i> , 2008)
Postsynaptic regions (GABAergic neurons)	GABA receptor (UNC-49)	(Gally and Bessereau, 2003)
Postsynaptic regions (cholinergic neurons)	Nicotinic acetylcholine receptor (UNC-29)	(Gally <i>et al.</i> , 2004)
Postsynaptic regions (cholinergic neurons)	Nicotinic acetylcholine receptor (LEV-10)	(Gally <i>et al.</i> , 2004)
Recycling endosome	EHD1 (RME-1)	(Hadwiger <i>et al.</i> , 2010)
Endoplasmic reticulum	The cytochrome P450 (CYP-33E1)	(Hadwiger <i>et al.</i> , 2010)
Golgi	Beta-1,3-glucuronyltransferase (SQV-8)	(Hadwiger <i>et al.</i> , 2010)
Mitochondria	Chaperonin (HSP-60)	(Hadwiger <i>et al.</i> , 2010)
Lysosomes	LAMP (LMP-1)	(Hadwiger <i>et al.</i> , 2010)
Clathrin-mediated endocytosis sites markers	Dynamin (DYN-1), the alpha-subunit of the adaptor complex 2 (APA-2)	(Hadwiger <i>et al.</i> , 2010)
Inhibitor transmitter (GABA)	GABA	(McIntire <i>et al.</i> , 1993)
Excitatory transmitter (5-HT)	Serotonin (5-HT)	(Weinshenker <i>et al.</i> , 1995)
Touch neurons	Acetylated alpha-tubulin	(Siddiqui <i>et al.</i> , 1989)
GABAergic neurons	Beta-tubulin (one isoform)	(Siddiqui <i>et al.</i> , 1989)
Cholinergic neurons	Choline acetyltransferase (ChAT)/ UNC-17	(Duerr <i>et al.</i> , 2008)
Dopaminergic neurons	Dopamine transporter (DAT-1)	(McDonald <i>et al.</i> , 2007)

### C. Electron Microscopy (EM)

Serial-section EM has been used to reconstruct the cellular architecture of the *C. elegans* nervous system. The brilliant work of John White using serial-section EM had provided us with the structure and connections of the worm nervous systems. Detailed methods about EM can be found in the chapter of this book by David Hall.

## IV. Functional Dissection of Signaling Pathways Controlling Neuronal Development

**Rationale:** To study the cellular processes and dissect the molecular signaling pathways underlying neuronal development, the first step is to characterize the normal developmental features and the second step is to identify and characterize mutants that affect the phenotypes. Further molecular and genetic manipulations will result in a comprehensive understanding of the signaling mode and pathways of the gene of interests. Below, we outline the general uses of transgenic labeling approaches to analyze of neuronal development.

## A. Phenotypic Characterizations

To address whether and how mutations in a gene alter the development of a neuron, the first step is to introduce a reporter into homozygous mutants through standard genetic crosses. The next step is to compare the expression patterns of the reporter in mutant versus wild-type animals. A rule of thumb in this kind of analysis is to maintain identical culture conditions and blinding genotypes when possible. If defects are observed, it is best to test a second transgenic marker or use other methods to confirm the phenotype. For example, if one wants to observe whether a gene affects axon guidance, one can cross a panel of neuronal morphology markers into mutants (Table I) to generate homozygous mutants carrying that marker. If abnormalities are observed, one also needs to confirm that the phenotypes are due to mutation of the gene by transgenic rescue experiments. Further studies include expressing the gene under specific neuronal promoters to determine whether the gene functions cell-autonomously, as well as dissecting the domain requirement of the protein.

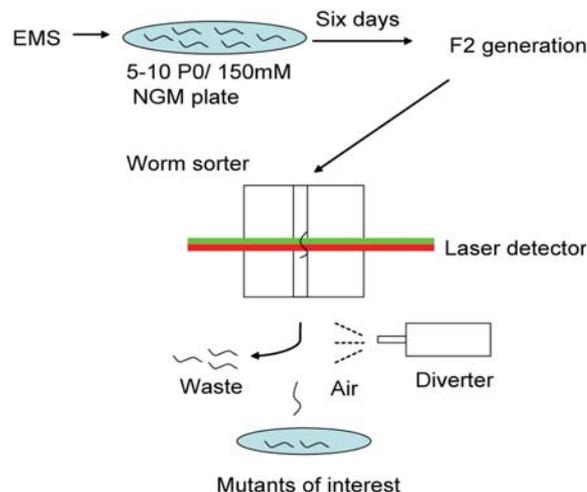
## B. Identifying Mutants in Forward Genetic Screens

Studying of the nervous system in *C. elegans* is facilitated by the ease of carrying out forward genetic screens. Such screens are the most effective and valuable way of discovering new genes and pathways. Before beginning, it is advised to read the article for general guidelines on designing a screen by [Jorgensen and Mango \(2002\)](#). A forward genetic screen combines standard mutagenesis in a range of mutant selection schemes as summarized below.

- (1) Mutagenesis: Two major types of mutagens are chemical mutagens including (1) ethyl methanesulfonate (EMS); (2) nitrosoguanidine (NTG); (3) diethyl sulfate; (4) *N*-Nitroso-*N*-ethylurea; (5) formaldehyde; (6) acetaldehyde; (7) diepoxycyclohexane; (8) diepoxybutane, and radiation mutagens including X-rays,  $\gamma$ -rays, UV light, and ionizing particles ([Anderson, 1995](#)). Among all these mutagens, EMS is the most potent and widely used. Because over 90% of mutations are G/C to A/T transitions, EMS is an excellent mutagen for generating nonsense or missense amino acid coding alternates. Another commonly used mutagen is UV irradiation, which frequently generates a large proportion of gene rearrangements (deletions or insertions) and most of these mutants are likely null alleles.
- (2) Genetic screen: The mostly commonly used procedure is manual visual inspection. Automated screening using worm sorters has also become feasible.
  - (a) Manual screen: For most screens, researchers prefer nonclonal F2 genetic screens. Briefly, mutagenized P0 worms are plated one worm per plate in seeded large plates and cultured for six days. The F2 generation is examined for phenotypes. Such a scheme allows for screening a large number of mutagenized genomes in a relatively short time, but lethal or sterile mutants are easily missed. In a standard clonal screen, after mutagenesis 20 to 30 late

L4 worms are selected as P0, and in the next or F1 generation, individual animals are plated one worm per plate. Three days later, one needs to go through all the F1 plates to find F2 worms with phenotypes of interest and keep them for further characterization. F2 worms from different F1 plates are likely independent mutants. The advantage of clonal screens is that lethal or sterile mutants can be maintained as heterozygotes and recovered from the screen. However, clonal screens can be time-consuming and labor intensive.

