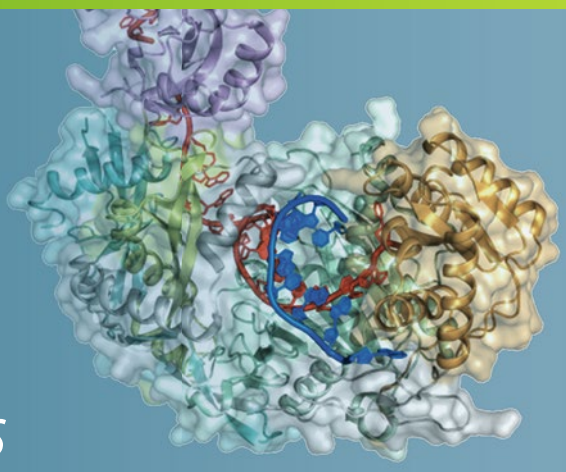
Methods in Molecular Biology 1680

Springer Protocols



Katsutomo Okamura
Kotaro Nakanishi *Editors*

Argonaute Proteins

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Argonaute Proteins

Methods and Protocols

Edited by

Katsutomo Okamura

Temasek Lifesciences Laboratory, National University of Singapore, Singapore

Kotaro Nakanishi

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA

Humana Press

Editors
Katsutomo Okamura
Temasek Lifesciences Laboratory
National University of Singapore
Singapore

Kotaro Nakanishi
Department of Chemistry and Biochemistry
The Ohio State University
Columbus, OH, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-7338-5 ISBN 978-1-4939-7339-2 (eBook)
DOI 10.1007/978-1-4939-7339-2

Library of Congress Control Number: 2017952703

Cover image: This image is the crystal structure of human Argonaute2 in complex with a guide and a target (PDB ID: 4W5O)

Printed on acid-free paper

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

Preface

Since the discovery that Argonautes play essential roles in RNA silencing in a variety of eukaryotes, biochemical properties and biological functions of Argonautes have been extensively studied using an assortment of techniques [1, 3]. Biochemical and structural analyses in the past decade clarified their detailed molecular architectures along with the mechanisms by which Argonautes bind guide and target nucleic acids and cleave targets [4, 5]. Despite the similarity of the overall domain structures, the roles of eukaryotic and prokaryotic Argonautes are quite diverse. Most of them seem to serve mainly as defense systems against foreign nucleic acids, such as viruses/bacteriophages and transposons, while some others are involved in gene regulation [6, 7].

Although the functionality of endogenous small regulatory RNAs was already described in developmental biology studies using *C. elegans* in the 1990s [8, 9], the importance of the underlying molecular mechanism was not recognized until the early 2000s. However, once the deep conservation of the small RNA processing enzymes, Dicers, and the effector proteins, Argonautes, was discovered, the scientific community quickly developed a large body of knowledge regarding biological functions of small regulatory RNA pathways. The development was further accelerated by advances in deep sequencing technologies that allowed massive identification of small RNA species in biological systems. Thanks to recent studies, we now have a seemingly complete catalog of Argonaute-bound small RNAs in a few model organisms as well as humans [10].

However, we are only beginning to understand biological roles for individual small RNAs. Furthermore, latest studies revealed that functions of small RNA-Argonaute complexes have been highly diversified during evolution ranging from DNA/chromatin modifications to mRNA cleavage/destabilization, presumably due to the specialized ability of Argonautes to recruit various downstream effectors [11]. Needless to say, a full understanding of such non-conserved functions of Argonautes is important to elucidate the molecular mechanisms behind biological phenomena in natural contexts. Meanwhile, studies of diverse Argonaute-mediated mechanisms are urged further in terms of the potential of Argonautes for applications as reprogrammable gene regulatory modules. Some of them may become important additions of artificial gene silencing/editing techniques to the current list of various RNAi- and CRISPR-mediated methods.

On the other hand, the vast diversity of Argonaute-mediated mechanisms makes it a difficult task to fully understand the biology of Argonautes in various organisms. The important steps include (1) Identification and expression analysis of guide nucleic acids and their targets. (2) Analysis of biochemical properties of Argonautes. (3) Biological functions of Argonautes. Furthermore, for all these studies, a common question is (4) How do we obtain materials and set up analysis platforms? This volume provides a comprehensive set of protocols that would help to address some of the important questions in these four areas.

1. *Identification and expression analysis of guide nucleic acids and their targets*
   The key step toward a full understanding of Argonaute functions is to comprehensively catalog the guide nucleic acids. This includes efficient cloning/sequencing of small RNAs bound to Argonautes (Chapter 1), precisely quantifying the absolute amounts of

Argonaute-bound small RNAs (Chapter 2), identifying intrinsic sequence preferences of Argonaute proteins (Chapter 3), and quantifying individual small RNA sequences in an isoform-specific manner (Chapter 4). For multicellular organisms, it is important to analyze the expression patterns of small RNA genes. In situ hybridization using whole-mount worm samples provides insights into biological functions of individual small RNA genes (Chapter 5). It is also critical to fully catalog targets of guide RNA-Argonaute complexes in tissues using genome-wide methods such as CLIP (cross-linking and immunoprecipitation: Chapter 6).

2. *Analysis of biochemical properties of Argonautes.*
   Thanks to the successful establishment of protocols to purify individual components of small RNA processing pathways as recombinant proteins, their enzymatic activities and biochemical properties can now be quantitatively measured with unprecedented accuracy. This volume includes protocols for quantitative analysis of substrate processing by Dicer enzymes (Chapter 7), binding kinetics between guide nucleic acids and effector Argonaute complexes in prokaryotes (Chapter 8) and eukaryotes (Chapters 9 and 10). In particular, these protocols offer a wide range of quantification methods from traditional enzymatic kinetics analyses to measurements of binding affinities based on fluorescence polarization and single-molecule measurements.

3. *Biological functions of Argonautes*
   To understand biological roles for Argonautes, it is inevitable to study their functions at cellular or organismal levels. The challenge is that Argonautes may play an extremely wide range of roles in various biological aspects. This volume provides protocols that allow us to study several aspects of Argonaute functions, including regulation of the high-order chromatin structure (Chapter 11), and in vivo miRNA functions in fish embryos (Chapter 12) and mouse embryonic brains (Chapter 13).

4. *Preparation of materials and analysis platforms.*
   To answer the above-mentioned questions, important steps are to prepare reliable materials and streamline their analytical pipelines. A protocol is described for an establishment of embryonic fibroblast cultures from Argonaute2 knockout mice (Chapter 14). Procedures are introduced for epitope tagging of the endogenous Argonaute1 locus in a *Drosophila* cell line using the CRISPR/Cas9 system (Chapter 15). For genome-wide analysis of small RNA sequencing data, a major challenge is to correctly identify the genomic origins of reads that are derived from repetitive sequences, such as tRNA genes. A streamlined bioinformatics resource provides analysis platform for small RNAs derived from tRNA genes and related sequences, some of which are known to bind Argonaute proteins (Chapter 16).

This collection of protocols covers methods to analyze various Argonaute proteins from a wide range of organisms to understand their properties at different levels, from the molecular to the organismal level. The editors hope that this volume will serve as a valuable resource. Lastly, we would like to sincerely thank all the authors who generously contributed their latest protocols to this volume of Methods in Molecular Biology *Argonaute Proteins.*

Singapore ***Katsutomo Okamura***
Columbus, OH, USA ***Kotaro Nakanishi***

References

1. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99(2):123–132
2. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. Science 293(5532):1146–1150
3. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
4. Nakanishi K (2016) Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? Wiley Interdiscip Rev RNA 7(5):637–660. doi:10.1002/wrna.1356
5. Kuhn CD, Joshua-Tor L (2013) Eukaryotic Argonautes come into focus. Trends Biochem Sci 38(5):263–271. doi:10.1016/j.tibs.2013.02.008
6. Makarova KS, Wolf YI, van der Oost J, Koonin EV (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol Direct 4:29. doi:10.1186/1745-6150-4-29
7. Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J (2014) The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol 21(9):743–753. doi:10.1038/nsmb.2879
8. Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell 75(5):843–854
9. Wightman B, Ha I, Ruvkun G (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell 75:855–862
10. Okamura K (2012) Diversity of animal small RNA pathways and their biological utility. Wiley Interdiscip Rev RNA 3(3):351–368. doi:10.1002/wrna.113
11. Burroughs AM, Ando Y, Aravind L (2014) New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. Wiley Interdiscip Rev RNA 5(2):141–181. doi:10.1002/wrna.1210

Contents

Contributors

Panagiotis Alexiou • *Division of Neuropathology, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*

Yoshiki Andachi • *Genetic Strains Research Center, National Institute of Genetics, Research Organization of Information and Systems, Mishima, Japan; Department of Genetics, SOKENDAI (The Graduate University for Advanced Studies), Mishima, Japan*

Daniel Cifuentes • *Department of Biochemistry, Boston University School of Medicine, Boston, MA, USA*

Daniel M. Dayeh • *Ohio State Chemistry Program, The Ohio State University, Columbus, OH, USA; Center for RNA Biology, Columbus, OH, USA*

Ryuya Fukunaga • *Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD, USA*

Pooja Gangras • *Department of Molecular Genetics, The Ohio State University, Columbus, OH, USA; Center for RNA Biology, The Ohio State University, Columbus, OH, USA*

Sanjay Ghosh • *MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK; Department of Biochemistry, University of Cambridge, Cambridge, UK*

Eling Goh • *Temasek Life Sciences Laboratory, National University of Singapore, Singapore; School of Biological Sciences, Nanyang Technological University, Singapore*

Shozo Honda • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*

Shintaro Iwasaki • *Institute of Molecular and Cellular Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; RNA Systems Biochemistry Laboratory, RIKEN, Wako, Saitama, Japan*

Yuka W. Iwasaki • *Department of Molecular Biology, Keio University School of Medicine, Tokyo, Japan*

Yohei Kirino • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*

Yuji Kohara • *Genetic Strains Research Center, National Institute of Genetics, Research Organization of Information and Systems, Mishima, Japan; Department of Genetics, SOKENDAI (The Graduate University for Advanced Studies), Mishima, Japan*

Dmitry A. Kretov • *Department of Biochemistry, Boston University School of Medicine, Boston, MA, USA*

Junan Li • *Division of Pharmacy Practice and Science, College of Pharmacy, The Ohio State University, Columbus, OH, USA*

Susan E. Liao • *Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD, USA*

Ji-Long Liu • *MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK; School of Life Science and Technology, ShanghaiTech University, Shanghai, China*
Phillipe Loher • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*
Justin M. Mabin • *Department of Molecular Genetics, The Ohio State University, Columbus, OH, USA; Center for RNA Biology, Columbus, OH, USA*
Manolis Maragkakis • *Division of Neuropathology, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
Tomohiro Miyoshi • *Center for Transdisciplinary Research, Niigata University, Nishi-ku, Niigata, Japan*
Zissimos Mourelatos • *Division of Neuropathology, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
Kensaku Murano • *Department of Molecular Biology, Keio University School of Medicine, Tokyo, Japan*
Takahisa Nakamura • *Divisions of Endocrinology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA*
Kotaro Nakanishi • *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA; Ohio State Chemistry Program, The Ohio State University, Columbus, OH, USA; Center for RNA Biology, The Ohio State University, Columbus, OH, USA*
Katsutomo Okamura • *Temasek Life Sciences Laboratory, National University of Singapore, Singapore; School of Biological Sciences, Nanyang Technological University, Singapore*
Hong-Duc Phan • *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA; Ohio State Biochemistry Program, Columbus, OH, USA*
Ming Poi • *Division of Pharmacy Practice and Science, College of Pharmacy, The Ohio State University, Columbus, OH, USA*
Isidore Rigoutsos • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*
Hiroshi M. Sasaki • *Institute for Molecular and Cellular Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA*
Joonbae Seo • *Divisions of Endocrinology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA*
Andrew M. Shafik • *Department of Biochemistry, Boston University School of Medicine, Boston, MA, USA*
Megumi Shigematsu • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*
Guramrit Singh • *Department of Molecular Genetics, The Ohio State University, Columbus, OH, USA; Center for RNA Biology, Columbus, OH, USA*
Haruhiko Siomi • *Department of Molecular Biology, Keio University School of Medicine, Tokyo, Japan*

Hisashi Tadakuma • *Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; Institute for Protein Research, Osaka University, Suita-shi, Osaka, Japan*

Aristeidis G. Telonis • *Computational Medicine Center, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*

Yukihide Tomari • *Institute for Molecular and Cellular Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan*

Anastassios Vourekas • *Division of Neuropathology, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*

Li Zeng • *Neural Stem Cell Research Lab, Research Department, National Neuroscience Institute, Singapore; Neuroscience and Behavioral Disorders Program, DUKE-NUS Graduate Medical School, Singapore*

Xiaoxia Zeng • *Neural Stem Cell Research Lab, Research Department, National Neuroscience Institute, Singapore*

Cai Zhang • *Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China*

Wei Zhang • *Neural Stem Cell Research Lab, Research Department, National Neuroscience Institute, Singapore*

Chapter 1

Cloning and Identification of Recombinant Argonaute-Bound Small RNAs Using Next-Generation Sequencing

Pooja Gangras, Daniel M. Dayeh, Justin W. Mabin, Kotaro Nakanishi, and Guramrit Singh

Abstract

Argonaute proteins (AGOs) are loaded with small RNAs as guides to recognize target mRNAs. Since the target specificity heavily depends on the base complementarity between two strands, it is important to identify small guide and long target RNAs bound to AGOs. For this purpose, next-generation sequencing (NGS) technologies have extended our appreciation truly to the nucleotide level. However, the identification of RNAs via NGS from scarce RNA samples remains a challenge. Further, most commercial and published methods are compatible with either small RNAs or long RNAs, but are not equally applicable to both. Therefore, a single method that yields quantitative, bias-free NGS libraries to identify small and long RNAs from low levels of input will be of wide interest. Here, we introduce such a procedure that is based on several modifications of two published protocols and allows robust, sensitive, and reproducible cloning and sequencing of small amounts of RNAs of variable lengths. The method was applied to the identification of small RNAs bound to a purified eukaryotic AGO. Following ligation of a DNA adapter to RNA 3′-end, the key feature of this method is to use the adapter for priming reverse transcription (RT) wherein biotinylated deoxyribonucleotides specifically incorporated into the extended complementary DNA. Such RT products are enriched on streptavidin beads, circularized while immobilized on beads and directly used for PCR amplification. We provide a stepwise guide to generate RNA-Seq libraries, their purification, quantification, validation, and preparation for next-generation sequencing. We also provide basic steps in post-NGS data analyses using Galaxy, an open-source, web-based platform.

Key words Small RNAs, Argonaute, Next-generation sequencing, Biotinylated dNTPs, Low RNA input

1 Introduction

Small RNAs are non-coding RNAs critical for most eukaryotes across animals, plants, and fungi. They are classified into microRNAs, small interfering RNAs, and Piwi-interacting RNAs based on their structure and function [1]. These RNAs range in size from ~19 to 25 nucleotides (nt) and participate in a wide variety of cellular processes such as heterochromatin formation, messenger RNA destabilization, translational repression, and transposon silencing [2, 3]. Small RNAs function within ribonucleoprotein complexes called

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_1,

RNA-induced silencing complex (RISC) following their incorporation into Argonaute proteins (AGOs) [4]. The loaded RNA serves as a guide to recruit the RISC to target RNAs with full or partial sequence complementarity. In the last decade, next-generation sequencing (NGS) has illuminated the significance of the base complementarity between the guide and target strands at the nucleotide level. This technology has enabled researchers to gain a greater understanding of small RNAs found in different biological samples as well as to reveal the diversity of small RNA sequences bound to the purified AGOs [5–9]. NGS-based approaches have also enabled high-throughput detection of long target RNAs [10]. Concurrently, strides are being made in expression and purification of recombinant AGOs for their biochemical and structural analysis. When purified from heterologous sources, AGOs co-purify small RNAs; the identification and analysis of such RNAs holds promise for unveiling the underlying principles of guide RNA selection by AGOs [11].

Here, we present a method for preparation of NGS libraries from small RNAs that co-purify with a yeast AGO. The described workflow is primarily based on a previously published method that extensively optimized each individual enzymatic step in RNA-Seq library preparation [12]. The specificity and sensitivity of the method was further improved by incorporating a feature from another kit-free RNA-Seq library preparation approach wherein complementary DNAs (cDNAs) from adapter-ligated small RNAs are specifically enriched using biotinylated deoxyribonucleotides (dNTPs) [13]. Since this chapter aims for small RNA-seq, the protocol does not include RNA fragmentation step that is required for NGS of long RNAs. Note that, however, this method is applicable to longer RNAs purified from any biological source, via a broad range of approaches such as RNA–protein immunoprecipitation, ribosome footprinting, and is even suitable for extracellular RNAs. In cases where long RNAs may need to be fragmented to input into this procedure, chemical methods of RNA fragmentation are much more desirable over enzymatic methods [14, 15].

A schematic of the whole procedure is depicted in Fig. 1. Isolated RNAs are first enzymatically treated to generate a 3′-hydroxyl group to allow ligation to a pre-adenylated DNA adapter. The adaptor-ligated RNAs are used to perform reverse transcription (RT) using biotinylated dNTPs and one of the 12 RT primers to produce biotinylated cDNAs containing a specific 5-nucleotide barcode (Table 1). Another feature of the RT primers

Fig. 1 (continued) The biotinylated RT product is shown captured on the streptavidin-conjugated magnetic beads (*gray sphere*) whereas unextended RT primer is shown to be selected away. (*6*) The *circularized* RT product is shown where the G at the 5′ terminus of the forward primer end is shown ligated to the 3′ end of the extended cDNA. (*7*) Illumina primers PE 1.0 and PE 2.0 are as *gray* and *black arrows*, respectively, and are shown bound to their complementary sequences in the *circularized* RT product. Note that biotinylated nucleotides in the cDNA do not interfere with PCR amplification. (*8*, *9*) and (*10*) describe main steps post cDNA library preparation

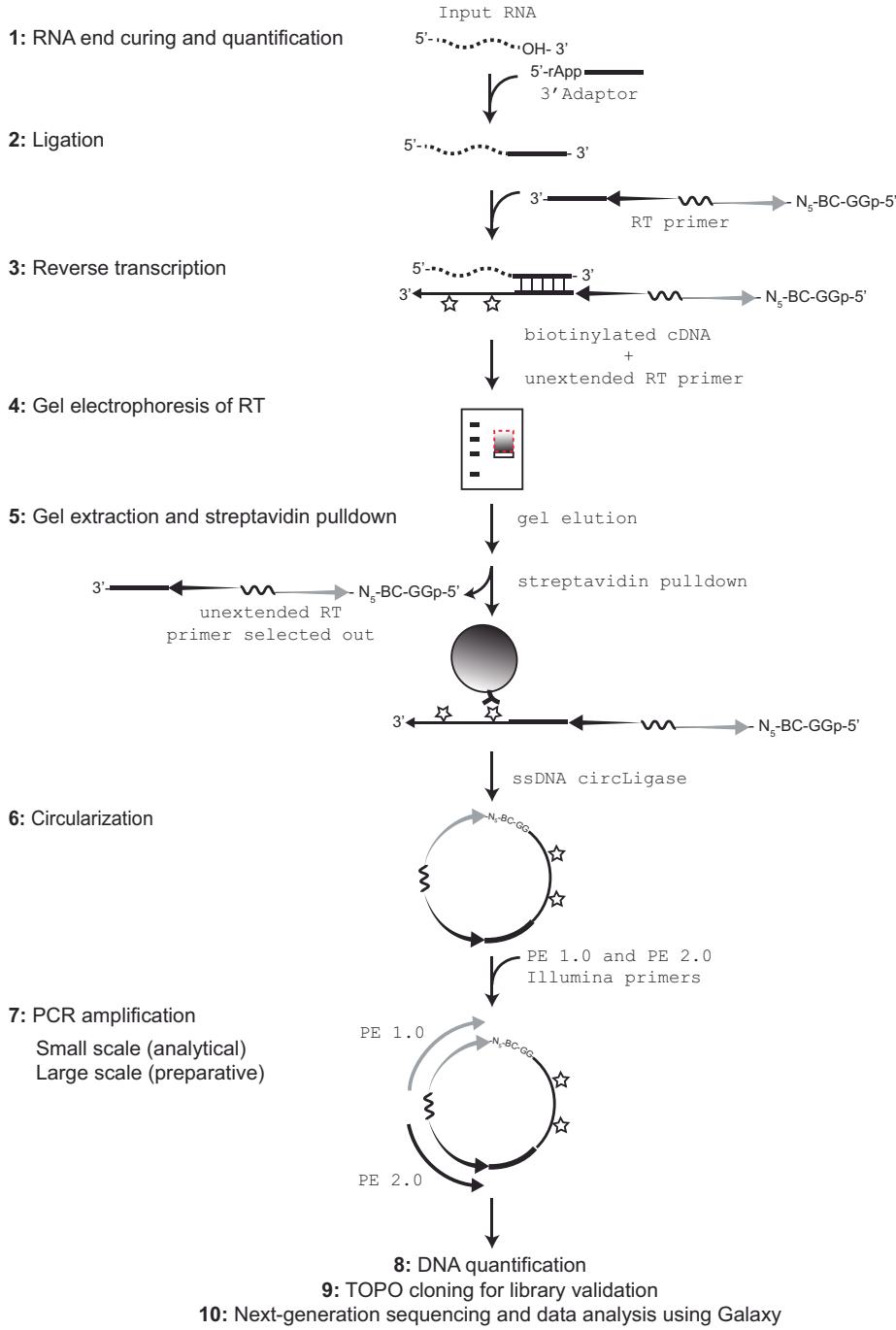


Fig. 1 A schematic of the steps in the NGS library prep procedure. (*1*) *Dotted wavy line* depicts input RNA with a 3′ hydroxyl. (*2*) Pre-adenylated 3′ adaptor is shown as a *thick black line* ligated to RNA 3′-end. (*3*) 3′ adapter complementary sequences in the RT primer are shown as a *thick black line*. In the RT primer, the sequence corresponding to forward sequencing primer is represented by a *gray tapered arrow*, and reverse primer is shown as a *black tapered arrow*. The hexaethylene glycol spacer SP18 that connects the forward and reverse primers is shown as a *solid wavy line*. At the 5′ end of the forward primer are a random 5-mer, the 5-nt TruSeq barcode sequence (BC) and two Gs. The 5′ phosphate on the terminal G is also shown. Following RT, the incorporated biotinylated dNTPs are represented by *stars*. (*4*) *Dotted box* marks the area of the gel that is excised out for elution of the RT product. (*5*)

Table 1

***Top*: Name and full sequence of each of the 12 RT oligos used in Subheading 3.3. The hexaethylene glycol spacer SP18 and 5′–3′ orientation of DNA sequences linked on its each end are indicated. The barcode sequence is *underlined*. *Bottom*: Name and full sequence of the PCR oligos used for amplification of cDNA libraries in Subheading 3.7. The position of the phosphorothioate linkage in each primer is shown**

RT primer sequences	
Name	**Sequence**
TruSeq_SE1	5′-pGGCACTANNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE2	5′-pGGGTAGCNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE3	5′-pGGTCGATNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE4	5′-pGGCCTCGNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE5	5′-pGGTGACANNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE6	5′-pGGTAGACNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE7	5′-pGGGCCCTNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE8	5′-pGGATCGGNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE9	5′-pGGACTGANNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE10	5′-pGGTGTTCNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′

TruSeq_SE11	5′-pGGTAAGTNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
TruSeq_SE12	5′-pGGAGATGNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-SPACER 18-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCTTGGCACCCGAGAATTCCA-3′
PCR Primers (Illumina PE primers; * indicates location of phosphorothioate bond)	
Name	**Sequence**
PE1.0 PCR Primer	5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′
PE2.0 PCR Primer	5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T-3′

is the inclusion of five random nucleotides at positions that will be the first five nucleotides sequenced (*see* Fig. 3d). This random sequence ensures sequence complexity during first few sequencing cycles, and also allows for removal of any sequencing reads that arise due to PCR duplication. Following RT, reaction is run on a gel, the desired product is extracted and enriched via streptavidin pulldown. The purified RT product is then circularized where the 3′-end is ligated to the 5′-end by circLigase, and the circular molecule serves as a template for PCRs. We describe the utility of small-scale PCRs to determine the appropriate number of cycles for production of the highest concentration of libraries without accumulation of nonspecific PCR products. Large-scale PCR under the determined conditions specifically amplifies the cDNA libraries for NGS. The purified PCR product then undergoes series of quantification and quality control steps before submission for NGS. After testing serially diluted synthetic RNAs as input, we have determined that our protocol can successfully generate complex cDNA libraries with as low as 0.01 pmol of RNA.

2 Materials

2.1 Equipment and General Supplies

1. Thermocycler (*see* **Note 1**).
2. Blue light transilluminator (VWR).
3. Typhoon scanner (or a similar imager compatible with phosphorescent screens and fluorescent DNA binding stains).
4. Phoshor-imager screens and cassettes.
5. Gel dryer (for vacuum and heat-assisted gel drying).
6. Access to Qubit fluorometer and Bioanalyzer or TapeStation.
7. 0.2 ml PCR tubes.
8. Vertical gel electrophoresis:

Apparatus	(C.B.S. scientific)
Metal heat sink	(C.B.S. scientific)
Large glass plates, 20 cm W × 28 cm H	(C.B.S. scientific)
Short glass plates, 20 cm W × 10 cm H	(Moliterno)
Spacers, 0.15 cm	(C.B.S. scientific)
Spacers, 0.1 cm	(C.B.S. scientific)
Combs	
8-well, 0.15 cm thick	(C.B.S. scientific)
16-well, 0.1 cm thick	(C.B.S. scientific)
8-well, 0.1 cm thick	(C.B.S. scientific)

9. 0.5× Tris/Borate/EDTA buffer (5× TBE: 54 g of Tris base, 27.5 g of boric acid, 20 ml of 0.5 M EDTA (pH 8.0)).
10. 40% (w/v) acrylamide:bisacrylamide (29:1 Accugel, National Diagnostics).
11. TEMED.
12. 10% Ammonium persulfate (APS).
13. SYBR® gold nucleic acid gel stain (ThermoFisher Scientific).
14. 25 bp DNA ladder (Invitrogen).
15. Glass baking tray for gel staining.
16. Standard clear sheet protector (e.g. Staples).
17. Razor blades (VWR).
18. Luer-lok™ Tip syringes (3 and 10 ml) and needles (22 gauge).
19. Magnetic rack (e.g. DynaMag™-2; ThermoFisher Scientific).
20. Corning® Costar® Spin-X columns (Sigma-Aldrich).
21. RNase-free water (ThermoFisher Scientific).
22. 10 μg/μl Glycogen (ThermoFisher Scientific).
23. 100% Ethanol (200 proof).

Materials to be prepared ahead of time in lab:

24. 3 M sodium acetate (pH 5.2).
25. 2 M magnesium chloride ($MgCl_2$).
26. 2× Denaturing Load Buffer: 3 ml 5× TBE, 1.8 g Ficoll Type 400, 6.3 g Urea, 3 mg bromophenol blue, 3 mg xylene cyanol, up to 15 ml ddH_2O. To get into solution, place tube in water in a beaker and boil on hot plate for 10–15 min. Add dyes after adjusting the volume to 15 ml. Store at 4 °C. Before use, heat to 50–60 °C to fully solubilize contents if necessary (*see* **Notes 2 and 3**).

2.2 RNA End Curing and Quantification

1. 10× CutSmart® buffer (NEB).
2. Calf intestine phosphatase (NEB).
3. Acid Phenol mix (Phenol:chloroform:iso-amyl alcohol, pH 4.5; ThermoFisher Scientific).
4. Low molecular weight single-stranded (ss) DNA ladder (Alfa Aesar).
5. Synthetic RNA oligo (any 20–40 nucleotide long sequence).
6. 1 mM ATP.
7. γ^{32}P-ATP 3000 Ci/mmol (Perkin Elmer).
8. 10× T4 Polynucleotide Kinase (PNK) buffer (NEB).
9. T4 PNK enzyme (NEB).

10. Saran wrap.
11. 3 M Whatman paper.
12. **Items 7–12** and **21–26** from Subheading 2.1.

Materials to be prepared ahead of time in lab:

13. 20% Urea-PAGE gel mix (20% acrylamide:bisacrylamide (stock: 40% (w/v)), 6 M Urea, 0.5× TBE).

2.3 Ligation

1. 50% PEG8000 (NEB; PEG8000 is currently supplied with T4RNL2 Tr. K227Q) (*see* **Note 4**).
2. 10× Ligation Buffer (10× T4RNL2 Tr. K227Q Buffer; NEB).
3. T4RNL2 Tr. K227Q enzyme (NEB).
4. 20 mM DTT.
5. 7 μM Pre-adenylated adaptor (miRCat-33® Conversion Kit from IDT) for single-end Sequencing:

5′-rAppTGGAATTCTCGGGTGCCAAGGddC-3′ (*see* **Notes 5** and **6**).

2.4 Reverse Transcription

1. 100 mM dATP, dCTP, dTTP, dGTP (NEB).
2. 100 mM DTT.
3. 1 mM Biotin-dATP (Metkinen Biotin-11-dATP or PerkinElmer Biotin-11-dATP).
4. 5 mM Biotin-dCTP (Trilink Biotin-16-AA-2′dCTP).
5. 10 μM TruSeq RT primers (*see* Table 1 for sequences).
6. Superscript III (SSIII) reverse transcriptase (Invitrogen).

Materials to be prepared ahead of time in lab:

7. 5× First-Strand Buffer w/o $MgCl_2$: 250 mM Tris-HCl (pH 8.3 at room temp), 375 mM KCl.
8. 4× dNTP mix: dGTP—0.25 mM; dTTP—0.25 mM; dATP—0.175 mM; biotin-dATP—0.075 mM; dCTP—0.1625 mM; biotin-dCTP—0.0875 mM.

2.5 Gel Electrophoresis of RT Product

1. **Items 7–14** from Subheading 2.1.

Materials to be prepared ahead of time in lab:

2. 10% Urea-PAGE gel mix (10% acrylamide:bisacrylamide (stock: 40% (w/v)), 6 M Urea, 0.5× TBE).
3. SYBR® gold staining solution: 20 μl of 10000× SYBR® gold in 200 ml of 0.5× TBE.

2.6 Gel Elution and Streptavidin Pulldown

1. **Items 15–20** from Subheading 2.1.
2. Hydrophilic Streptavidin magnetic beads (NEB).

Materials to be prepared ahead of time in lab:

3. DNA Elution Buffer: 300 mM NaCl, 1 mM EDTA.
4. Streptavidin bead wash buffer: 0.5 M NaOH, 20 mM Tris-HCl pH 7.5, 1 mM EDTA.
5. Streptavidin bead resuspension buffer: 10 mM Tris HCl pH 7.5, 0.1 mM EDTA, 0.3 M NaCl.

2.7 Circularization

1. CircLigase™ reaction kit (CircLigase reaction buffer, CircLigase enzyme and 50 mM $MnCl_2$) (Epicentre Biotechnologies).
2. 1 mM ATP.
3. 5 M Betaine solution (Sigma-Aldrich).

2.8 PCR Amplification

1. Phosphorothioate Illumina PE1.0 and PE2.0 primers (*see* Table 1 for sequences).
2. 5× Q5 reaction buffer (NEB).
3. Q5 DNA polymerase (NEB).
4. 10 mM dNTPs.
5. 6× DNA gel loading dye (NEB).
6. SYBR® gold staining solution: 20 μl of 10,000× SYBR® gold in 200 ml of 0.5× TBE.
7. **Items 8–16** from Subheading 2.1.

2.9 DNA Quantification and Sample Preparation

1. **Items 8–13** from Subheading 2.1.
2. 100 bp DNA ladder (NEB).

2.10 TOPO Cloning to Validate Deep Seq Libraries

1. TOPO® TA Cloning® Kit for Subcloning, without competent cells (Invitrogen).
2. 10× Taq buffer (NEB).
3. 10 mM dATP.
4. GoTaq® Hot Start Green Master Mix (Promega).
5. Taq polymerase (5000 U/ml, NEB).
6. Bacterial Competent cells (TOP10 or DH5α).
7. PCR purification columns (e.g. Qiagen).
8. M13 forward and M13 reverse primers.