- (b) Automated screen: various machines that automatically sort animals based on transgene expression have helped reduce labor and are beginning to be widely used (Fig. 4) (Doitsidou *et al.*, 2008). A common type of automatic screen is shown below:
- (i) Equipment: A worm sorter is a modified flow cytometry instrument called the COPAS™ Biosort from Union Biometrica, which is adjusted to analyze and sort small living organisms on the basis of their optical density, size, and fluorescence (Pulak, 2006).
  - (ii) Strain: It is advisable to have two transgenic markers in the starting strain. For example, one labels the neurons of interest using GFP and the other labels other cells as an internal control using RFP. The doubly labeled transgenic animals are necessary for decreasing false positives. Following standard mutagenesis, 5 to 10 P0 worms are plated in a large plate (150 mm) and cultured for six days until mixed F2/F3 generations



**Fig. 4** Illustration of an automated screen strategy. Mixed F2 worms are sorted by a worm sorter basing on the fluorescence intensity of neurons of interest. This image is based on Doitsidou *et al.* (2008). (For color version of this figure, the reader is referred to the web version of this book.)

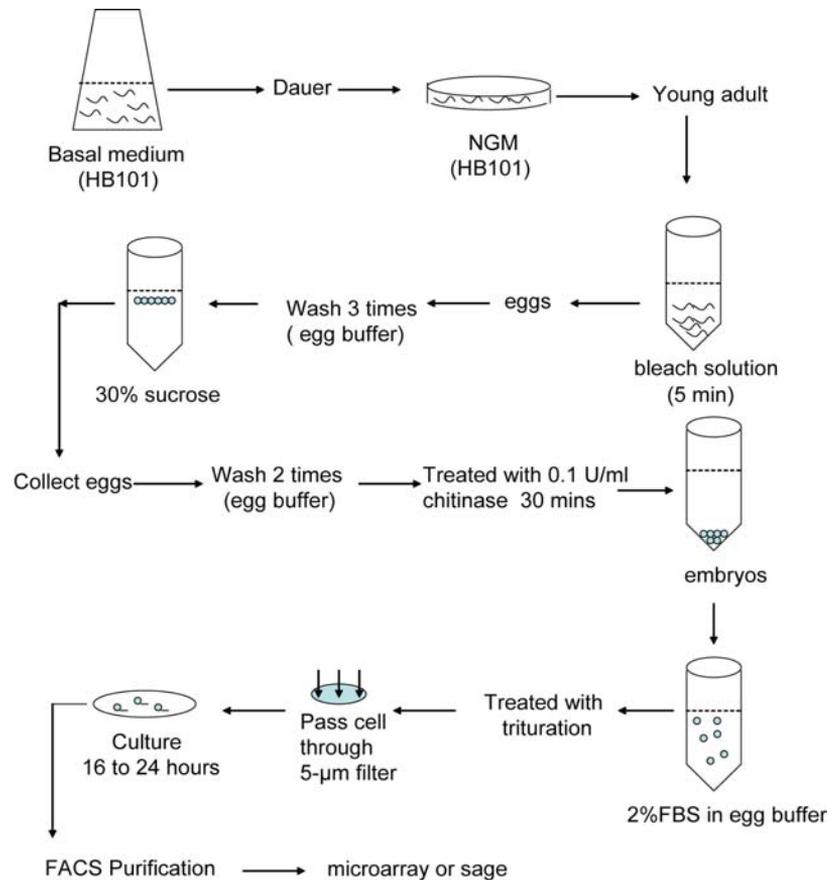
are obtained. The animals are then collected from the P0 plates; one half is passed through the worm sorter to isolate mutants, and the other half is kept for recovery of possible lethal/sterile mutants.

- (iii) Sorting: The sorter is gated to exclude eggs, small L1, and debris using gating parameters time of flight versus extinction. Worms are then sorted based on fluorescence parameters (green fluorescence intensity versus Red fluorescence intensity [GI/RI]). Any mutants that affect GI/RI ratio are kept for further analysis.
  - (iv) Confirmation: For the sorted positive candidates, visual inspection under compound microscopes is used to confirm the phenotypes.
- (3) Analysis of mutants: the phenotypes of mutants should be confirmed for two generations to make sure they are true-breeding. Further analysis of the mutants follows standard guideline, including backcrossing and genetic mapping.

### C. Gene Expression Profiling of Neurons

GFP labeling at transgenic *C. elegans* lines is now widely used to isolate subgroups of neurons for transcriptome analysis. This is largely due to the maturation of the procedure for *in vitro* culturing of dissociated neurons (Strange *et al.*, 2007). Four steps for isolating GFP positive neurons are illustrated in Fig. 5 and described below.

- (1) *C. elegans* culture: Multiple methods can be used to achieve large-scale culture. One common method is a combination of dauer liquid culture and growth on large NGM plates (15 cm). In brief, worms are cultured in 500 mL s-Basal medium (100 mM NaCl, 50 mM KH<sub>2</sub>PO<sub>4</sub>.KOH [pH 6.0], 10 mg/mL cholesterol) with 35 g/mL streptomycin and fed with 10 mL streptomycin-resistant HB101 until they become dauer. These dauer animals are plated on 2–4 mL HB101-seeded 15-cm NGM plates and grown until they reach young gravid adults.
- (2) Dissection and cell culture: Synchronized adult hermaphrodites are lysed for 5 min using bleach solution (0.5 M NaOH and 1% NaOCl). Eggs are then collected and washed three times in sterile egg buffer (118 mM NaCl, 48 mM KCl, 2 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, and 25 mM HEPES [pH 7.3]). Eggs are floated on 30% sucrose by centrifugation to remove adult carcasses. The egg layer is collected and washed twice with the egg buffer and eggs are treated with 0.1 U/mL chitinase (in egg buffer) for 30 min to destroy eggshells. Embryos are pelleted to remove chitinase, and the pellet is resuspended in 2% fetal bovine serum (FBS) in the egg buffer. Embryos are dissociated by trituration. Intact embryos, larvae, and clumps are removed by passing the cell suspension through a 5 μm filter. Filtered cells are plated onto poly-D-lysine-coated cell culture plates in Leibovitz's L-15 medium containing 10% FBS and antibiotics. 16–18 hours later, cells are collected by gentle wash and trituration, centrifuged for 5



**Fig. 5** A flowchart for isolating GFP positive neurons in gene profiling analysis. (For color version of this figure, the reader is referred to the web version of this book.)

min at 900g and redissolved in the egg buffer containing 2% FBS to  $1-2 \times 10^7$  cells/mL concentration.

- (3) FACS purification: GFP-positive cells are purified by flow cytometry. Dead cells and debris are eliminated by a combination of light scatter and propidium iodide negative gates. Cell sorting region for GFP-positive cells is defined by green (530/30 nm) versus orange (575/22 nm) fluorescence and light scatter gate.
- (4) Microarray or SAGE analysis: After getting purified GFP-labeled neurons, RNA isolation, microarray hybridization, and analysis will be applied to get the expression profile of specific subgroups of neurons.