Materials to be prepared ahead of time in lab:

9. Luria-Bertani + Ampicillin (LB + Amp) plates: 10 g/l Tryptone, 10 g/l NaCl, 5 g/l Yeast Extract, 15 g/l Agar; pH 7.4; add 100 μg/ml Ampicillin after autoclaving.

2.11 Next-Generation Sequencing and Data Analysis Using Galaxy

1. User account on Galaxy (https://usegalaxy.org), an open source, web-based environment for analysis and manipulation of NGS data.

3 Methods

3.1 RNA End Curing and Quantification

1. There are two important considerations for the RNAs that will be input into the method described here. First, this method requires a hydroxyl group at the RNA 3′-ends. Most small RNAs have a 3′-hydroxyl group and hence can be directly input into ligation reaction in **step 1** of Subheading 3.2. RNAs used for the NGS library preparation in our laboratories were co-purified with a fragment of *Kluyveromyces polysporus* Argonaute1 (KpAGO). The recombinant protein was expressed and purified from *E. coli* cells as reported previously [11]. The size of these RNAs falls into two distinct size classes that peak around 13 and 19 nucleotides (Fig. 2a). Consistent with previous crystal structures of eukaryotic AGOs, these RNAs possess a 5′-monophosphate [11, 16–19]. However, the exact nature of these RNA 3′-ends (3′-hydroxyl or 3′-phosphate) is not known. Therefore, we first describe a Calf Intestine phosphatase (CIP)-mediated dephosphorylation step to ensure that all RNAs end in a 3′-hydroxyl. This treatment will be necessary in other cases when a small RNA (or any RNA for that matter) does not end in a 3′-hydroxyl group (e.g. chemical or enzymatic nucleolysis yields RNAs with 3′-phosphate). Second, this method can be executed most reliably if input RNA concentration is precisely known. To quantify input RNA, we prefer to use a radioactive method that compares the number of 5′-ends in the RNA sample to a small synthetic RNA oligo of known size and concentration. This is a very sensitive approach that is particularly beneficial when input RNA amounts are low (nanomolar range) and cannot be reliably quantified using spectroscopic methods. As small RNAs often have a 5′-monophosphate, the CIP treatment described here also converts the 5′-phosphate to a 5′-hydroxyl group, which can be subsequently labeled with a radioactive phosphate using T4 polynucleotide kinase (PNK) (*see* **Note 7**).

RNA End Curing

2. Add the following to an eppendorf tube:

10× CutSmart® buffer	2 µl
RNA	*x* µl
CIP enzyme	1 µl
RNase-free water	to 20 µl

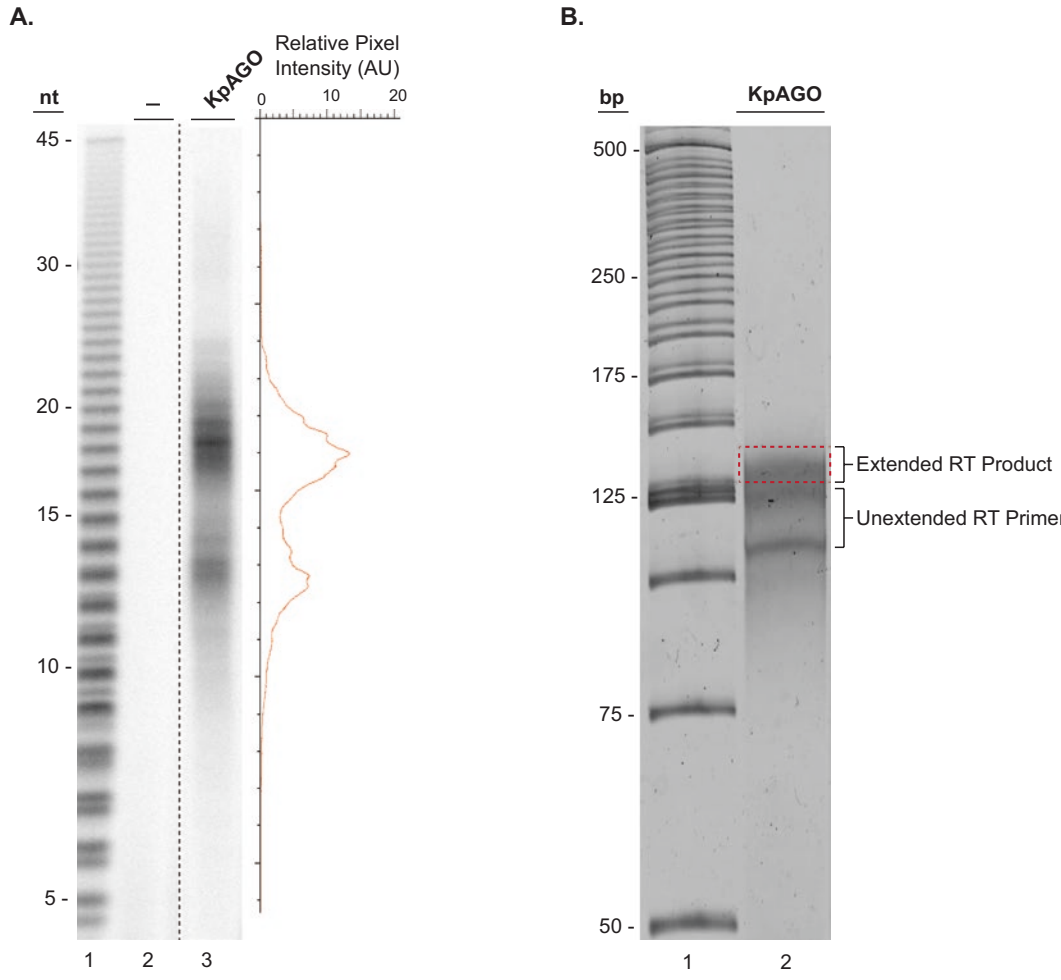


Fig. 2 Input RNA size estimation and reverse transcription product size selection. (**a**) A 16% denaturing urea PAGE of ^{32}P-labeled endogenous *E. coli* RNAs co-purified from KpAGO (*lane 3*) and a mock sample (*lane 2*). On the *right*, the pixel intensity profile of *lane 3* shows the presence of two distinct RNA populations centered at 13 and 19 nt in length from the purified KpAGO. A hydrolyzed poly uridine 45-mer was used to estimate relative size of the input RNA (*lane 1*). (**b**) A SYBR® gold stained 10% denaturing urea-PAGE gel showing the extended RT product smear located above the unextended RT primer (*lane 2*). The excised gel region for elution of the DNA product is indicated by a *dashed red rectangle*. The 25-bp DNA ladder is in *lane 1*

3. Incubate at 37 °C for 30 min (*see* **Note 8**).
4. To the same tube, add the following:

500 mM EDTA	1 µl
Water	380 µl
Acid Phenol mix	400 µl

5. Vortex for at least 1 min at room temperature.
6. Centrifuge at 12,000 × *g* at room temperature for 5 min.
7. Carefully pipet 350 µl of the upper aqueous phase into a new eppendorf tube (*see* **Note 9**).
8. Add the following to the tube:

3 M sodium acetate (pH 5.2)	40 µl
2 M Magnesium chloride	10 µl
Glycogen	2 µl
100% Ethanol	1000 µl (2.5× of aqueous volume)

9. Precipitate RNA overnight at −20 °C.
10. Pellet RNA at 12,000 × *g* in a refrigerated (4 °C) centrifuge for 30 min. Carefully discard supernatant and wash with 1 ml of 70% ethanol by spinning at 12,000 × *g* at 4 °C for 5 min.
11. Discard supernatant and quick spin the tube. With a 20 µl pipet, carefully remove any residual liquid. Air dry the pellet for 1 min.
12. Resuspend RNA pellet in a small volume (≤10 µl).

RNA Quantification

13. Add the following to a new eppendorf tube on ice:

10× PNK buffer	1 µl
RNA (from **step 12**)	1 µl
1 mM ATP	0.5 µl
γ^{32}P-ATP	20 µCi
T4 PNK	0.5 µl
RNase-free water	to 10 µl
(*see* **Note 10**)	

14. Set-up parallel T4 PNK reactions to label 0.1 pmol of a synthetic RNA oligo, and 1 µl low molecular weight ssDNA ladder.
15. Incubate at 37 °C for 30 min.
16. Add 10 µl of 2× Denaturing Load Buffer to each reaction.
17. Heat at 65 °C for 2–5 min.

18. Prepare a 20% Urea-PAGE gel (20 × 27 × 0.1 cm) with 16 wells (*see* **Note 11**).
19. Pre-run gel at 35 W (constant power setting) for at least 30 min (*see* **Note 12**).
20. Load 10 μl of the radiolabeled input RNA, 10 μl of the radiolabeled RNA oligonucleotide, and 2 μl of the radiolabeled low molecular weight DNA ladder from **step 16** above (*see* **Note 13**).
21. Run the gel at 35 W until bromophenol blue migrates 3/4th the length of the gel.
22. Carefully lift-off gel onto 3M Whatman filter paper and cover with saran wrap (*see* **Note 14**).
23. Dry gel for 1 h at 80 °C in a gel dryer (*see* **Note 15**).
24. Expose gel to phosphorimager screen overnight to visualize ^{32}P-labeled RNA.
25. Quantify the volume of signal of the input RNA and the RNA oligo used as standard.
26. Determine molar concentration of the input RNA by comparing its signal to that of the known RNA oligonucleotide.
27. If needed, adjust the concentration of the remaining input RNA (**step 12**) such that 3′-ends are between 0.015 and 0.5 pmol/μl (μM). This will yield ~0.05–2 pmol of 3′-ends in 3.8 μl, which is optimal for efficient ligation reaction in the next step, and eventually yields a complex library.

3.2 Ligation

1. Thaw the following and keep on ice: 50% PEG8000, 10× T4RNL2 Tr. K227Q Buffer, 20 mM DTT, and 7 μM pre-adenylated adaptor.
2. Add the following to a PCR tube on ice:

7 μM pre-adenylated adaptor	1.0 μl
RNA	3.8 μl
RNase-free water	to 4.8 μl

3. Incubate in a thermocycler at 65 °C for 10 min. Rapidly cool to 16 °C and hold for 5 min. Samples can be on hold at 4 °C.
4. Transfer tubes from thermocycler to ice and add the following:

10× Ligation buffer	1.5 µl
50% PEG8000	7.5 µl
20 mM DTT	0.75 µl
T4RNL2 Tr. K227Q	0.45 µl
Final volume	15.0 µl

5. Incubate in a thermocycler at 30 °C for 6 h.
6. Heat inactivate the enzyme at 65 °C for 20 min. Samples can be on hold at 4 °C.

3.3 Reverse Transcription

1. Thaw the following and keep on ice: RT Primer, 4× dNTP mix, 5× First-Strand Buffer without $MgCl_2$, 100 mM DTT.
2. To the 15 µl ligation mix from **step 6** in Subheading 3.2, add the following:

10 µM TruSeq RT primer	1.0 µl
10 mM dNTP mix	11.25 µl
RNase-free water	6.8 µl

3. Incubate at 65 °C for 5 min. Hold at 4 °C for at least a minute.
4. Transfer tubes to ice and add the following:

5× FS w/o $MgCl_2$ Buffer	9.0 µl
100 mM DTT	2.25 µl
SSIII (200 U/µl)	1.2 µl
Final volume	45.0 µl

5. Incubate at 55 °C for 30–45 min.
6. Heat inactivate RT reaction at 70 °C for 15 min. The sample can be on hold at 4 °C.

3.4 Gel Electrophoresis of RT Product

1. To RT reaction (from **step 6** in Subheading 3.3), add 45 µl of 2× denaturing urea load buffer.
2. Prepare 10% denaturing Urea-PAGE gel (20 × 28 × 0.15 cm) with eight 1.7 cm-wide wells (*see* **Note 16**).
3. Pre-run the gel at 35 W for at least 30 min (*see* **Note 17**).

4. Load samples using P20 tips. For size markers, mix 6 μl of 25 bp ladder (Invitrogen) with 34 μl of 2× urea denaturing buffer and load 20 μl on either side of the sample (*see* **Note 18**).
5. Run gel at 35 W until bromophenol blue is 3/4th from the top of the gel.
6. Stain the gel with SYBR® gold for 5 min using SYBR® gold staining solution.
7. Transfer gel to a plastic sheet protector cut open on three sides (*see* **Note 19**).
8. Scan on a Typhoon scanner using SYBR® gold compatible excitation (520 nm) and emission (580 nm) filters.
9. Print an image of the gel for documentation and to reliably identify the desired bands/regions to be excised. Figure 2b shows a gel where the to-be-excised fragment with the extended RT product is indicated (*see* **Note 20**).
10. To excise the desired size gel fragments, place gel on a blue light transilluminator and cut out bands using razor blades (*see* **Note 21**).

3.5 Gel Elution and Streptavidin Pulldown

1. To crush a gel fragment, place it in the barrel of a 3 ml syringe (piston removed), and extrude the gel fragment using the piston into a 2 ml eppendorf tube.
2. Add 800 μl of DNA elution buffer to the eppendorf tube with gel pieces. Nutate overnight at room temperature (*see* **Note 22**).
3. For each sample, aliquot 10 μl of streptavidin beads into a 1.5 ml eppendorf tube. Add 200 μl streptavidin bead wash buffer. Flick tube with a finger to uniformly suspend the beads. Place tube in a magnetic rack to capture the beads. Carefully pipet-off the supernatant and discard. Repeat wash steps at least three times (*see* **Note 23**).
4. After the final wash, resuspend the magnetic beads in 10 μl of streptavidin bead resuspension buffer.
5. Transfer gel slurry from **step 2** above to Spin-X column placed in its collection tube. Spin at 10,000 × *g* for 3 min, at room temperature (*see* **Note 24**).
6. If needed, pool elution into a single 2 ml eppendorf tube. Add streptavidin magnetic beads, and nutate for 2–3 h at room temperature (*see* **Note 25**).
7. Place the tube on a magnetic rack to capture magnetic beads and discard supernatant.
8. Resuspend beads in 10 μl of RNase-free water.

3.6 Circularization

1. Thaw the following reagents and keep on ice: CircLigase Reaction Buffer, 1 mM ATP, 50 mM $MnCl_2$.
2. Transfer the bead slurry (10 µl; **step 8** in Subheading 3.5) to a PCR tube and add the thawed reagents as described below:

RT Product	10.0 µl (Bead slurry)
CircLigase Reaction Buffer	2.0 µl
1 mM ATP	1.0 µl
50 mM $MnCl_2$	1.0 µl
5 M Betaine	4.0 µl
CircLigase	1.0 µl
RNase-free water	to 20 µl

3. Mix the contents of the PCR tube well by gently flicking the tube. Quick spin, and incubate the reaction at 60 °C for 4 h.
4. Heat inactivate CircLigase by heating to 80 °C for 10 min in a thermocycler.

3.7 PCR Amplification

1. The number of PCR cycles required to generate enough cDNA for deep sequencing will depend on the amount of the RT product, and in turn, on the amount of input RNA. The exact number of cycles needed for each sample will have to be empirically determined. If RT product was abundant enough to be readily visible on the urea-PAGE gel in **steps 8–10** of Subheading 3.4, 6–10 PCR cycles will yield sufficient amount of DNA (Fig. 3a). However, if input RNA amounts were lower and RT product was faint or undetectable in **step 8** in Subheading 3.4, one will have to test the higher number of PCR cycles (between 15 and 18 cycles). Perform small-scale test PCRs first to test the appropriate number of PCR cycles for a sample. Another goal of small-scale test PCRs is to identify a minimum number of PCR cycles for library amplification. This is very important to preserve the biological complexity of the input RNA sample, and also to prevent the appearance of artifactual PCR products (*see* Fig. 3a, lane 4). In our experience, these two conditions can be easily met by limiting amplification to the lowest PCR cycle number that yields distinct PCR product and where free (unused) primer pool is minimally depleted.

Small-Scale Test PCRs

2. Thaw the following and keep on ice once thawed: phosphorothioate PE1.0 and PE2.0 primers and 5× Q5 polymerase buffer, dNTPs.

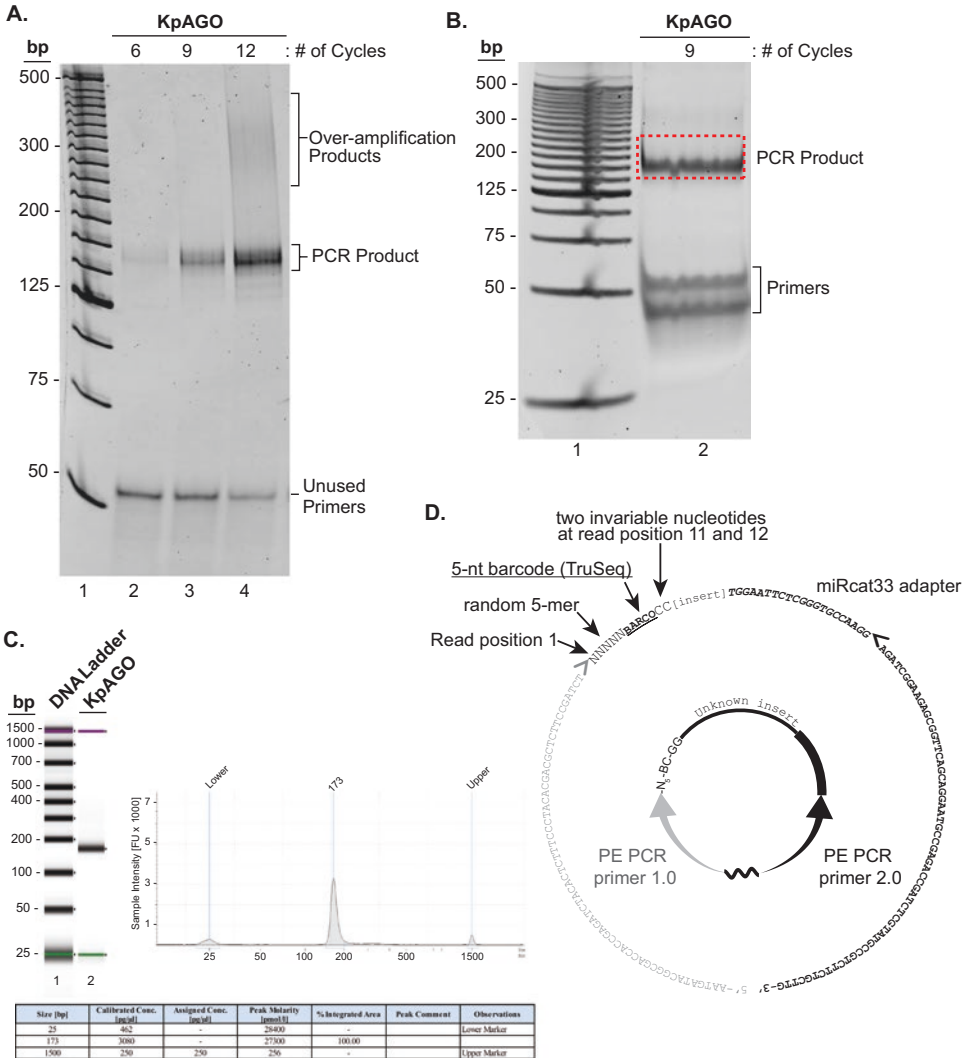


Fig. 3 PCR amplification and size estimation of NGS libraries. (**a**) SYBR® gold stained 8% nondenaturing PAGE-gel showing specific PCR products, over-amplification products, and unused primers (*lanes 2–4*; each species is indicated on the *right*) in small-scale PCRs performed for increasing number of cycles indicated on *top* of each lane. The 25-bp DNA ladder is in *lane 1*. (**b**) SYBR® gold stained 8% nondenaturing PAGE showing the product band and unused primers (labeled on the *right*) from large-scale PCR (*lane 2*). The 25-bp DNA ladder is in *lane 1*. The gel fragment excised to purify DNA for NGS is indicated by a *red-dotted rectangle*. (**c**) The Bioanalyzer report of the purified PCR product from (**b**). On the *left* is a virtual gel image of the DNA ladder (*lane 1*) and cDNA library of KpAGO-bound RNAs (*lane 2*). On the *right* is the histogram showing the size of the cDNA library along with size markers (*top*) and a table with a summary of size and amount quantification of the library (*bottom*). Note that while Bioanalyzer analysis also provides sample concentration (*see* table), fluorescent DNA binding dye based quantification provides more accurate quantification of DNA amount. (**d**) The *circularized* cDNA product as in Fig. 1 is shown in the *center*. The corresponding sequence of the final PCR product in 5′–3′ direction is shown as the *outer circle*, and its different segments are individually labeled. *Right* and *left pointing angled brackets* indicate the start position of DNA synthesis during PCR. The *gray right-pointing angled bracket* also indicates where the Illumina first-read sequencing primer begins extension

3. Add the following to a PCR tube:

5× Q5 buffer	9.0 µl
10 mM dNTPs	0.9 µl
10 µM PE 1.0	2.25 µl
10 µM PE 2.0	2.25 µl
Circularization reaction	4–6 µl
Q5 polymerase	0.45 µl
RNase-free water	to 45 µl
(*see* **Notes 26** and **27**)	

4. Split the reaction into three equal parts (*see* **Note 28**).
5. Subject each reaction to a different number of PCR cycles (e.g. 5, 8 and 11 cycles) using the following amplification conditions:

   98 °C—30 s.

   98 °C—5 s.

   65 °C—10 s.

   72 °C—15 s.

   72 °C—2 min.

   12 °C—hold.

   (repeat steps in **bold** for the desired number of cycles).
6. Mix PCR reactions with 6× DNA loading dye and load on 20 × 28 × 0.1 cm 8% nondenaturing PAGE (*see* **Notes 29** and **30**).
7. Run gel at 35 W until bromophenol blue is 3/4th from the top of the gel.
8. Stain the gel with SYBR® gold for 5 min in SYBR® gold staining solution.
9. Transfer the gel to a clear plastic sheet protector as in **step 7** in Subheading 3.4 to scan on a Typhoon scanner. The optimal number of cycles needed for sufficient amplification of the library will correspond to the minimum number of cycles that provide readily detectable levels of expected size PCR products such that there is <10% depletion of free primers and complete absence of aberrant slow migrating DNA products (*see* **Notes 31** and **32**).

Large-Scale PCRs

10. Redo 2 × 45 µl PCRs (assembled as in **step 3** above) for a fixed number of PCR cycles determined above.

11. Separate large-scale PCR products on a short (20 × 12 × 0.1 cm) 8% nondenaturing PAGE at constant volts (150 V max) (*see* **Note 33**).
12. Stain the gel in SYBR® gold and image on a Typhoon scanner as described in **step 8** in Subheading 3.4 (*see* **Note 34**).
13. Excise gel fragment containing PCR products of the expected size. Figure 3b shows a gel with the excised fragment containing the large-scale PCR product.
14. Crush gel fragment and perform DNA elution as described in **steps 1** and **2** of Subheading 3.5 (*see* **Notes 35** and **36**).
15. Precipitate DNA by adding 2.5 volumes of 100% ethanol and 10 μg glycogen. Store at −20 °C for at least 2 h.
16. Pellet DNA at 12,000 × *g* at 4 °C for 30 min. Wash the DNA pellet two times with 70% ethanol. Remove residual ethanol via quick spin and fine pipetting.
17. Immediately resuspend pellet in 20 μl water (*see* **Notes 37** and **38**).

3.8 DNA Quantification and Sample Preparation for Next-Generation Sequencing

1. To prepare DNA sample for next-generation sequencing, both the size and the amount of PCR product have to be carefully quantified. NGS libraries usually result in low sub-nanogram DNA yields, and therefore require highly sensitive quantification methods. DNA size in a library is best quantified via automated chip-based electrophoresis systems such as Bioanalyzer (high sensitivity DNA chip) or TapeStation (DNA and RNA ScreenTape). The Bioanalyzer trace and stats for the library prepared from the example input sample is shown in Fig. 3c. DNA concentration of a library can be accurately quantified using a Qubit fluorometer. These instruments are now part of workflows at most NGS facilities, which often provide access to these services for a small per sample charge. If access to these instruments is not available, precise size and amount quantification of NGS library can be carried out using standard lab equipment as in the following steps.
2. Prepare an 8% native PAGE short gel as in **step 11** of Subheading 3.7.
3. Load 2 μl of NGS library DNA along with 1.5, 0.5, and 0.15 μl of 100 bp NEB ladder on the gel. Run gel at 150 V until bromophenol blue dye front runs a 3/4th length of the gel.
4. Stain the gel with SYBR® gold as described above (**step 8** of Subheading 3.7).
5. Scan the gel on a Typhoon scanner.
6. Quantify DNA using software such as ImageQuant. Use the amount of DNA in a specific band of the ladder (the one most

comparable to intensity of your PCR bands) as a reference for quantification.

7. After quantification, a small amount of PCR product can be used for cloning into T-tailed vectors (*see* Subheading 3.9).
8. Convert DNA concentration from ng/μl to nM (fmol/μl) (*see* **Note 39**).
9. Assuming 40 fmol of DNA will be loaded in one lane of an Illumina flow cell, different barcoded libraries can be pooled depending on the number of reads desired for each library to obtain 40 fmol of DNA in a particular volume. For example, if three samples are to be mixed to get reads at a ratio of 1:1:2, mix them at 10 fmol: 10 fmol: 20 fmol (*see* **Note 40**).
10. If the pooled library is to be quantified again prior to sequencing, obtain the average size of the DNA in the pooled mix as follows:

    ((fmol of library 1 × Average BP Size of Library 1) + (fmol of library 2 × Average BP Size of Library 2) + … (fmol of library *n* × Average BP Size of Library *n*))/40 total fmol
11. Determine the concentration of the pooled library via Qubit, and using the equation in **Note 39** and average bp size from **step 10**, verify if nM concentration of your pool matches to that expected.

3.9 TOPO Cloning to Validate NGS Library

1. A small amount of PCR product can be cloned into T-tailed TOPO TA-cloning vector and transformed into *E. coli*. Using universal primers, inserts from a handful of bacterial colonies can be PCR amplified and sequenced via Sanger sequencing to validate the library before deep-sequencing.
2. The Q5 polymerase used to PCR amplify NGS libraries is a proofreading enzyme. So, unlike Taq, this polymerase does not leave a 3′-A overhang. To add an A-overhang at the ends of the PCR product to enable cloning into T-tailed TOPO vector, mix the following in a PCR tube:

PCR product	4.0 μl (*see* **Note 41**)
10× Taq Buffer	0.5 μl
10 mM dATP	0.125 μl
Taq polymerase	0.25 μl
Water	5 μl (0.125 μl)

3. Incubate the PCR tube in a thermocycler at 72 °C for 20 min.
4. Mix the following in a PCR tube:

A-tailed PCR product	4.0 µl
Vector (pCR2.1-TOPO)	1.0 µl
Salt solution	1.0 µl

5. Incubate the PCR tube at room temperature for 5 min then move it to ice.
6. Transform 4.0 µl of above mix into 50 µl of chemical competent Top10 or DH5α cells.
7. Plate the transformed bacteria on the pre-warmed LB + Amp plate.
8. To identify insert-containing clones, first prepare the following mix in a PCR tube:

GoTaq Master Mix	10 µl
10 µM M13 forward primer	0.8 µl
10 µM M13 reverse primer	0.8 µl
Water	to 20 µl
(*see* **Note 42**)	

9. Pick a single colony with a pipet tip or toothpick and mix into the PCR mix.
10. Perform PCR in a thermocycler using following conditions:

    95 °C for 5 min.

    95 °C for 1 min.

    55 °C for 30 s.

    72 °C for 1 min.

    Repeat steps in bold, 29 times.

    72 °C for 10 min.

    4 °C forever.
11. Run 5 µl of the product on a 1.5% agarose gel.
12. For each PCR reaction that yields a single, desired-size PCR product, purify the remaining PCR reaction using PCR purification columns and perform Sanger sequencing with M13 reverse primer.

3.10 Next-Generation Sequencing and Data Analysis Using Galaxy

1. Sequence the NGS libraries prepared by this procedure on the HiSeq or MiSeq Illumina platforms in the single-end 50 bp format. These libraries use the standard Illumina first-read sequencing primer for both the barcode and insert sequencing.

The final sequence of the DNA product from the procedure is shown in Fig. 3d. The data from the sequencing pipeline is delivered in the fastq format.

2. Log into a user account on the Galaxy server (https://usegalaxy.org/) [20].
3. Upload the fastq file to Galaxy server with "Get data → Upload file".
4. Obtain a basic quality control report for fastq data file using "NGS: QC and manipulation → FastQC".
5. Groom the fastq file into fastqsanger format using "NGS: QC and manipulation → Fastq Groomer".
6. Select the high quality reads using "NGS: QC and manipulation → Filter by quality" (*see* **Note 43**).
7. Convert fastq file into fasta format using "NGS:Convert Formats → FASTQ to FASTA" (*see* **Note 44**).
8. Remove first five nucleotides (random 5-mer) from each read using "NGS: QC and manipulation → Trim sequences".
9. Upload a text file with the exact sequence of the barcode sequence (positions 6–12 on the RT primers) using "Get data → Upload file" (*see* **Note 45**).
10. Using the barcode text file and the output from **step 5** above, split reads based on barcode using "NGS: QC and manipulation → Barcode splitter". Specify that the barcodes are at the start of the sequence (*see* **Note 46**).
11. Sequences with or without barcode sequences are output separately in an html file. Convert the output file of barcode-containing sequences into a fasta file for subsequent analysis as follows. In the "Results" pane, click on the "eye" icon to view the html barcode splitter output table. For each individual barcode-containing sequence file, right click on the hyperlink to capture the "Copy Link Location" URL of the file. Go to the tool "Get Data → Upload File" and paste the URL into the " URL/Text:" box. Submit and the data will load as a fasta format dataset.
12. Remove first seven nucleotides (barcode + two invariable Cs) from each read using "NGS: QC and manipulation → Trim sequences" by keeping bases 8 through 50.
13. Remove the 3′-adapter using "NGS: QC and manipulation → Clip" tool. Indicate minimum sequence length after clipping as "1" and supply miRCat-33® sequence as a custom sequence to be clipped.
14. Compute lengths of reads using the tool "FASTA manipulation → Compute sequence length".

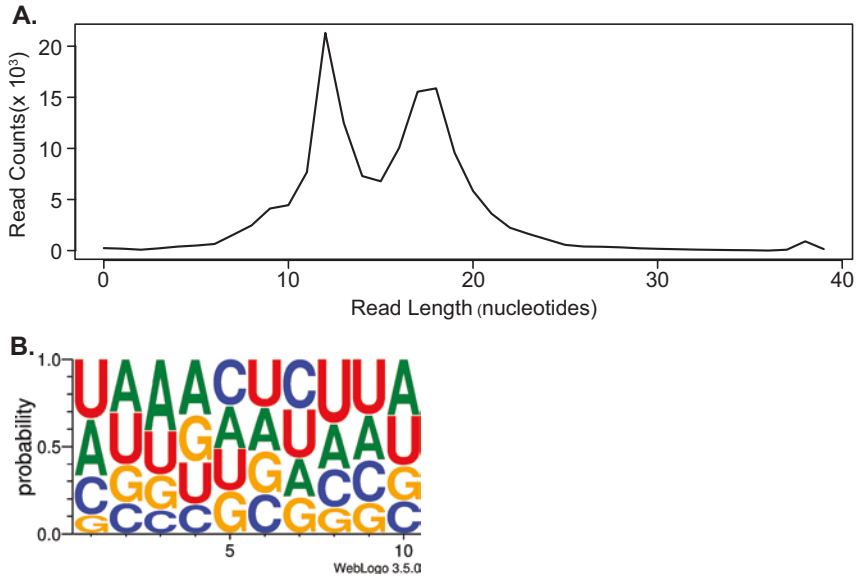


Fig. 4 Size and nucleotide composition of KpAGO-bound RNAs from NGS of the cDNA library. (**a**) A histogram showing the read counts (*y*-axis) corresponding to different lengths (*x*-axis) in the KpAGO cDNA library. ~1.39 million reads are represented in the plot. (**b**) A web-logo showing the probability distribution (*y*-axis) of each of the four nucleotide bases at positions 1–10 (*x*-axis) of the KpAGO-bound RNAs. ~1.29 million reads ≥10 nt were used for this analysis

15. Plot histogram of read lengths using "Graph/Display Data → Histogram". Figure 4a shows the read length profile for the input sample shown in Fig. 2a.
16. To determine sequence composition of first ten positions of the inserts, first use "NGS: QC and manipulation → Trim sequences" tool to keep bases 1 through 10. Then use the tool "Motif Tools → Sequence Logo" to generate a web-logo. Figure 4b shows the web-logo for the first ten positions of the KpAGO-bound RNAs (*see* **Note 47**).
17. The trimmed sequences obtained in **step 10** above can be used for alignment to reference genome using tools available in "NGS: Mapping" tool-shed.