#### D. Laser Surgery of Cells and Processes

Genetically targeted cell disruption and laser-ablation have been used to determine the function of neurons. To genetically eliminate an identified neuron, one can express apoptotic genes such as *ced-3* and *ced-4*, “degenerin family” genes such as *mec-4*, *mec-10*, *deg-1*, and *unc-8*, or toxins such as ricin A and diphtheria toxin-A under neuronal promoters (Harbinder *et al.*, 1997; Shaham and Horvitz, 1996). The caveats of these genetic methods are lack specific promoters to express cell-toxic proteins, disrupting neuronal development and affecting surrounding cells. In these cases, physically killing cells by laser has its advantages. Laser-ablation was traditionally done by first identifying the cell under Nomarski microscope, and then followed by laser operation. Transgenic labeling has greatly simplified the cell identification process. More recently, laser has been successfully used to sever axons or dendrites. We summarize the two types of laser ablation below.

- (1) Cell ablation: To make a traditional laser-ablation system, one needs a laser source, a microscope, and optics to direct the beam of the laser into the microscope objective. The fluorescent compound microscope used to find targeting neurons should have an objective with a numerical aperture of at least 1.25 to focus the laser beam. The laser pulse energy should be bigger than 5 mJ in order to kill cells. Optics is used to shape the beam to enter the specimen from a full range of angles to a focus at a point on the image plane. The traditional cell-ablation system has been widely used in the *C. elegans* field even before the discovery of GFP; the detailed information of this system can be found in Bargmann and Avery’s review paper “Laser killing of cells in *Caenorhabditis elegans*” (Bargmann and Avery, 1995).
- (2) Axon severing: A new application of laser ablation is to severing axons or nerve processes (Wu *et al.*, 2007; Yanik *et al.*, 2004). Two types of lasers can be used to perform axotomy.
  - (a) Femtosecond lasers: A KMLabs MTS Ti-Sapphire oscillator (Kapteyn-Murnane Laboratories, Boulder, CO) pumped by a Verdi V5 (Coherent, Santa Clara, CA) or a KMLabs Cascade laser that can be operated in mode-locked continuous wave (80 MHz) or cavity-dumped (1–100 kHz) modes can be used to generate laser. The pulse energy can be attenuated using neutral density filters and controlled pulse delivery with an electromechanical shutter (Uniblitz VS14 with VMM-T1 controller; Vincent Associates, Rochester, NY).
  - (b) Conventional laser: Axotomy can also be performed using a Photonics Micropoint VSL pulsed UV laser (Photonics Instruments, St. Charles, IL). The laser beam is delivered to an imaging microscope equipped for simultaneous laser and GFP illumination via a Photonics Instruments adaptor, and surgery can be performed by using a Plan Neofluar  $\times 100$ /N.A. 1.3 objectives.

- (c) Axotomy procedure: At this point, axotomy is usually performed on anesthetized animals. The common anesthetics are levamisole (6 mM) (Ou *et al.*, 2010), tetramisole (0.05%) (Brenner, 1974), 1-phenoxy-2-propanol (1%) (Wu *et al.*, 2007), muscimol (10 mM) (Hammarlund *et al.*, 2009), or simply embedding in high-percentage agar pad (10%). The key is to make sure that the anesthetics allow rapid immobilization and recovery of the operated animals, and do not interfere with fluorescent labeling. The regeneration rate and pattern of axons can be measured by confocal imaging at different time points of postaxotomy. Advanced methods using microfluidics to immobilize animals either by cooling or pressure have also been reported for axotomy on nonanesthetized animals (Guo *et al.*, 2008), although such methods are not yet widely used because of technical complications.

## E. Questions and Answers

- Are there common methods for quantifying neuronal defects?  
Most imaging quantification methods for fluorescent light microscopy can be used to record expression patterns and levels of fluorescent reporters. For fluorescence intensity measurement, it is important to maintain identical image-capture conditions, such as laser power, exposure time, and region of interest. Several published studies have also devised home-built software programs for comparing a large number of samples (Burbea *et al.*, 2002; Hung *et al.*, 2007).
- How does one know if a genetic screen has worked?  
A gold standard to determining whether a screen has worked is the frequency of mutants with behavioral or morphological defects. In the pioneer screen done by Sydney Brenner, 69 such mutants (uncoordinated, long, dumpy, small, roller, and blistered mutants) were found in 318 independent F1 plates (Brenner, 1974). This indicates that in a successful screen, at least 20 these mutants can be found in every 100 F1 plates. Also the percentage of the lethal or sterile mutants can be used to determine how well a screen has worked. Usually, nearly 30% of F1 plates will have lethal or sterile mutants.
- How can one reduce false positives in an automated screen?  
To minimize false positives in an automatic screen, one needs to set up a good internal control. The best internal control will be another fluorescent protein, which is usually RFP if the reporter is GFP, expressed in the adjacent tissue of the neuron of interest. Also one needs to select a line that has good expression level for both the reporter internal control. To eliminate false positives, one always needs to confirm the phenotypes by using compound microscopes before doing further experiments.
- What should one do when not getting enough cultured neurons?

The process of *C. elegans* cell culture includes three major steps: egg isolation, cell dissociation, and cell seeding. The reasons for low cell yield are low egg yield, low yield of dissociated cells, and cells failing to attach to growth substrate. To get enough eggs for cell culture, one needs to make sure that worms are gravid before lysis, and that worms are completely digested during lysis. Using glass pipettes during worm and egg transferring can reduce loss. In the cell-dissociation process, one needs to make sure that eggshells are completely digested by chitinase and that all debris is removed before filtering. In the cell-seeding step, one needs to make sure that all coverslips are thoroughly acid-washed and rinsed, and that growth substrate has been coated properly with lectin.

- How can one increase the survival rate in laser surgery experiments?

One major reason for worm death during laser surgery is from keeping worms in anesthetics for too long. To avoid this, one should minimize surgery time (less than 15 min). Because the tolerance to anesthetics varies for different mutants of adjusting anesthetic concentration is also helpful for increasing survival rate. In the axotomy experiments, neurons could be killed if cut sites are too close to the cell bodies. Usually one should sever axons more than 30  $\mu\text{m}$  away from cell bodies.

---

---

---

## V. Examples of Neuronal Development

Since the start of deciphering the nervous system in *C. elegans*, understanding how it is developed has been a major interest for generations of *C. elegans* researchers. The findings from *C. elegans* have made enormous contributions to the discovery of fundamentally conserved mechanisms. Here, we briefly mention some key studies in neuronal development, in the context of particular use of the techniques described above.

### A. Neuronal Fate Specification and Differentiation

*C. elegans* neurons are generally classified by morphology, lineage, and position. The combination of single-cell labeling, genetic manipulation and molecular profiling have been a standard approach to dissecting the genes controlling and executing the specification of neuron type. Such studies have given a rich understanding of fate specification and differentiation for *C. elegans* neurons. Recent successes in combining automated screens with whole-genomic sequencing will surely greatly speed up the progress of our knowledge of cell-type determinants. For example, a group of genes involved in dopamine neuron fate determination were identified by automated screen using *dat-1* promoter driven GFP transgenic worms (Doitsidou *et al.*, 2008).