4 Notes

1. All reactions should preferably be performed in a thermocycler. For all reactions above 25 °C, use the heated lid at 110 °C.
2. To keep samples as free from contaminants as possible, all pipetting should be done with filter tips.
3. All homemade buffers should be filter sterilized before use to remove contaminants.

4. Pipet very slowly to accurately pipet viscous solutions containing PEG8000. PEG8000 is susceptible to oxidation in the presence of air. Every time a new vial is opened, make small aliquots and store at −20 °C. Dispose of an aliquot after it has been used a couple times.
5. All primers should be gel purified or HPLC purified.
6. This linker is no longer available from IDT in pre-synthesized form. It can be ordered as a custom oligo from any oligo synthesis service, preadenylated using Mth RNA ligase (NEB) and gel purified.
7. If input RNA requires only 3′-end dephosphorylation, it can be achieved using the 3′-phosphatase activity of T4 PNK in a reaction where RNA is incubated with the enzyme in the absence of ATP.
8. Do not use heat to inactivate CIP as RNA is susceptible to hydrolysis at elevated temperature, particularly in the presence of metal ions present in CIP buffer (Mg^{2+} and Zn^{2+}). Instead, quench reaction with EDTA and remove CIP by Phenol:Chloroform extraction as described.
9. Recovering only 350 μl of the aqueous phase prevents contamination by interphase/organic phase constituents.
10. The phosphorylation steps using T4 PNK (**steps 13–27** in Subheading 3.1) should be performed only with a fraction of the input RNA and most of the RNA should be saved for input into **step 1** of Subheading 3.2.
11. Allow the gel to polymerize for at least 1 h at room temperature.
12. Before and after the pre-run, flush wells with running buffer using a syringe fitted with a 22-gauge needle. Clamp metal heat sinks onto the plates during the pre-run and the final run to prevent cracking of plates by overheating.
13. It is best to skip at least one lane between each sample and RNA/DNA standards.
14. The bottom chamber buffer will be significantly radioactive as unincorporated radioactive ATP will run-off into it from the gel.
15. If the signal is strong and only a few hours of exposure is needed, this step can be skipped and a saran-wrapped gel can be exposed to the phosphorimager screen.
16. Broad wells help reduce the height of loaded sample in the well and reduce sample streaking. Those made using the described 8-well comb on 20 cm wide gel makes 1.7 cm wide well.
17. Before and after pre-run, flush wells before loading and running sample. Clamp metal heat sinks onto the plates during the pre-run and the final run to prevent cracking of plates by overheating.

18. The unextended RT primer migrates ~120 nt. Depending on the size of starting RNA sample, the extended RT product will appear above the unextended primer as a specific size band (in the case of small RNAs) or as a streak (in the case of RNAs of variable length).
19. A standard sheet protector open from three sides allows easy handling of stained PAGE gels without contributing any background fluorescence.
20. If the input is on the lower end of the recommended amount, the RT product may be hard to visualize. In such a case, adjustment of brightness/contrast on images from fluorescent scanner may help to locate the product. In any case, a print copy of the gel image can serve as a guide for excision of the gel region where the RT product is expected to migrate. Numerous times we have successfully made NGS libraries even when RT product is not readily visible.
21. Use a new razor for each sample.
22. The volume of DNA elution buffer used may have to be adjusted to ensure that the gel fragment slurry is mixing well on the nutator. We typically start with 800 μl and add more elution buffer if needed. If gel fragment is too large, split into multiple tubes and add sufficient elution buffer to each tube to allow thorough mixing.
23. For multiple samples, enough beads for $n + 1$ samples can be washed in the same tube using larger volumes of wash buffer. Washed beads can then be split into aliquots.
24. To easily pipet gel slurry, cut a 1000 μl pipet tip about 0.5 cm from the tip to widen its bore. Also, more than one Spin-X column may be needed if the slurry volume is large.
25. Pre-washed streptavidin magnetic beads can also be incubated with gel pieces in elution buffer overnight during **step 2** above. In this case, DO NOT crush gel pieces in **step 1** in Subheading 3.5 as beads tend to stick to the jagged edges of gel fragments.
26. Mix the circularization reaction well by pipetting before adding to PCR reaction.
27. Do NOT exceed the suggested amount of circularization reaction as a template in PCR reaction as it may affect PCR conditions.
28. To test a wider range of PCR cycles, reaction can be split into four equal parts.
29. Use 18-well comb or one that provides several narrow wells per gel.
30. Run a parallel negative control PCR without a circularized template to use as a comparison for size and starting amount of PCR primers.

31. The PCR product from unextended but circularized RT is expected to be 151 bp. Hence, insert containing PCR product will be equivalent to the sum of 151 and expected insert size in bp.
32. Under PCR over-amplification conditions, both depletion of primers and appearance of aberrant slow migrating PCR products can be seen readily; compare lanes 2 and 3 with lane 4 in Fig. 3a for the primer depletion and over-amplification artifacts.
33. Use a wide well comb so as to fit all the 90 μl of the PCRs in a single well sized about 1.7 cm. It is extremely important to perform electrophoresis under conditions that minimize heat generation. Heat produced during electrophoresis can lead to denaturation of double-stranded DNA library, and confound DNA quantifications.
34. Scanning the gel provides documentation for records and enables careful analysis of PCR product sizes before proceeding ahead with next-generation sequencing.
35. Use of blue-light transilluminator minimizes DNA damage that may be caused by short-wave UV.
36. As the volume of gel pieces is similar to those in steps in Subheading 3.5, **steps 1** and **2** of Subheading 3.5 can be followed exactly.
37. Air drying of DNA pellet for even a short duration can lead to DNA denaturation and confound precise quantification in subsequent steps.
38. Glycogen added during DNA precipitation may interfere with sequencing reaction. To remove glycogen completely, purify the DNA using Zymo Research DNA clean & concentrator™ columns as described by the manufacturer.
39. A rough conversion factor for 200 bp DNA product is 7.58 [(ng/μl) × 7.58 = nM]. Use the following conversion formula for a more precise calculation:

nM = ((QUBIT in ng per μl) × 1,000,000)/(((BP SIZE) × 607.4) + 157.9)

40. Directly consult with the NGS sequencing facility regarding volume and concentration of DNA needed for sequencing. Aim for 1.5–2 times the required volume at the given concentration so that additional pooled sample is available for quantification via Bioanalyzer and Qubit, if necessary.
41. DNA from multiple barcoded libraries can be pooled and cloned together.
42. To screen n colonies, prepare a master-mix for $n + 1$ samples.

43. As seen in the FastQC output, quality score >28 will keep only the highest quality reads.
44. There are two tools with the same name in the galaxy toolshed. We recommend to use the one developed by the Galaxy team.
45. The required format of the text file is described in the galaxy tool "NGS: QC and manipulation → Barcode splitter".
46. The number of mismatches allowed and the number of deletions allowed in barcodes can be changed to optimize the output.
47. Sequence Logo tool requires all input sequences to be of the same length.

Acknowledgements

Our research is supported by start-up funds from The Ohio State University (to G.S. and K.N.), a seed-grant from the Center for RNA Biology, OSU (to G.S. and K.N.), a Center for RNA Biology Fellowship, OSU (to D.M.D.) and a Graduate Student Pelotonia Fellowship (to D.M.D). We acknowledge Erin Heyer and Melissa Moore from University of Massachusetts Medical School, Worcester for their critical insights to streamline this procedure in our laboratory.

References

1. Hirose T, Mishima Y, Tomari Y (2014) Elements and machinery of non-coding RNAs: toward their taxonomy. EMBO Rep 15(5):489–507. doi:10.1002/embr.201338390
2. Chu CY, Rana TM (2007) Small RNAs: regulators and guardians of the genome. J Cell Physiol 213(2):412–419. doi:10.1002/jcp.21230
3. Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10(2):126–139. doi:10.1038/nrm2632
4. Nakanishi K (2016) Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? Wiley Interdiscip Rev RNA 7(5):637–660. doi:10.1002/wrna.1356
5. Liu XS, Fan BY, Pan WL, Li C, Levin AM, Wang X,Zhang RL, Zervos TM, Hu J, Zhang XM, Chopp M, Zhang ZG (2017) Identification of miRNomes associated with adult neurogenesis after stroke using Argonaute 2-based RNA sequencing. RNA Biol 14(5):488–499. doi:10.1080/15476286.2016.1196320
6. Spornraft M, Kirchner B, Haase B, Benes V, Pfaffl MW, Riedmaier I (2014) Optimization of extraction of circulating RNAs from plasma-enabling small RNA sequencing. PLoS One 9(9):e107259. doi:10.1371/journal.pone.0107259
7. Vongrad V, Imig J, Mohammadi P, Kishore S,Jaskiewicz L, Hall J, Gunthard HF, Beerenwinkel N, Metzner KJ (2015) HIV-1 RNAs are not part of the argonaute 2 associated RNA interference pathway in macrophages. PLoS One 10(7):e0132127. doi:10.1371/journal.pone.0132127
8. Wang T, Li R, Wen L, Fu D, Zhu B, Luo Y, Zhu H (2015) Functional analysis and RNA sequencing indicate the regulatory role of argonaute 1 in tomato compound leaf development. PLoS One 10(10):e0140756. doi:10.1371/journal.pone.0140756
9. Zhao J, Luo R, Xu X, Zou Y, Zhang Q, Pan W (2015) High-throughput sequencing of RNAs isolated by cross-linking immunoprecipitation

(HITS-CLIP) reveals Argonaute-associated microRNAs and targets in Schistosoma japonicum. Parasit Vectors 8:589. doi:10.1186/s13071-015-1203-9

10. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460(7254):479–486. doi:10.1038/nature08170
11. Nakanishi K, Weinberg DE, Bartel DP, Patel DJ (2012) Structure of yeast Argonaute with guide RNA. Nature 486(7403):368–374. doi:10.1038/nature11211
12. Heyer EE, Ozadam H, Ricci EP, Cenik C, Moore MJ (2015) An optimized kit-free method for making strand-specific deep sequencing libraries from RNA fragments. Nucleic Acids Res 43(1):e2. doi:10.1093/nar/gku1235
13. Sterling CH, Veksler-Lublinsky I, Ambros V (2015) An efficient and sensitive method for preparing cDNA libraries from scarce biological samples. Nucleic Acids Res 43(1):e1. doi:10.1093/nar/gku637
14. Singh G, Kucukural A, Cenik C, Leszyk JD, Shaffer SA, Weng Z, Moore MJ (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell 151(4):750–764. doi:10.1016/j.cell.2012.10.007
15. Wery M, Descrimes M, Thermes C, Gautheret D, Morillon A (2013) Zinc-mediated RNA fragmentation allows robust transcript reassembly upon whole transcriptome RNA-Seq. Methods 63(1):25–31. doi:10.1016/j.ymeth.2013.03.009
16. Nakanishi K, Ascano M, Gogakos T, Ishibe-Murakami S, Serganov AA, Briskin D, Morozov P, Tuschl T, Patel DJ (2013) Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. Cell Rep 3(6):1893–1900. doi:10.1016/j.celrep.2013.06.010
17. Schirle NT, MacRae IJ (2012) The crystal structure of human Argonaute2. Science 336(6084):1037–1040. doi:10.1126/science.1221551
18. Elkayam E, Kuhn CD, Tocilj A, Haase AD, Greene EM, Hannon GJ, Joshua-Tor L (2012) The structure of human argonaute-2 in complex with miR-20a. Cell 150(1):100–110. doi:10.1016/j.cell.2012.05.017
19. Faehnle CR, Elkayam E, Haase AD, Hannon GJ, Joshua-Tor L (2013) The making of a slicer: activation of human Argonaute-1. Cell Rep 3(6):1901–1909. doi:10.1016/j.celrep.2013.05.033
20. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44:W3–W10. doi:10.1093/nar/gkw343

Chapter 2

Quantification of miRNAs Co-Immunoprecipitated with Argonaute Proteins Using SYBR Green-Based qRT-PCR

Hong-Duc Phan, Junan Li, Ming Poi, and Kotaro Nakanishi

Abstract

MicroRNAs (miRNAs) are small non-coding RNAs that trigger post-transcriptional gene silencing. These RNAs need to be associated with the Argonaute proteins to be functional. This assembly begins with loading of a miRNA duplex, followed by the ejection of one of the strands (passenger). The remaining strand (guide) together with the Argonaute protein forms a ribonucleoprotein effector complex (the RNA-induced silencing complex, RISC). Mutation on the Argonaute protein, if affecting either step of the RISC assembly, impacts the function of miRNAs. Therefore, any observation of decreased miRNA level of mutants will provide insights into the role of those amino acid residues in the mechanical function of the Argonaute protein. In this chapter, we introduce a method to relatively quantify a specific miRNA co-immunoprecipitated with wild type and mutant Argonaute proteins from HEK293T cells, using Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR). Spiking a synthetic exogenous miRNA as an internal control with RNA extraction prior to cDNA synthesis will normalize the C_t values obtained from the qRT-PCR assays and enable us to quantify the relative level of Argonaute-bound miRNA.

Key words qRT-PCR, miRNA, Argonaute, Immunoprecipitation, SYBR Green

1 Introduction

miRNAs are single-stranded RNAs of about 22 nucleotides (nt) that are generated from the endogenous hairpin-shaped transcript [1, 2]. In human, miRNA duplexes are loaded into four Argonaute paralogs (Ago1-4) [1, 3]. After passenger-strand ejection, the remaining guide and the Argonaute protein form the RNA-induced silencing complex (RISC) [4, 5]. This ribonucleoprotein complex also provides a scaffold for the components essential for translational repression and deadenylation [3, 4, 6]. To understand the versatile functions of Argonaute proteins, previous studies compared the physiological activities between wild type (WT) and the mutant proteins. For example, FLAG-tagged human Argonaute (FLAG-Ago) proteins were expressed in HEK293T cells and

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_2,

purified by immunoprecipitation with anti-FLAG antibody from the cell lysate [7]. The miRNA-binding activities of Argonaute mutants were evaluated by Northern blot using a probe for a particular miRNA [7–9]. This method has been routinely used to identify residues on Argonaute proteins playing important role(s) during loading of miRNA and/or passenger-strand ejection [10–12]. Despite its high specificity, Northern blot has some disadvantages, such as being time-consuming, low throughput, and low sensitivity for small RNA detection [13–15].

Recently, Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) is becoming more popular as an alternative method to Northern blot. However, there are two major obstacles preventing the use of qRT-PCR in quantifying miRNA co-immunoprecipitated with FLAG-Ago. First is the primer design for qRT-PCR. Historically, qRT-PCR has been developed for detection of mRNA; therefore, the techniques and knowledge based on mRNA detection may not be applicable for miRNA, primarily due to the much shorter length. The use of stem-loop qRT-PCR may partially solve the problems, but this method requires the use of stem-loop reverse transcription primers (also known as TaqMan® MicroRNA assay), which has been licensed to Applied Biosystems™ [16, 17]. Another obstacle is that a reference gene such as a housekeeping gene needs to be used as an internal control to normalize variances between samples [18]. Previous studies often used fortuitously bound RNA such as GAPDH as a reference gene for the normalization of the C_t value when employing qRT-PCR to detect a specific mRNA co-immunoprecipitated with Argonaute proteins [19–21]. This approach may introduce considerable variance in the quantification of miRNA co-immunoprecipitated with Argonaute mutants when the mutation affects the binding of such RNAs.

To address the issue of primer design, we employed SYBR Green-based qRT-PCR with polyT adaptor primers (Fig. 1) [22]. This approach is cost-effective, especially in the case of systematic quantification of different miRNAs co-immunoprecipitated with a large number of Argonaute protein mutants. In this method, miRNAs will be first polyadenylated by poly(A) polymerase (Fig. 1, step 1), prior to the complementary DNA (cDNA) synthesis by reverse transcription using an oligo-dT adaptor primer composed of a polyT primer and a universal tag on its 5′ side (Fig. 1, step 2). Then, an equal amount of an exogenous miRNA, cel-miR-39, was spiked as the internal control into the immunoprecipitated Argonaute proteins whose amount was adjusted based on western blot analysis before RNA extraction (Fig. 2) [23]. The difference in threshold cycles for the miRNA of interest and the spiked cel-miR-39 enables us to normalize RNA extraction from the different immunoprecipitated proteins.

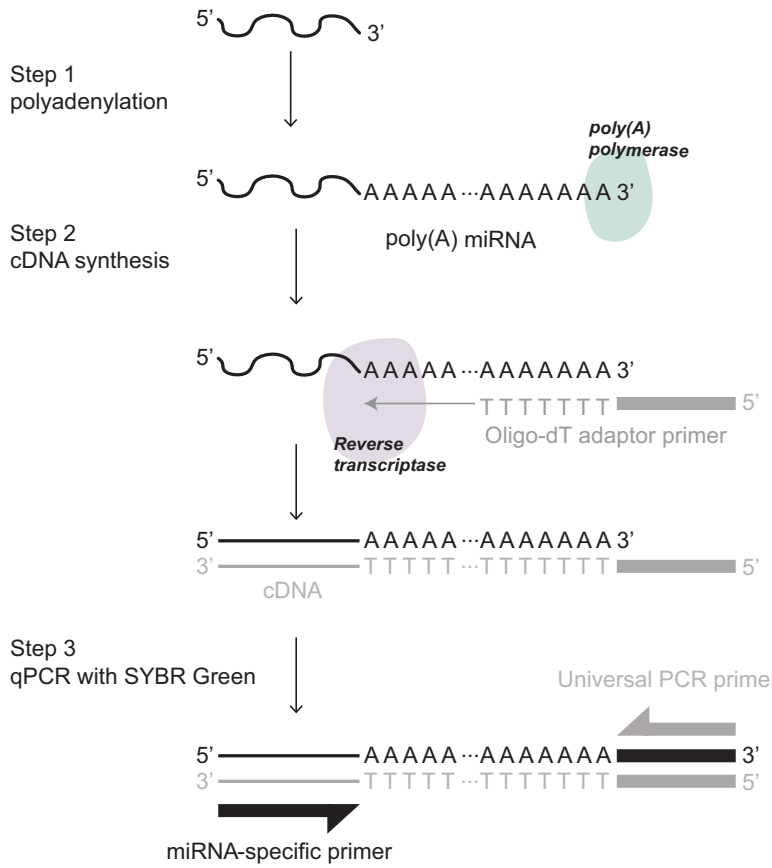


Fig. 1 Schematic representation of polyT adaptor qRT-PCR method. Step 1: poly(A) tail is added to miRNA by poly(A) polymerase. Step 2: cDNA is synthesized by reverse transcriptase using oligo-dT adaptor primer. Step 3: cDNA is amplified by SYBR Green-based qPCR with miRNA-specific and universal PCR primers

2 Materials

1. Dynabeads Protein G (Thermo Fisher Scientific).
2. Magnetic rack, DynaMag-2 Magnet (Thermo Fisher Scientific) or equivalent.
3. Ultrasonic processor.
4. Bench top tube rotator.
5. Vortex mixer.
6. Anti-FLAG M2 antibody, F1804 (Sigma). The antibody is diluted 1000 times with the blotting buffer containing 5% of skim milk and 1% of sodium azide as working solution for Western blot.
7. Lysis buffer: 30 mM HEPES-KOH (pH 7.4), 100 mM KOAc, 2 mM $Mg(OAc)_2$, 10 μg/μl Aprotinin*, 10 μM Leupeptin*,

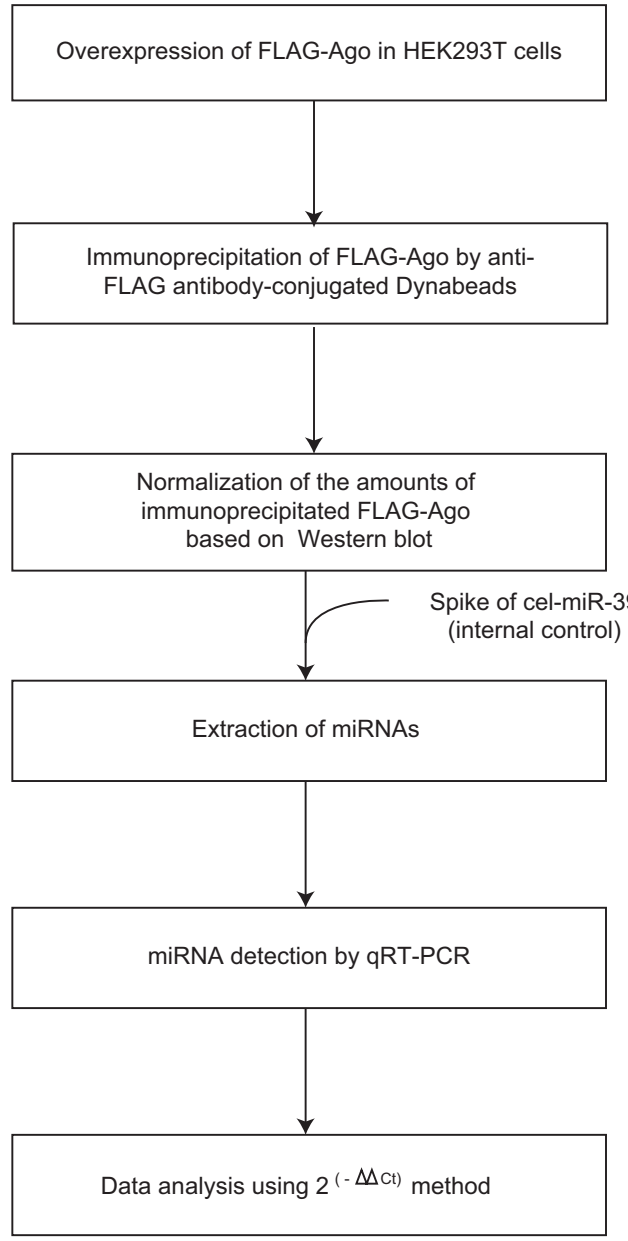


Fig. 2 Workflow of SYBR Green-based qRT-PCR of miRNAs using an exogenous miRNA as an internal control

1 μM Pepstatin*****, and 1 mM phenylmethylsulfonyl fluoride (PMSF)***** (*****: freshly prepared).

8. Wash buffer: 30 mM HEPES-KOH (pH 7.4), 100 mM KOAc, 2 mM $Mg(OAc)_2$, 800 mM NaCl, and 1% Triton X-100.
9. Cel-miR-39 (GE Dharmacon): 5′ UCACCGGGUGUAAAU CAGCUUG 3′.

10. Proteinase K reaction buffer: 200 mM Tris–HCl (pH 7.5), 25 mM EDTA (pH 8.0), 300 mM NaCl, and 2% sodium dodecyl sulfate (SDS) (w/v).
11. Proteinase K, 20 mg/ml.
12. TRIzol LS reagent (Thermo Fisher Scientific).
13. Direct-zol RNA kits (Zymo Research).
14. TransIT-X2 Dynamic Delivery System (Mirus Bio).
15. qScript microRNA cDNA Synthesis Kit (Quanta Bio).
16. PerfeCTa SYBR Green SuperMix (Quanta Bio)
17. Universal PCR primer (Quanta Bio).
18. Forward primer to detect miR-19b and cel-miR-39 (ordered from Quanta Bio) (*see* **Note 1**).

Cel-miR-39	GCAGCTGATTTCGTCTTGGTAA
miR-19b	GTGCAAATCCATGCAAAACTG

19. 2× SDS-loading dye: 125 mM Tris–HCl (pH 6.8), 4% Sodium dodecyl sulfate (SDS), 20% Glycerol, 0.04% Bromophenol blue, and 100 mM beta-mercaptoethanol* (*freshly prepared).
20. Opti-MEM reduced serum media (Gibco).
21. Nitrocellulose blotting membrane (GE Healthcare).
22. Electrophoresis chamber (Bio-RAD).
23. Blotting buffer: 20 mM Tris-base, 150 mM NaCl, and 0.001% Tween 20.
24. Blocking buffer: 20 mM Tris-base, 150 mM NaCl, 0.001% Tween with 10% (w/v) skim milk.
25. Novex NuPAGE SDS-PAGE gel (Invitrogen).
26. Image StudioLite (Li-cor).
27. Odyssey® Imaging Systems (Li-cor).
28. IRDye® 800CW Goat anti-Mouse IgG (H + L), 0.1 mg (fluorescent-labeled secondary antibody, Li-cor). The antibody is diluted at 1:16,000 with dilution buffer containing 50% PBS (pH 7.4), 50% Odyssey buffer and 0.01% Tween as working solution for Western blot.
29. pCAGEN-FLAG-Ago2 WT [24].
30. pCAGEN-FLAG-Ago2 Y529E is made using the following forward and reverse site-directed mutagenesis primers: Forward: 5′-CGGCAAGACGCCCGTGGAAGCCGAG-3′, Reverse: 5′-ACGCGCTTGACCTCGGCTTCCACGGGC-3′.
31. The gene of MBP is amplified from pMAL vector (Novagen) with the following forward and reverse primers and processed with EcoRI and NotI. Then, the fragment is cloned into

pCAGEN vector using the same restriction enzyme sites. Forward: 5′-GAGTGAATTCATGAAAATCGAAGAAGGTAAACTGGTAAT-3′, Reverse: 5′-CACTGCGGCCGCTCAATTAGTCTGCGCGGCTGC-3′ (EcoRI and NotI sites are underscored).

3 Methods

We quantified miR-19b, one of the most abundant miRNAs in HEK293T cells [25] and compare the miRNA-binding activities between the FLAG-Ago2 WT and Y529E mutant. The latter was previously reported as a mutant deficient in miRNA binding by Northern blot analysis [12]. Non-tagged MBP was used as a negative control.

3.1 Overexpression of MBP, FLAG-Ago2 WT, and FLAG-Ago2Y529E Proteins in HEK293T Cells

1. Transfect the mixture of 10 μg plasmids, 45 μl of TransIT-X2 reagent, and 1.5 ml OMEM into the 10 cm plates of 80% confluent HEK293T cells, according to the manufacturer manual's instruction. Cultivate the cells for 48–50 h.
2. After rinsing the cells once with 10 ml cold PBS, detach the cell off the plates by vigorously pipetting the cells with PBS. Harvest the cell by centrifugation at 1000 × *g* for 10 min at 4 °C. Measure the weight of the cell pellet (*x* mg). Hereafter, handle the samples at 4 °C through the whole procedure.
3. Resuspend the cell pellet with five times of the pellet weight (5× μl) of lysis buffer in a 1.7 ml microcentrifuge tube. Place the tube in an ice bucket and sonicate at 30% amplitude using Microtip for 20 s in 4 s burst with at least 60-s intervals. Incubate on ice for 5 min, and centrifuge at 17,000 × *g* for 20 min at 4 °C. Collect the supernatant and aliquot into 100 μl fraction in microcentrifuge tubes. Then, flash-freeze the lysate in liquid nitrogen and store them at −80 °C.

3.2 FLAG-Ago Immunoprecipitation Using Anti-FLAG M2 Antibody

1. For each FLAG-Ago sample, aliquot 60 μl of Dynabeads Protein G into a 1.7 ml microcentrifuge tube and resuspend with 1 ml of Lysis buffer. Set the tubes on the magnetic rack and remove the solution. Repeat this wash process three times.
2. Resuspend the beads in 60 μl of Lysis buffer, followed by mixing with 2 μl of 1 mg/ml anti-FLAG M2 antibody. Shake the tube on the vortex mixer at room temperature for 1 h.
3. Resuspend the antibody-conjugated beads with 1 ml of Lysis buffer. Set the tube on the magnetic rack and remove the solution. After repeating this wash process three times, discard the supernatant.

4. Quickly thaw the frozen 100 μl-cell lysate aliquot from **step 3** of Subheading 3.1, and incubate with the antibody-conjugated beads at 4 °C for 2 h on a benchtop tube rotator.
5. Add 1 ml of Wash buffer to the beads and resuspend the mixture. Set the tubes on the magnetic rack and discard the supernatant. Repeat this wash process five times.

3.3 Semi-quantification of Immunoprecipitated Protein by Western Blot

We followed the standard western blot protocol [26] with some modifications.

1. Take 3 μl of the beads from **step 5** in Subheading 3.2 and mix with 2 μl of 2× SDS-loading dye and 7 μl of water to the final volume of 12 μl. Heat the sample at 90 °C for 5 min.
2. After spinning down the sample, load the supernatant on the SDS-PAGE gel. Run the gel at a constant voltage of 180 V until the bromophenol blue reaches to the bottom.
3. Transfer the sample from the gel to the nitrocellulose blotting membrane in the electrophoresis chamber at a constant voltage of 80 V for 1.5 h at 4 °C.
4. Block the membrane with 20 ml of blocking buffer for 1 h.
5. Incubate the membrane for 2 h at room temperature with 10 ml of anti-FLAG M2 antibody working solution.
6. After discarding the antibody, add 20 ml of blotting buffer to the membrane and incubate them for 10 min. Repeat this wash process three times.
7. Incubate the membrane for 1 h with 40 ml of secondary antibody working solution.
8. After discarding the second antibody, add 20 ml of Blotting buffer to the membrane, and incubate them for 10 min. Repeat this wash process three times.
9. Fluorescence from the fluorescent-labeled secondary antibody is detected in the 800 channel of Odyssey Imaging System.
10. Quantify the fluorescent intensity from the immunoprecipitated Ago protein using Image StudioLite software (Fig. 3).

3.4 Extraction of RNA from Immunoprecipitated FLAG-Ago (see Note 2)

1. Adjust the amount of beads based on the intensities of immunoprecipitated Ago of the Western blot data from **step 10** in Subheading 3.3.
2. Add 250 μl of proteinase K buffer and 10 μl of 20 mg/ml proteinase K to the protein immunoprecipitated beads. Incubate the mixture at 50 °C for 30 min.
3. Spike 1 μl of 10 nM of cel-miR-39 into each sample as an internal control (*see* **Note 3**).

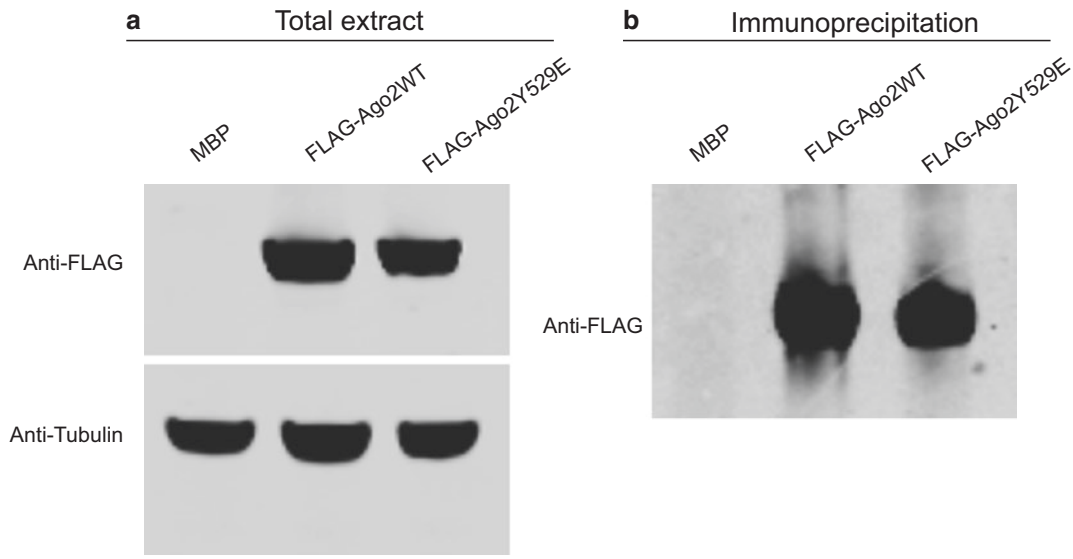


Fig. 3 Expression and immunoprecipitation of FLAG-Ago2 WT and Y529E. (**a**) Expression levels of FLAG-Ago2 WT and Y529E were detected by anti-FLAG antibody. Anti-tubulin antibody was used to detect tubulin as an internal control. Non-tagged MBP was used as a negative control. Western blot analysis was triplicated. (**b**) Immunoprecipitated FLAG-Ago2 WT and Y529E were detected by anti-FLAG antibody

4. Add 200 μl of TRIzol LS reagent to each sample and vortex the mixture for 30 s, followed by incubation at room temperature for 5 min. Centrifuge at 17,000 × *g* for 10 min to remove the beads and debris. Transfer the supernatant to a new tube.
5. Mix the supernatant with an equal volume of 100% ethanol and transfer the mixture to the spin column of Direct-zol RNA kits. Hereafter, follow the instruction of the kit to extract the total RNA.
6. Elute the total RNA from the spin column with 20 μl of RNase-free water. The RNA eluents can be stored at −80 °C until they are used for cDNA synthesis step.