## B. Axon Guidance and Dendrite Morphogenesis

Following fate determination, neurons will grow two types of processes: dendrites, which collect information from upstream cells, and axons, which pass information to downstream targets. Axon guidance and dendrite morphogenesis involve complex interactions between intrinsic pathways and extrinsic cues. In their pioneering work, Hedgecock *et al.* identified the first group of genes that are required for axon growth and guidance using fluorescein dye-filling technique (Hedgecock *et al.*, 1985). With the use of fluorescent protein labeling which allows for clear visualization of axon paths, the roles of many genes essential to axon guidance including the conserved UNC-6/Netrin and the SLT-1/slit signaling pathways have been deeply investigated. One such example of axon guidance studies is from the study of mechanosensory neuron AVM. The guidance of AVM axons is coordinated by both attraction to a ventral netrin signal (*via* the UNC-40/DCC receptor) and repulsion from a dorsal Slit signal (*via* the SAX-3/Robo receptor) (Culotti, 1994; Hao *et al.*, 2001).

It is generally perceived that most *C. elegans* neurons' dendrites have relatively simple morphology, based on images of EM reconstruction and static antibody staining. One great advantage of using transgenic XFP labeling of neurons is the increasing realization that numerous *C. elegans* neurons exhibit complex morphology and that the axon and dendritic arbors display dynamic features. For example, cell-type labeling revealed that the PVD neuron elaborates complex dendrite tree. A recent study has discovered that a cell fusion gene *eff-1* is required for the dendritic tree formation of PVD neurons (Oren-Suissa *et al.*, 2010). Additionally, transgenic double-labeling of neurons with their neighboring cells has also begun to reveal intercellular mechanisms in the development of neurons. For example, through a genetic screen using cell-type specific XFP labeling for amphid neurons and glial cells, it was shown that formation of amphid dendrites utilizes a “retrograde extension” mode (Heiman and Shaham, 2009).

## C. Synapse Formation and Specificity

Just like axon guidance can be clearly studied at single process resolution, single synapses can be studied *in vivo* by expressing synaptic fluorescent proteins (Nonet, 1999). Using XFP labeling synapses in visual genetic screens, multiple signaling pathways have been identified to regulate formation of pre-synaptic terminals and postsynaptic structures (Jin, 2005). Further studies in other organisms have shown that most of these synapse-formation molecules are functionally conserved from *C. elegans* to mammals. For example, the conserved giant PHR protein, RPM-1, was discovered by its effects on synapse organization. Loss of function of *rpm-1* reduces synapse number and alters synapse morphology (Schaefer *et al.*, 2000; Zhen *et al.*, 2000). Further mechanistic dissection revealed that RPM-1 acts as a ubiquitin E3 ligase to control activation of the conserved DLK-1 MAP kinase signaling pathway (Nakata *et al.*,

2005; Schaefer *et al.*, 2000; Yan *et al.*, 2009; Zhen *et al.*, 2000). Such regulatory mechanisms were later shown to be conserved from *Drosophila* to mammals (Jin and Garner, 2008).

Another highlight of synapse formation research is the increasing findings on molecular control of synapse specificity. In *C. elegans*, most neurons form *en passant* synapses in restricted regions of axons. An interesting question is how neurons know where to form synapses. Using single-cell synapse labeling, new cell surface molecules have been identified to control specific synapse position (Shen and Bargmann, 2003). Additionally, several reports have revealed that classical axon guidance clues and their receptors, such as WNT/Frizzles, UNC-6/UNC-5, and UNC-6/UNC-40, function in cell-type specific synapse pattern formation (Colon-Ramos *et al.*, 2007; Klassen and Shen, 2007; Poon *et al.*, 2008). Moreover, the interactions between cognate receptor and ligand at synapses differ from those in axon guidance.

#### D. Maintenance of Neuronal Architecture

*C. elegans* undergoes four larval stages with adapted behaviors. Recent studies show that there is tremendous plasticity in the maintenance of cell-body and process position. For example, the L1/ neuroglial-related transmembrane protein SAX-7/LAD-1 is required for maintenance of cell-body and axon position during development (Sasakura *et al.*, 2005; Wang *et al.*, 2005). In *sax-7* loss of function mutants, most axons and processes are normally positioned in newly hatched larva and early larva stages and become misplaced in later larva stages. SAX-7 works both in neurons and the epidermis to maintain neuronal position. Another example is the FGF receptor EGL-15, which is expressed in the epidermis adjacent to neuron but not in neurons themselves (Bulow *et al.*, 2004). These studies show that maintenance of neuron position in different developmental stages depends on cell-cell or cell-ECM (extra cellular matrix) adhesion. In addition to these adhesion molecules, the environmental insults are also important for maintaining of neuronal architecture. For example, hypoxia can upregulate the Eph receptor VAB-1 through HIF-1 (hypoxia-inducible factor 1) to affect axon guidance and neuronal migration (Pocock and Hobert, 2008).

---

---

---

## VI. Outlook

In summary, the defined anatomy of *C. elegans* and genetic manipulations have made *C. elegans* a favorite model organism for examining the development and function of the nervous system. The invention of fluorescent proteins for labeling individual neurons or organelles has revolutionized the way to analyze development of neuronal structures and connectivity. New methods such as machine-based automated screening, neuronal-specific RNAi knockdown, and whole genome

sequencing are making an unprecedented impact in the identification of genes and functional pathways. Research in *C. elegans* will undoubtedly continue to lead the discovery of conserved principles controlling neuronal development.