3.5 cDNA Preparation

1. Take 7 μl of RNA for the poly(A) tailing reaction, following the qScript microRNA cDNA Synthesis Kit's instructions: incubate the reaction at 37 °C for 1 h, followed by another incubation at 70 °C for 5 min.

Reagents	Volume (μl)
5× Poly(A) tailing buffer	2
RNA	7
Poly(A) polymerase	1

2. Set up the cDNA synthesis reaction, following the qScript microRNA cDNA Synthesis Kit's instructions: incubate the reaction at 42 °C for 20 min, followed by another incubation at 85 °C for 5 min.

Reagents	Volume (μl)
Poly(A) tailing reaction	10
microRNA cDNA reaction mix	9
qScript reverse transcriptase	1

3. Transfer 4 μl of the cDNA product successively to 16 μl of RNase-free water to generate fivefold serial dilution samples and store them at −20 °C.

3.6 cDNA Amplification by SYBR Green qPCR

1. Set up the qPCR reaction to amplify miR-19b as follows:

Reagents	Volume (μl)
2× SYBR Green premix	5
10 μM miR-19b primer	0.2
10 μM Universal PCR primer	0.2
cDNA (from **step 3** in Subheading 3.5)	4
Double distilled water	0.6

2. Set up the qPCR reaction to amplify cel-miR-39 as follows:

Reagents	Volume (μl)
2× SYBR Green premix	5
10 μM cel-miR-39 primer	0.2
10 μM Universal PCR primer	0.2
cDNA (from **step 3** in Subheading 3.5)	4
Double distilled water	0.6

3. Run the PCR as follows:
   Pre-incubation (1 cycle): 95 °C for 2 min.
   PCR (40 cycles):
   Denaturation: 95 °C for 5 s.
   Annealing: 60 °C for 30 s (to collect fluorescence signals).

3.7 Data Analysis

The miRNA-binding activity of Ago2 Y529E mutant was compared to that of WT by quantifying their bound miR-19b (*see* **Note 4**).

1. The threshold cycle for miR-19b co-immunoprecipitated with FLAG-Ago2 WT is normalized to the spiked cel-miR-39 using the following equation:

$$\Delta C_{\mathrm{t,WT}} = C_{\mathrm{t,WT}}\left(\text{miR-19b}\right) - C_{\mathrm{t,WT}}\left(\text{cel-miR-39}\right)$$

   where $C_{\mathrm{t,WT}}$ (miR-19b) and $C_{\mathrm{t,WT}}$ (cel-miR-39) are the threshold cycles for miR-19b and cel-miR-39 extracted from FLAG-Ago2WT sample, respectively.

2. The threshold cycle for miR-19b co-immunoprecipitated with FLAG-Ago2 Y529E is normalized to the spiked cel-miR-39 using the following equation:

$$\Delta C_{\mathrm{t,Y529E}} = C_{\mathrm{t,Y529E}}\left(\text{miR-19b}\right) - C_{\mathrm{t,Y529E}}\left(\text{cel-miR-39}\right)$$

   where $C_{\mathrm{t,\ Y529E}}$ (miR-19b) and $C_{\mathrm{t,\ Y529E}}$ (cel-miR-39) are the threshold cycles for miR-19b and cel-miR-39 extracted from FLAG-Ago2 Y529E sample, respectively.

3. The relative amount of miR-19b co-immunoprecipitated with FLAG-Ago2 Y529E to that with FLAG-Ago2 WT is given by

$$\frac{X_{\mathrm{N,Y529E}}}{X_{\mathrm{N,WT}}} = \frac{2^{-\Delta C_{\mathrm{t,Y529E}}}}{2^{-\Delta C_{\mathrm{t,WT}}}} = 2^{-\left(\Delta C_{\mathrm{t,Y529E}} - \Delta C_{\mathrm{t,WT}}\right)} = 2^{-\Delta\Delta C_{\mathrm{t}}}$$

   where $X_{\mathrm{N,WT}}$ and $X_{\mathrm{N,Y529E}}$ are the normalized initial amounts of miR-19b co-immunoprecipitated with FLAG-Ago2 WT and Y529E mutant, respectively, and $\Delta\Delta C_{\mathrm{t}} = \Delta C_{\mathrm{t,Y529E}} - \Delta C_{\mathrm{t,WT}}$ [27].

4. $X_{\mathrm{N,Y529E}}/X_{\mathrm{N,WT}}$ is shown in bar graph with two-sided Student's *t*-test values calculated (Fig. 4). The relative amount of miR-19 co-immunoprecipitated with MBP to that with FLAG-Ago2 WT (i.e. $X_{\mathrm{N,MBP}}/X_{\mathrm{N,WT}}$) is calculated in the same way.

4 Notes

1. A list of pre-designed primers to detect miRNAs has been validated and published on the website of the manufacturer (http://www.quantabio.com/products/microrna-profiling). For miRNAs without pre-validated primers, a guide note for primer design has also been published (http://www.quantabio.com/products/microrna-profiling).
2. It is important to use RNase-free materials in Subheadings 3.4–3.6.

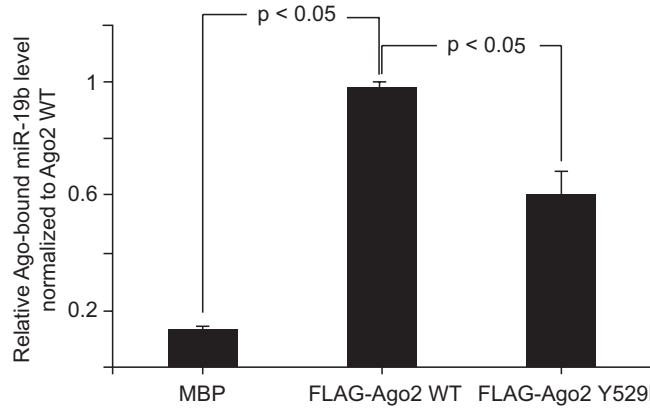


Fig. 4 The amount of miR-19b bound to FLAG-Ago2 Y529E and MBP relative to FLAG-Ago2 WT. Relative quantification of miR-19b level in FLAG-AGo2 Y529E, WT and MBP. FLAG-Ago2 Y529E bound less amount of miR-19b than FLAG-Ago2 WT ($p < 0.05$), which is consistent with the previous report [12]. Experiments were conducted three times

3. Alternative exogenous miRNAs, which are not expressed in the sample, can be spiked in the sample as the internal control. The level of exogenous spiked miRNAs should be in the linear range of C_t value of the miRNA of interest.
4. This method is applicable for detecting any miRNAs by changing the specific miRNA primers.

Acknowledgments

We thank G. Singh for providing HEK293T cells and reagents for us. This work was supported by the PRESTO from the Japan Science and Technology (JST) Agency (JPMJPR13L7), The Ohio State University Start-up Fund, and The Ohio State University Center for RNA Biology Seed Grant to K.N.

References

1. Ha M, Kim VN (2014) Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15(8):509–524. doi:10.1038/nrm3838
2. Chong MM, Zhang G, Cheloufi S, Neubert TA, Hannon GJ, Littman DR (2010) Canonical and alternate functions of the microRNA biogenesis machinery. Genes Dev 24(17): 1951–1960. doi:10.1101/gad.1953310
3. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
4. Nakanishi K (2016) Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? Wiley Interdiscip Rev RNA 7(5): 637–660. doi:10.1002/wrna.1356
5. Kawamata T, Tomari Y (2010) Making RISC. Trends Biochem Sci 35(7):368–376. doi:10.1016/j.tibs.2010.03.009
6. Peters L, Meister G (2007) Argonaute proteins: mediators of RNA silencing. Mol Cell 26(5):611–623. doi:10.1016/j.molcel.2007.05.001

7. Meister G, Landthaler M, Peters L, Chen PY, Urlaub H, Luhrmann R, Tuschl T (2005) Identification of novel argonaute-associated proteins. Curr Biol 15(23):2149–2155. doi:10.1016/j.cub.2005.10.048
8. Hauptmann J, Kater L, Loffler P, Merkl R, Meister G (2014) Generation of catalytic human Ago4 identifies structural elements important for RNA cleavage. RNA 20(10):1532–1538. doi:10.1261/rna.045203.114
9. Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell 15(2):185–197. doi:10.1016/j.molcel.2004.07.007
10. Iwasaki S, Sasaki HM, Sakaguchi Y, Suzuki T, Tadakuma H, Tomari Y (2015) Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature 521(7553):533–536. doi:10.1038/nature14254
11. Gu S, Jin L, Huang Y, Zhang F, Kay MA (2012) Slicing-independent RISC activation requires the argonaute PAZ domain. Curr Biol 22(16):1536–1542. doi:10.1016/j.cub.2012.06.040
12. Rudel S, Wang Y, Lenobel R, Korner R, Hsiao HH, Urlaub H, Patel D, Meister G (2011) Phosphorylation of human Argonaute proteins affects small RNA binding. Nucleic Acids Res 39(6):2330–2343. doi:10.1093/nar/gkq1032
13. Lee RC, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science 294(5543):862–864. doi:10.1126/science.1065329
14. Damm K, Bach S, Muller KM, Klug G, Burenina OY, Kubareva EA, Grunweller A, Hartmann RK (2015) Improved northern blot detection of small RNAs using EDC crosslinking and DNA/LNA probes. Methods Mol Biol 1296:41–51. doi:10.1007/978-1-4939-2547-6_5
15. Beckmann BM, Grunweller A, Weber MH, Hartmann RK (2010) Northern blot detection of endogenous small RNAs (approximately14 nt) in bacterial total RNA extracts. Nucleic Acids Res 38(14):e147. doi:10.1093/nar/gkq437
16. Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS, Chen C (2008) Real-time PCR quantification of precursor and mature microRNA. Methods 44(1):31–38. doi:10.1016/j.ymeth.2007.09.006
17. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 33(20):e179. doi:10.1093/nar/gni178
18. Kozera B, Rapacz M (2013) Reference genes in real-time PCR. J Appl Genet 54(4):391–406. doi:10.1007/s13353-013-0173-x
19. Beitzinger M, Meister G (2011) Experimental identification of microRNA targets by immunoprecipitation of Argonaute protein complexes. Methods Mol Biol 732:153–167. doi:10.1007/978-1-61779-083-6_12
20. Loedige I, Gaidatzis D, Sack R, Meister G, Filipowicz W (2013) The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Res 41(1):518–532. doi:10.1093/nar/gks1032
21. Hassan T, Smith SG, Gaughan K, Oglesby IK, O'Neill S, McElvaney NG, Greene CM (2013) Isolation and identification of cell-specific microRNAs targeting a messenger RNA using a biotinylated anti-sense oligonucleotide capture affinity technique. Nucleic Acids Res 41(6):e71. doi:10.1093/nar/gks1466
22. Shi R, Chiang VL (2005) Facile means for quantifying microRNA expression by real-time PCR. Biotechniques 39(4):519–525
23. Kroh EM, Parkin RK, Mitchell PS, Tewari M (2010) Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods 50(4):298–301. doi:10.1016/j.ymeth.2010.01.032
24. Yoda M, Cifuentes D, Izumi N, Sakaguchi Y, Suzuki T, Giraldez AJ, Tomari Y (2013) Poly(A)-specific ribonuclease mediates 3′-end trimming of Argonaute2-cleaved precursor microRNAs. Cell Rep 5(3):715–726. doi:10.1016/j.celrep.2013.09.029
25. Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T (2008) Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA 14(12):2580–2596. doi:10.1261/rna.1351608
26. Mahmood T, Yang PC (2012) Western blot: technique, theory, and trouble shooting. N Am J Med Sci 4(9):429–434. doi:10.4103/1947-2714.100998
27. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25(4):402–408. doi:10.1006/meth.2001.1262

Chapter 3

Gateway to Understanding Argonaute Loading of Single-Stranded RNAs: Preparation of Deep Sequencing Libraries with In Vitro Loading Samples

Eling Goh and Katsutomo Okamura

Abstract

Identification of sequences preferred by individual RNA-binding proteins (RBPs) has been accelerated by recent advances in the quantitative analysis of protein–RNA interactions on a massive scale, and such experiments have even revealed hidden sequence specificity of RBPs that were assumed to be non-specific. Argonaute (AGO) proteins bind diverse guide small RNAs and were believed to have no sequence specificity besides the preference for particular bases at the 5′ nucleotide. However, we recently showed that short single-stranded RNAs (ssRNAs) are loaded to AGOs in vivo and in cell extracts with detectable sequence preferences. To study the sequence specificity, we established a protocol for preparing the oligo-specific deep-sequencing library. The protocol includes in vitro loading assay that uses RNA oligos containing randomized nucleotides at the first five positions and also splinted-ligation that specifically amplifies the introduced oligo RNA species from a complex mixture of endogenous small RNAs and exogenously introduced RNA oligos. With the current sequencing depth, this procedure will allow quantitative profiling of interactions between the AGO and ~1000 ssRNA species with different sequences. The method would aid in studying the mechanism behind the selective loading of ssRNAs to AGOs and may potentially be applied to study interactions between RNA and other RNA-binding proteins.

Key words Argonaute, Single-stranded RNA, High-throughput sequencing, In-vitro

1 Introduction

RNA interference (RNAi) is a gene-silencing phenomenon triggered by microRNA (miRNA) and small interfering RNA (siRNA) duplexes loaded to Argonaute (AGO) proteins [1]. Upon loading, the passenger strand of each duplex is released, leaving behind just the guide strand, resulting in the formation of RNA-induced silencing complex (RISC) [2]. Strand selection by AGO is actuated by the identity of 5′ terminal nucleotide and asymmetrical thermodynamic stability of the duplex [3]. While much is known about the mechanistic and functional aspect of AGO-duplex loading, loading of single-stranded

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_3,

RNAs (ssRNAs) to AGOs has only recently been brought to attention. Accumulating evidence suggests that a fraction of cellular non-structured ssRNAs such as the loop regions released from miRNA hairpins are loaded to AGOs to form active RISCs [4–6]. Furthermore, with various chemical modifications extending the half-lives, ssRNA oligos successfully exhibited the anticipated RNAi effect in vivo, with better penetration to target tissues compared with the duplex siRNA counterparts [7–11].

We previously proposed that loading of endogenous ssRNAs to fly AGOs is sequence-selective and recapitulated-selective loading by introducing ^{32}P-labeled ssRNA oligos harboring *dme-mir-317* and *dme-mir-34* loop sequences to in-vitro cytoplasmic extracts (Fig. 1a). The preferential loading of AGO1 and AGO2 observed was consistent with that of in-vivo processed *mir-317* and *mir-34* loops (Fig. 1a).

To understand the sequence selectivity of AGOs, an unbiased method to screen for preferentially loaded sequences selected from a large pool of RNA species was needed. The challenge with establishing such a method is to specifically amplify the synthetic RNA oligos that have partially randomized sequence from a complex sample containing vastly diverse endogenous RNAs (Fig. 1b). In this chapter, we describe an approach to specifically amplify a particular RNA species with partially randomized sequences after in-vitro AGO loading for the construction of high-throughput sequencing (HTS) libraries (Fig. 2).