## References

- Abrams, B., Grill, B., Huang, X., and Jin, Y. (2008). Cellular and molecular determinants targeting the *Caenorhabditis elegans* PHR protein RPM-1 to perisynaptic regions. *Dev. Dyn.* **237**, 630–639.
- Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y., and Hobert, O. (2001). A regulatory cascade of three homeobox genes, *ceh-10*, *ttx-3* and *ceh-23*, controls cell fate specification of a defined interneuron class in *C. elegans*. *Development* **128**, 1951–1969.
- Anderson, P. (1995). Mutagenesis. *Methods Cell Biol.* **48**, 31–58.
- Aurelio, O., Boulin, T., and Hobert, O. (2003). Identification of spatial and temporal cues that regulate postembryonic expression of axon maintenance factors in the *C. elegans* ventral nerve cord. *Development* **130**, 599–610.
- Bacaj, T., Tevlin, M., Lu, Y., and Shaham, S. (2008). Glia are essential for sensory organ function in *C. elegans*. *Science* **322**, 744–747.
- Bargmann, C. I., and Avery, L. (1995). Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol.* **48**, 225–250.
- Bauer Huang, S. L., Saheki, Y., VanHoven, M. K., Torayama, I., Ishihara, T., Katsura, I., van der Linden, A., Sengupta, P., and Bargmann, C. I. (2007). Left–right olfactory asymmetry results from antagonistic functions of voltage-activated calcium channels and the Raw repeat protein OLRN-1 in *C. elegans*. *Neural Dev.* **2**, 24.
- Bessereau, J. L., Wright, A., Williams, D. C., Schuske, K., Davis, M. W., and Jorgensen, E. M. (2001). Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74.
- Boulin, T., Pocock, R., and Hobert, O. (2006). A novel Eph receptor-interacting IgSF protein provides *C. elegans* motoneurons with midline guidepost function. *Curr. Biol.* **16**, 1871–1883.
- Bounoutas, A., Kratz, J., Emtage, L., Ma, C., Nguyen, K. C., and Chalfie, M. (2010). Microtubule depolymerization in *Caenorhabditis elegans* touch receptor neurons reduces gene expression through a p38 MAPK pathway. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3982–3987.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94.
- Bulow, H. E., Boulin, T., and Hobert, O. (2004). Differential functions of the *C. elegans* FGF receptor in axon outgrowth and maintenance of axon position. *Neuron* **42**, 367–374.
- Burbea, M., Dreier, L., Dittman, J. S., Grunwald, M. E., and Kaplan, J. M. (2002). Ubiquitin and AP180 regulate the abundance of GLR-1 glutamate receptors at postsynaptic elements in *C. elegans*. *Neuron* **35**, 107–120.
- Celniker, S. E., Dillon, L. A., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., Kellis, M., Lai, E. C., Lieb, J. D., MacAlpine, D. M., Micklem, G., Piano, F., Snyder, M., Stein, L., White, K. P., and Waterston, R. H. (2009). Unlocking the secrets of the genome. *Nature* **459**, 927–930.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Chang, C., Adler, C. E., Krause, M., Clark, S. G., Gertler, F. B., Tessier-Lavigne, M., and Bargmann, C. I. (2006). MIG-10/lamellipodin and AGE-1/PI3K promote axon guidance and outgrowth in response to slit and netrin. *Curr. Biol.* **16**, 854–862.
- Chang, S., Johnston Jr., R. J., and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev.* **17**, 2123–2137.
- Charlie, N. K., Thomure, A. M., Schade, M. A., and Miller, K. G. (2006). The Dunce cAMP phosphodiesterase PDE-4 negatively regulates G alpha(s)-dependent and G alpha(s)-independent cAMP pools in the *Caenorhabditis elegans* synaptic signaling network. *Genetics* **173**, 111–130.
- Chelur, D. S., and Chalfie, M. (2007). Targeted cell killing by reconstituted caspases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2283–2288.

- Christensen, M., and Strange, K. (2001). Developmental regulation of a novel outwardly rectifying mechanosensitive anion channel in *Caenorhabditis elegans*. *J. Biol. Chem.* **276**, 45024–45030.
- Chuang, C. F., Vanhove, M. K., Fetter, R. D., Verselis, V. K., and Bargmann, C. I. (2007). An innexin-dependent cell network establishes left–right neuronal asymmetry in *C. elegans*. *Cell* **129**, 787–799.
- Clark, S. G., and Chiu, C. (2003). *C. elegans* ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. *Development* **130**, 3781–3794.
- Colon-Ramos, D. A., Margeta, M. A., and Shen, K. (2007). Glia promote local synaptogenesis through UNC-6 (netrin) signaling in *C. elegans*. *Science* **318**, 103–106.
- Crump, J. G., Zhen, M., Jin, Y., and Bargmann, C. I. (2001). The SAD-1 kinase regulates presynaptic vesicle clustering and axon termination. *Neuron* **29**, 115–129.
- Culotti, J. G. (1994). Axon guidance mechanisms in *Caenorhabditis elegans*. *Curr. Opin. Genet. Dev.* **4**, 587–595.
- Doitsidou, M., Flames, N., Lee, A. C., Boyanov, A., and Hobert, O. (2008). Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nat. Methods* **5**, 869–872.
- Duerr, J. S., Frisby, D. L., Gaskin, J., Duke, A., Asermely, K., Huddleston, D., Eiden, L. E., and Rand, J. B. (1999). The cat-1 gene of *Caenorhabditis elegans* encodes a vesicular monoamine transporter required for specific monoamine-dependent behaviors. *J. Neurosci.* **19**, 72–84.
- Duerr, J. S., Han, H. P., Fields, S. D., and Rand, J. B. (2008). Identification of major classes of cholinergic neurons in the nematode *Caenorhabditis elegans*. *J. Comp. Neurol.* **506**, 398–408.
- Etchberger, J. F., and Hobert, O. (2008). Vector-free DNA constructs improve transgene expression in *C. elegans*. *Nat. Methods* **5**, 3.
- Feinberg, E. H., Vanhove, M. K., Bendesky, A., Wang, G., Fetter, R. D., Shen, K., and Bargmann, C. I. (2008). GFP reconstitution across synaptic partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353–363.
- Finney, M., and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in *C. elegans*. *Cell* **63**, 895–905.
- Fire, A., Harrison, S. W., and Dixon, D. (1990). A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**, 189–198.
- Forrester, W. C., and Garriga, G. (1997). Genes necessary for *C. elegans* cell and growth cone migrations. *Development* **124**, 1831–1843.
- Frant, A. R., Russel, S., and Ruvkun, G. (2005). Functional genomic analysis of *C. elegans* molting. *PLoS Biol.* **3**, e312.
- Frank, C. A., Hawkins, N. C., Guenther, C., Horvitz, H. R., and Garriga, G. (2005). *C. elegans* HAM-1 positions the cleavage plane and regulates apoptosis in asymmetric neuroblast divisions. *Dev. Biol.* **284**, 301–310.
- Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383.
- Gally, C., and Bessereau, J. L. (2003). GABA is dispensable for the formation of junctional GABA receptor clusters in *Caenorhabditis elegans*. *J. Neurosci.* **23**, 2591–2599.
- Gally, C., Eimer, S., Richmond, J. E., and Bessereau, J. L. (2004). A transmembrane protein required for acetylcholine receptor clustering in *Caenorhabditis elegans*. *Nature* **431**, 578–582.
- Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., Yip, K. Y., Robilotto, R., Rechsteiner, A., Ikegami, K., Alves, P., Chateigner, A., Perry, M., Morris, M., Auerbach, R. K., Feng, X., Leng, J., Vielle, A., Niu, W., Rhrissorakrai, K., Agarwal, A., Alexander, R. P., Barber, G., Brdlik, C. M., Brennan, J., Brouillet, J. J., Carr, A., Cheung, M. S., Clawson, H., Contrino, S., Dannenberg, L. O., Dernburg, A. F., Desai, A., Dick, L., Dose, A. C., Du, J., Egelhofer, T., Ercan, S., Euskirchen, G., Ewing, B., Feingold, E. A., Gassmann, R., Good, P. J., Green, P., Gullier, F., Gutwein, M., Guyer, M. S., Habegger, L., Han, T., Henikoff, J. G., Henz, S. R., Hinrichs, A., Holster, H., Hyman, T., Iniguez, A. L., Janette, J., Jensen, M., Kato, M., Kent, W. J., Kephart, E., Khivansara, V., Khurana, E., Kim, J. K., Kolasinska-Zwierz, P., Lai, E. C., Latorre, I., Leahey, A., Lewis, S., Lloyd, P., Lochovsky, L., Lowdon,

- R. F., Lubling, Y., Lyne, R., MacCoss, M., Mackowiak, S. D., Mangone, M., McKay, S., Mecnas, D., Merrihew, G., Miller, D. M., 3rd., Muroyama, A., Murray, J. I., Ooi, S. L., Pham, H., Phippen, T., Preston, E. A., Rajewsky, N., Ratsch, G., Rosenbaum, H., Rozowsky, J., Rutherford, K., Ruzanov, P., Sarov, M., Sasidharan, R., Sboner, A., Scheid, P., Segal, E., Shin, H., Shou, C., and Slack, F. J., *et al.* (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* **330**, 1775–1787.
- Guo, S. X., Bourgeois, F., Chokshi, T., Durr, N. J., Hilliard, M. A., Chronis, N., and Ben-Yakar, A. (2008). Femtosecond laser nanoaxotomy lab-on-a-chip for *in vivo* nerve regeneration studies. *Nat. Methods* **5**, 531–533.
- Hadwiger, G., Dour, S., Arur, S., Fox, P., and Nonet, M. L. (2010). A monoclonal antibody toolkit for *C. elegans*. *PLoS One* **5**, e10161.
- Hallam, S., Singer, E., Waring, D., and Jin, Y. (2000). The *C. elegans* NeuroD homolog *cnd-1* functions in multiple aspects of motor neuron fate specification. *Development* **127**, 4239–4252.
- Hallam, S. J., and Jin, Y. (1998). *lin-14* regulates the timing of synaptic remodelling in *Caenorhabditis elegans*. *Nature* **395**, 78–82.
- Hammarlund, M., Nix, P., Hauth, L., Jorgensen, E. M., and Bastiani, M. (2009). Axon regeneration requires a conserved MAP kinase pathway. *Science* **323**, 802–806.
- Hao, J. C., Yu, T. W., Fujisawa, K., Culotti, J. G., Gengyo-Ando, K., Mitani, S., Moulder, G., Barstead, R., Tessier-Lavigne, M., and Bargmann, C. I. (2001). *C. elegans* slit acts in midline, dorsal–ventral, and anterior–posterior guidance via the SAX-3/Robo receptor. *Neuron* **32**, 25–38.
- Harbinder, S., Tavernarakis, N., Herndon, L. A., Kinnell, M., Xu, S. Q., Fire, A., and Driscoll, M. (1997). Genetically targeted cell disruption in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. U. S. A.* **94**, 13128–13133.
- Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. *Dev. Biol.* **111**, 158–170.
- Heiman, M. G., and Shaham, S. (2009). DEX-1 and DYF-7 establish sensory dendrite length by anchoring dendritic tips during cell migration. *Cell* **137**, 344–355.
- Hobert, O., Moerman, D. G., Clark, K. A., Beckerle, M. C., and Ruvkun, G. (1999). A conserved LIM protein that affects muscular adherens junction integrity and mechanosensory function in *Caenorhabditis elegans*. *J. Cell Biol.* **144**, 45–57.
- Hu, C. D., Chinenov, Y., and Kerppola, T. K. (2002). Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell* **9**, 789–798.
- Hu, C. D., Grinberg, A. V., and Kerppola, T. K. (2005). Visualization of protein interactions in living cells using bimolecular fluorescence complementation (BiFC) analysis. *Curr. Protoc. Protein Sci.* **Chapter 19**, Unit 19 10.
- Huang, X., Cheng, H. J., Tessier-Lavigne, M., and Jin, Y. (2002). MAX-1, a novel PH/MyTH4/FERM domain cytoplasmic protein implicated in netrin-mediated axon repulsion. *Neuron* **34**, 563–576.
- Hung, W., Hwang, C., Po, M. D., and Zhen, M. (2007). Neuronal polarity is regulated by a direct interaction between a scaffolding protein, Neurabin, and a presynaptic SAD-1 kinase in *Caenorhabditis elegans*. *Development* **134**, 237–249.
- Hutter, H. (2003). Extracellular cues and pioneers act together to guide axons in the ventral cord of *C. elegans*. *Development* **130**, 5307–5318.
- Jin, Y. (2005). Synaptogenesis. WormBook, (<http://www.wormbook.org/>) pp. 1–11.
- Jin, Y., and Garner, C. C. (2008). Molecular mechanisms of presynaptic differentiation. *Annu. Rev. Cell Dev. Biol.* **24**, 237–262.
- Johnston, R. J., and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**, 845–849.
- Jorgensen, E. M., and Mango, S. E. (2002). The art and design of genetic screens: *Caenorhabditis elegans*. *Nat. Rev. Genet.* **3**, 356–369.

- Juo, P., and Kaplan, J. M. (2004). The anaphase-promoting complex regulates the abundance of GLR-1 glutamate receptors in the ventral nerve cord of *C. elegans*. *Curr. Biol.* **14**, 2057–2062.
- Klassen, M. P., and Shen, K. (2007). Wnt signaling positions neuromuscular connectivity by inhibiting synapse formation in *C. elegans*. *Cell* **130**, 704–716.
- Knobel, K. M., Davis, W. S., Jorgensen, E. M., and Bastiani, M. J. (2001). UNC-119 suppresses axon branching in *C. elegans*. *Development* **128**, 4079–4092.
- Lanjuin, A., VanHoven, M. K., Bargmann, C. I., Thompson, J. K., and Sengupta, P. (2003). Otx/otd homeobox genes specify distinct sensory neuron identities in *C. elegans*. *Dev. Cell* **5**, 621–633.
- Lesch, B. J., Gehrke, A. R., Bulyk, M. L., and Bargmann, C. I. (2009). Transcriptional regulation and stabilization of left–right neuronal identity in *C. elegans*. *Genes Dev.* **23**, 345–358.
- Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C. C., and Kain, S. R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* **273**, 34970–34975.
- Lim, Y. S., Mallapur, S., Kao, G., Ren, X. C., and Wadsworth, W. G. (1999). Netrin UNC-6 and the regulation of branching and extension of motoneuron axons from the ventral nerve cord of *Caenorhabditis elegans*. *J. Neurosci.* **19**, 7048–7056.
- Lundquist, E. A., Reddien, P. W., Hartwig, E., Horvitz, H. R., and Bargmann, C. I. (2001). Three *C. elegans* Rac proteins and several alternative Rac regulators control axon guidance, cell migration and apoptotic cell phagocytosis. *Development* **128**, 4475–4488.
- Macosko, E. Z., Pokala, N., Feinberg, E. H., Chalasani, S. H., Butcher, R. A., Clardy, J., and Bargmann, C. I. (2009). A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175.
- Mangone, et al. Mangone, M., Manoharan, A. P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S. D., Mis, E., Zegar, C., Gutwein, M. R., Khivansara, V., Attie, O., Chen, K., Salehi-Ashtiani, K., Vidal, M., Harkins, T. T., Bouffard, P., Suzuki, Y., Sugano, S., Kohara, Y., Rajewsky, N., Piano, F., Gunsalus, K. C., and Kim, J. K. (2010). The landscape of *C. elegans* 3'UTRs. *Science* **329**, 432–435.
- Matsuda, T., Miyawaki, A., and Nagai, T. (2008). Direct measurement of protein dynamics inside cells using a rationally designed photoconvertible protein. *Nat. Methods* **5**, 339–345.
- McDonald, P. W., Hardie, S. L., Jensen, T. N., Carvelli, L., Matthies, D. S., and Blakely, R. D. (2007). Vigorous motor activity in *Caenorhabditis elegans* requires efficient clearance of dopamine mediated by synaptic localization of the dopamine transporter DAT-1. *J. Neurosci.* **27**, 14216–14227.
- McIntire, S. L., Jorgensen, E., Kaplan, J., and Horvitz, H. R. (1993). The GABAergic nervous system of *Caenorhabditis elegans*. *Nature* **364**, 337–341.
- McIntire, S. L., Reimer, R. J., Schuske, K., Edwards, R. H., and Jorgensen, E. M. (1997). Identification and characterization of the vesicular GABA transporter. *Nature* **389**, 870–876.
- Mehta, N., Loria, P. M., and Hobert, O. (2004). A genetic screen for neurite outgrowth mutants in *Caenorhabditis elegans* reveals a new function for the F-box ubiquitin ligase component LIN-23. *Genetics* **166**, 1253–1267.
- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: Extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970.
- Mohler, W. A., Simske, J. S., Williams-Masson, E. M., Hardin, J. D., and White, J. G. (1998). Dynamics and ultrastructure of developmental cell fusions in the *Caenorhabditis elegans* hypodermis. *Curr. Biol.* **8**, 1087–1090.
- Much, J. W., Slade, D. J., Klampert, K., Garriga, G., and Wightman, B. (2000). The fax-1 nuclear hormone receptor regulates axon pathfinding and neurotransmitter expression. *Development* **127**, 703–712.
- Nakata, K., Abrams, B., Grill, B., Goncharov, A., Huang, X., Chisholm, A. D., and Jin, Y. (2005). Regulation of a DLK-1 and p38 MAP kinase pathway by the ubiquitin ligase RPM-1 is required for presynaptic development. *Cell* **120**, 407–420.
- Nass, R., Hall, D. H., Miller 3rd, D. M., and Blakely, R. D. (2002). Neurotoxin-induced degeneration of dopamine neurons in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. U. S. A.* **99**, 3264–3269.
- Nonet, M. L. (1999). Visualization of synaptic specializations in live *C. elegans* with synaptic vesicle protein–GFP fusions. *J. Neurosci. Methods* **89**, 33–40.

- Nonet, M. L., Saifee, O., Zhao, H., Rand, J. B., and Wei, L. (1998). Synaptic transmission deficits in *Caenorhabditis elegans* synaptobrevin mutants. *J. Neurosci.* **18**, 70–80.
- Nonet, M. L., Staunton, J. E., Kilgard, M. P., Fergestad, T., Hartwig, E., Horvitz, H. R., Jorgensen, E. M., and Meyer, B. J. (1997). *Caenorhabditis elegans* rab-3 mutant synapses exhibit impaired function and are partially depleted of vesicles. *J. Neurosci.* **17**, 8061–8073.
- Oren-Suissa, M., Hall, D. H., Treinin, M., Shemer, G., and Podbilewicz, B. (2010). The fusogen EFF-1 controls sculpting of mechanosensory dendrites. *Science* **328**, 1285–1288.
- Ortiz, C. O., Etchberger, J. F., Posy, S. L., Frokjaer-Jensen, C., Lockery, S., Honig, B., and Hobert, O. (2006). Searching for neuronal left/right asymmetry: Genomewide analysis of nematode receptor-type guanylyl cyclases. *Genetics* **173**, 131–149.
- Ou, C. Y., Poon, V. Y., Maeder, C. I., Watanabe, S., Lehrman, E. K., Fu, A. K., Park, M., Fu, W. Y., Jorgensen, E. M., Ip, N. Y., and Shen, K. (2010). Two cyclin-dependent kinase pathways are essential for polarized trafficking of presynaptic components. *Cell* **141**, 846–858.
- Palmer, R. E., Inoue, T., Sherwood, D. R., Jiang, L. I., and Sternberg, P. W. (2002). *Caenorhabditis elegans* cog-1 locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. *Dev. Biol.* **252**, 202–213.
- Patterson, G. H., and Lippincott-Schwartz, J. (2002). A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* **297**, 1873–1877.
- Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88.
- Pilon, M., Peng, X. R., Spence, A. M., Plasterk, R. H., and Dosch, H. M. (2000). The diabetes autoantigen ICA69 and its *Caenorhabditis elegans* homologue, ric-19, are conserved regulators of neuroendocrine secretion. *Mol. Biol. Cell* **11**, 3277–3288.
- Pocock, R., and Hobert, O. (2008). Oxygen levels affect axon guidance and neuronal migration in *Caenorhabditis elegans*. *Nat. Neurosci.* **11**, 894–900.
- Poon, V. Y., Klassen, M. P., and Shen, K. (2008). UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites. *Nature* **455**, 669–673.
- Poyurovsky, M. V., Jacq, X., Ma, C., Karni-Schmidt, O., Parker, P. J., Chalfie, M., Manley, J. L., and Prives, C. (2003). Nucleotide binding by the Mdm2 RING domain facilitates Arf-independent Mdm2 nucleolar localization. *Mol. Cell* **12**, 875–887.
- Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226.
- Pujol, N., Torregrossa, P., Ewbank, J. J., and Brunet, J. F. (2000). The homeodomain protein CePHOX2/CEH-17 controls antero-posterior axonal growth in *C. elegans*. *Development* **127**, 3361–3371.
- Pulak, R. (2006). Techniques for analysis, sorting, and dispensing of *C. elegans* on the COPAS flow-sorting system. *Methods Mol. Biol.* **351**, 275–286.
- Richmond, J. E., Davis, W. S., and Jorgensen, E. M. (1999). UNC-13 is required for synaptic vesicle fusion in *C. elegans*. *Nat. Neurosci.* **2**, 959–964.
- Rongo, C., Whitfield, C. W., Rodal, A., Kim, S. K., and Kaplan, J. M. (1998). LIN-10 is a shared component of the polarized protein localization pathways in neurons and epithelia. *Cell* **94**, 751–759.
- Sagasti, A., Hobert, O., Troemel, E. R., Ruvkun, G., and Bargmann, C. I. (1999). Alternative olfactory neuron fates are specified by the LIM homeobox gene *lim-4*. *Genes Dev.* **13**, 1794–1806.
- Saheki, Y., and Bargmann, C. I. (2009). Presynaptic CaV2 calcium channel traffic requires CALF-1 and the alpha(2)delta subunit UNC-36. *Nat. Neurosci.* **12**, 1257–1265.
- Sarafi-Reinach, T. R., Melkman, T., Hobert, O., and Sengupta, P. (2001). The *lin-11* LIM homeobox gene specifies olfactory and chemosensory neuron fates in *C. elegans*. *Development* **128**, 3269–3281.
- Sasakura, H., Inada, H., Kuhara, A., Fusaoka, E., Takemoto, D., Takeuchi, K., and Mori, I. (2005). Maintenance of neuronal positions in organized ganglia by SAX-7, a *Caenorhabditis elegans* homologue of L1. *EMBO J.* **24**, 1477–1488.
- Schaefer, A. M., Hadwiger, G. D., and Nonet, M. L. (2000). *rpm-1*, a conserved neuronal gene that regulates targeting and synaptogenesis in *C. elegans*. *Neuron* **26**, 345–356.

- Shaham, S. (2006). Glia–neuron interactions in the nervous system of *Caenorhabditis elegans*. *Curr. Opin. Neurobiol.* **16**, 522–528.
- Shaham, S., and Horvitz, H. R. (1996). Developing *Caenorhabditis elegans* neurons may contain both cell-death protective and killer activities. *Genes Dev.* **10**, 578–591.
- Shen, K., and Bargmann, C. I. (2003). The immunoglobulin superfamily protein SYG-1 determines the location of specific synapses in *C. elegans*. *Cell* **112**, 619–630.
- Siddiqui, S. S., Aamodt, E., Rastinejad, F., and Culotti, J. (1989). Anti-tubulin monoclonal antibodies that bind to specific neurons in *Caenorhabditis elegans*. *J. Neurosci.* **9**, 2963–2972.
- Sieburth, D., Ch'ng, Q., Dybbs, M., Tavazoie, M., Kennedy, S., Wang, D., Dupuy, D., Rual, J. F., Hill, D. E., Vidal, M., Ruvkun, G., and Kaplan, J. M. (2005). Systematic analysis of genes required for synapse structure and function. *Nature* **436**, 510–517.
- Sieburth, D., Madison, J. M., and Kaplan, J. M. (2007). PKC-1 regulates secretion of neuropeptides. *Nat. Neurosci.* **10**, 49–57.
- Strange, K., Christensen, M., and Morrison, R. (2007). Primary culture of *Caenorhabditis elegans* developing embryo cells for electrophysiological, cell biological and molecular studies. *Nat. Protoc.* **2**, 1003–1012.
- Troemel, E. R., Sagasti, A., and Bargmann, C. I. (1999). Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell* **99**, 387–398.
- Tsalik, E. L., and Hobert, O. (2003). Functional mapping of neurons that control locomotory behavior in *Caenorhabditis elegans*. *J. Neurobiol.* **56**, 178–197.
- Vanhoven, M. K., Bauer Huang, S. L., Albin, S. D., and Bargmann, C. I. (2006). The claudin superfamily protein nsy-4 biases lateral signaling to generate left–right asymmetry in *C. elegans* olfactory neurons. *Neuron* **51**, 291–302.
- Vashlishan, A. B., Madison, J. M., Dybbs, M., Bai, J., Sieburth, D., Ch'ng, Q., Tavazoie, M., and Kaplan, J. M. (2008). An RNAi screen identifies genes that regulate GABA synapses. *Neuron* **58**, 346–361.
- Wang, X., Kweon, J., Larson, S., and Chen, L. (2005). A role for the *C. elegans* L1CAM homologue lad-1/sax-7 in maintaining tissue attachment. *Dev. Biol.* **284**, 273–291.
- Weinshenker, D., Garriga, G., and Thomas, J. H. (1995). Genetic and pharmacological analysis of neurotransmitters controlling egg laying in *C. elegans*. *J. Neurosci.* **15**, 6975–6985.
- White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the Nematode *Caenorhabditis Elegans*. *Phil Trans. R. Soc. London* **314**, 1–340.
- Wu, Z., Ghosh-Roy, A., Yanik, M. F., Zhang, J. Z., Jin, Y., and Chisholm, A. D. (2007). *Caenorhabditis elegans* neuronal regeneration is influenced by life stage, ephrin signaling, and synaptic branching. *Proc. Natl Acad. Sci. U. S. A.* **104**, 15132–15137.
- Yampolsky, I. V., Kislukhin, A. A., Amatov, T. T., Shcherbo, D., Potapov, V. K., Lukyanov, S., and Lukyanov, K. A. (2008). Synthesis and properties of the red chromophore of the green-to-red photoconvertible fluorescent protein Kaede and its analogs. *Bioorg. Chem.* **36**, 96–104.
- Yan, D., Wu, Z., Chisholm, A. D., and Jin, Y. (2009). The DLK-1 kinase promotes mRNA stability and local translation in *C. elegans* synapses and axon regeneration. *Cell* **138**, 1005–1018.
- Yanik, M. F., Cinar, H., Cinar, H. N., Chisholm, A. D., Jin, Y., and Ben-Yakar, A. (2004). Neurosurgery: Functional regeneration after laser axotomy. *Nature* **432**, 822.
- Yeh, E., Kawano, T., Weimer, R. M., Bessereau, J. L., and Zhen, M. (2005). Identification of genes involved in synaptogenesis using a fluorescent active zone marker in *Caenorhabditis elegans*. *J. Neurosci.* **25**, 3833–3841.
- Zhai, R. G., and Bellen, H. J. (2004). The architecture of the active zone in the presynaptic nerve terminal. *Physiology (Bethesda)* **19**, 262–270.
- Zhang, S., Ma, C., and Chalfie, M. (2004). Combinatorial marking of cells and organelles with reconstituted fluorescent proteins. *Cell* **119**, 137–144.
- Zhen, M., Huang, X., Bamber, B., and Jin, Y. (2000). Regulation of presynaptic terminal organization by *C. elegans* RPM-1, a putative guanine nucleotide exchanger with a RING-H2 finger domain. *Neuron* **26**, 331–343.

- Zhen, M., and Jin, Y. (1999). The liprin protein SYD-2 regulates the differentiation of presynaptic termini in *C. elegans*. *Nature* **401**, 371–375.
- Zheng, Y., Brockie, P. J., Mellem, J. E., Madsen, D. M., and Maricq, A. V. (1999). Neuronal control of locomotion in *C. elegans* is modified by a dominant mutation in the GLR-1 ionotropic glutamate receptor. *Neuron* **24**, 347–361.
- Zipkin, I. D., Kindt, R. M., and Kenyon, C. J. (1997). Role of a new Rho family member in cell migration and axon guidance in *C. elegans*. *Cell* **90**, 883–894.