This is achieved through the combination of (1) multiple gel purification steps to eliminate species of undesired sizes after AGO immunoprecipitation and ligation steps (Fig. 3), (2) splinted-ligation to specifically add a 3′-linker to AGO-loaded 5′ end-randomized ssRNA oligo, and (3) oligo-specific RT-PCR to further enrich cDNA molecules derived from the desired RNA species. By optimizing splinted-ligation conditions (Fig. 4), this protocol could achieve strong enrichment of the introduced oligo and depletion of endogenous species, resulting in >80% specificity (Fig. 5). With great sequencing depth and high specificity of the library, it is anticipated that the broad spectrum of AGO sequence preference, if any, would be distinctly revealed.

2 Materials

Items are listed in the order of the procedure.

2.1 ^{32}P ssRNA Oligo Labeling for In Vitro Loading

1. 1 μM partially randomized RNA oligonucleotides, HPLC purified: NNNNNUGAAAUGCAAGCAAG (N: randomized region with hand-mixed nucleotides, underlined: backbone sequence of choice).

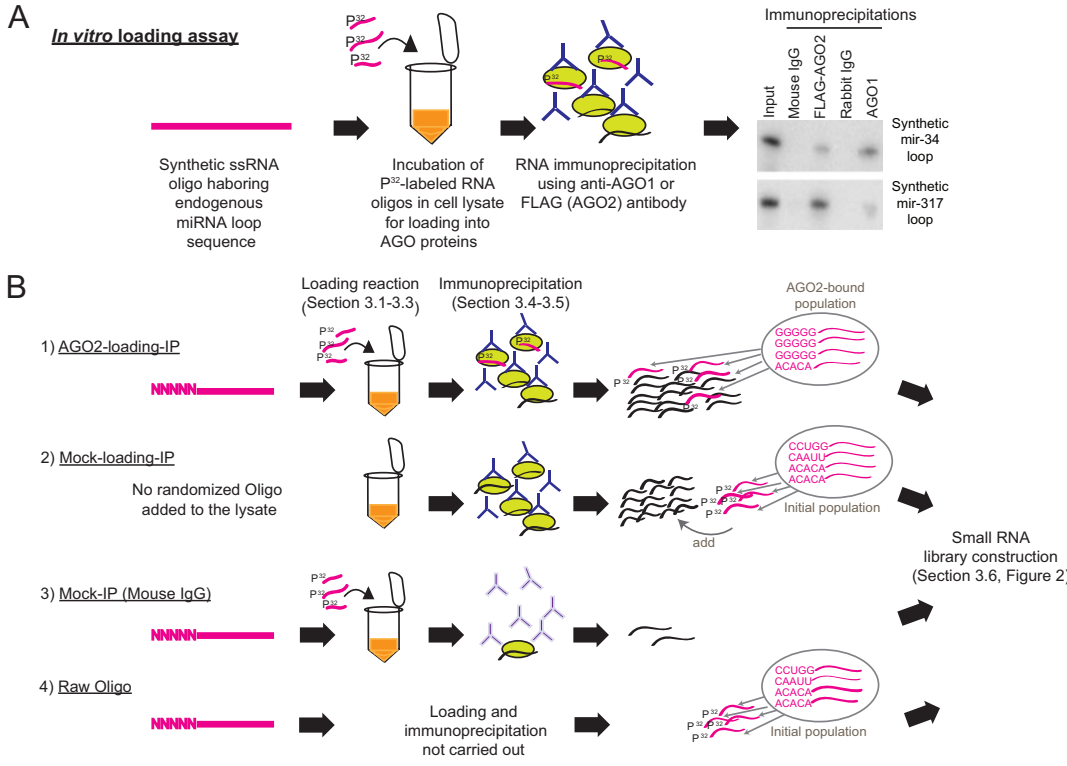


Fig. 1 Applications of ssRNA in vitro loading assay. (**a**) In vitro loading assay recapitulates AGO loading trend of in vivo processed miRNA loops. ^{32}P-labeled ssRNA oligos harboring endogenous miRNA loop sequences were introduced into S2R+/FLAG-HA-AGO2 cell lysate for in vitro AGO loading for an hour. RNA extracted from immunopurified AGO1 and FLAG-AGO2 were resolved on a denaturing gel and the bands were visualized and quantified by phosphorimager. AGO1 prefers *mir-34* loop while AGO2 prefers *mir-317* loop, which are consistent with a previous study [5]. (**b**) In vitro loading of ssRNA oligo coupled with small RNA library construction enables the study of AGO protein sequence selectivity. "AGO2-loading-IP" requires the loading of ^{32}P-labeled ssRNA oligo in the cell lysate, followed by FLAG-AGO2 immunoprecipitation. The resulting extracted RNA pool contains mainly endogenous small RNAs with a minor population of the ^{32}P-labeled ssRNA oligo. To obtain the baseline, the "mock-loading-IP" experiment was performed. This does not involve the loading of the ssRNA oligo, but instead, the ^{32}P-labeled ssRNA oligo is only added after the RNA is extracted from AGO2-IP sample. Therefore, the composition of the initial population used for the loading could be determined by analyzing the "mock loading" library. Enrichment/depletion in the AGO2-bound population will be determined by the comparison of the relative abundances between the "AGO2-loading-IP" and "Mock-loading-IP" libraries. In this example, 5′ GGGGG sequence is highly enriched in "AGO2-loading-IP" since the initial population as reflected in "Mock-loading-IP" has little of 5′GGGGG sequences. On the other hand, 5′ACACA sequence is depleted as the initial population carries an abundance of it yet less is present in "AGO2-loading-IP." In addition, we made two additional control libraries to exclude potential possible artifacts, non-specific binding ("mock IP") and cloning artifacts ("raw oligo"). *See* **Note 3** for more details

2. T4 Polynucleotide Kinase (NEB).
3. 10× PNK Buffer (NEB).
4. 6000 Ci/mmol 10 mCi/ml [γ-^{32}P]-ATP EasyTide (PerkinElmer).
5. Heat block.

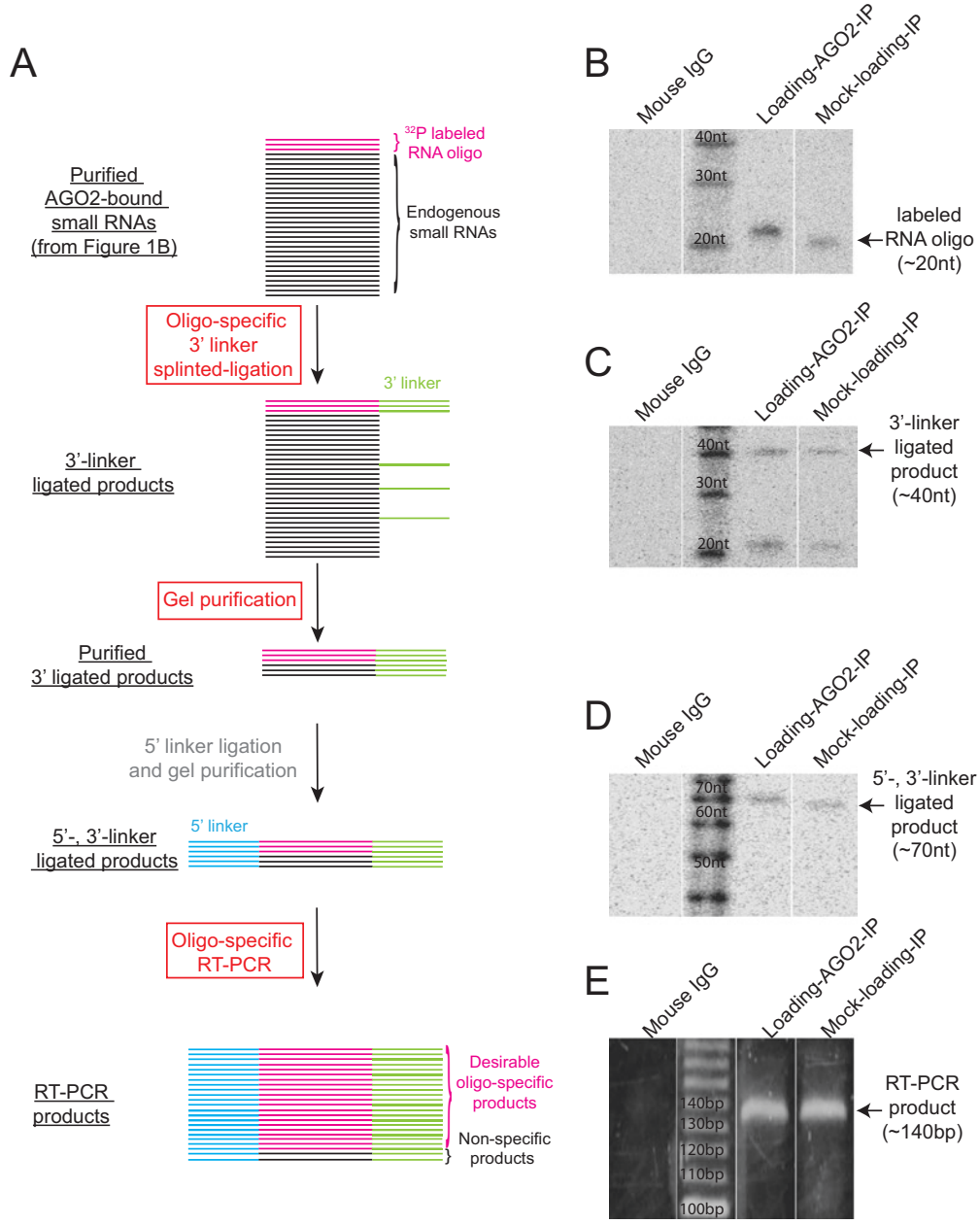


Fig. 2 Preparation of HTS library enriching synthetic RNA loaded in vitro. (**a**) Procedure flow of library construction. *Pink*: AGO2-bound RNA oligos; *Black*: Ago2-bound endogenous small RNAs; *Green*: 3′ linkers ligated to the 3′ ends of RNAs; *Blue*: 5′ linkers ligated to the 5′ ends of RNAs; *Red boxes*: Critical steps that enhance library specificity. After in vitro loading of ^{32}P-labeled partially randomized oligo, the small RNA population extracted from the resulting AGO complexes consists mainly of endogenous small RNAs (*black*) and little of the loaded oligos (*pink*). However, splinted-ligation step (*red boxed*) adds 3′-linkers (*green*) to the 3′ ends of RNA in a sequence-specific manner and the loaded oligos are strongly enriched after the non-ligated endogenous RNAs are eliminated by gel purification (*red boxed*). During the 5′ linker ligation reaction, 5′ linkers (*blue*) are ligated to the 5′ ends of RNAs but this does not confer any specificity. The gene-specific RT-PCR step (*red boxed*) uses primers that partially base-pair with the oligo sequences and further enriches the oligo sequences

6. MicroSpin G25 columns (GE Healthcare).
7. Scintillation cocktail Ultima Gold F (PerkinElmer).
8. 20 ml scintillation vial.
9. Scintillation counter.

2.2 Cell Lysate Preparation for In Vitro Loading

1. 150 cm^2 Tissue Culture Flasks, Plug Seal Cap.
2. S2R+ cells expressing FLAG-HA-tagged AGO2 (S2R+/FLAG-HA-AGO2) [12]; the plasmid for tagged AGO2 expression was a kind gift from the Hannon lab [13].
3. 50 ml Falcon Tube.
4. Hypotonic Buffer: 30 mM HEPES-KOH (pH 7.4), 2 mM Magnesium acetate, 5 mM DTT. Dissolve 1 tablet/10 ml cOmplete ULTRA Tablets Mini (Roche) before use.
5. 25G PrecisionGlide Needle.
6. 1 ml Tuberculin Syringe.
7. 15 ml Falcon Tube.
8. Protein Assay Dye Reagent Concentrate (Bio-Rad).
9. Liquid nitrogen.

2.3 In Vitro Loading Reaction

1. 1 M Potassium Acetate (KOAc).
2. 5 mM Adenosine 5′-triphosphate (ATP) disodium salt solution.
3. 100 mM Dithiothreitol (DTT).
4. RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen).

2.4 Preparation of Antibody-Bound Beads

1. Dyna beads protein G (Novex).
2. Citrate phosphate buffer: 22.4 mM citric acid, 64.8 mM Na_2HPO_4, pH 5.0.

Fig. 2 (continued) and amplify them. The resulting PCR products consist mainly of oligo sequences instead of endogenous RNAs, as opposed to the initial AGO2-IP population. (**b**) Gel image of the small RNA population extracted from in vitro loading samples. The band above 20 nt corresponds to the ^{32}P-labeled partially randomized oligos that were loaded to AGO2 proteins. In the lane of "mock-loading-IP" sample, ^{32}P-labeled partially randomized oligos were added to the endogenous RNAs extracted from the immunoprecipitated AGO2. We performed an additional control experiment to determine the extent of background binding of RNA oligos by performing mock IP using normal mouse IgG after in vitro loading reaction (*see* Fig. 1b for the detailed experimental design.). No signal was detected in this lane. (**c**) Results of the 3′-end splinted linker ligation. Successfully ligated oligos at their 3′ ends are seen at ~40 nt, whereas non-ligated products remained at ~20 nt. Only the band of ~40 nt is purified from the gel. (**d**) Results of the 5′-end linker ligation. Successfully ligated oligos at their 5′ ends seen at ~70 nt are to be purified from the gel. (**e**) Results of PCR products. Bands of ~140 bp are purified for HTS. Non-template controls (NTC) were performed to verify that there was no contamination of the template DNA (not shown)

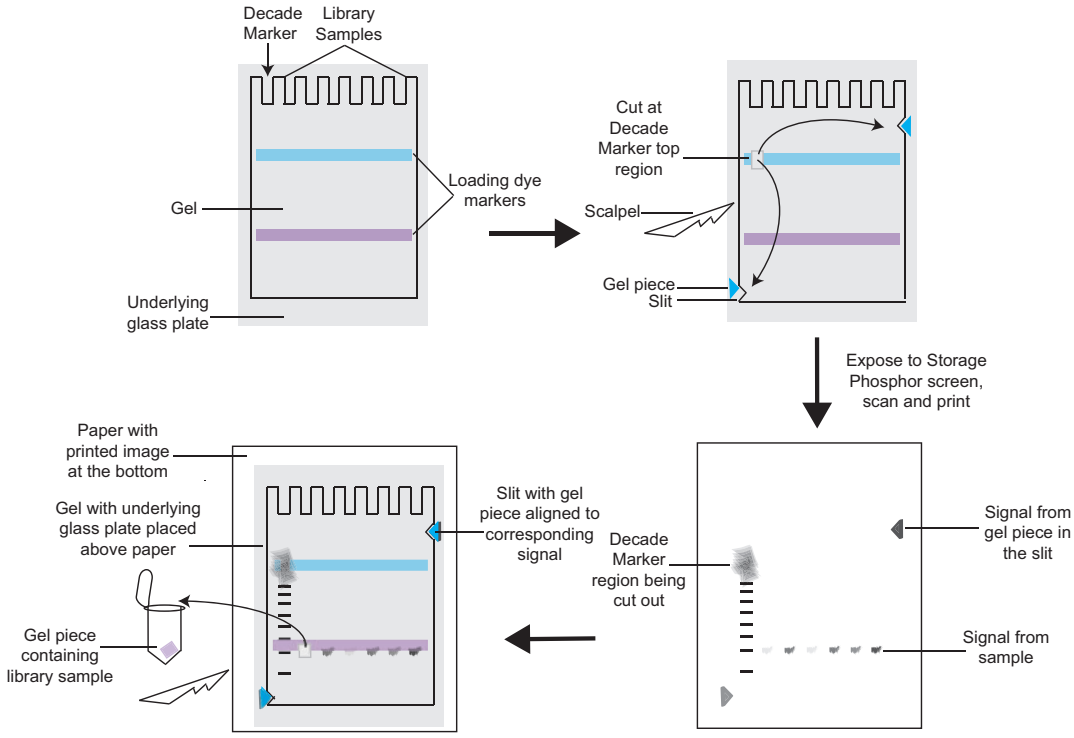


Fig. 3 Gel cutting procedure for HTS library construction. Procedure for gel cutting in library construction as described in Subheading 3.6, **steps 8** and **9**. After gel electrophoresis, remove the top glass plate so that the gel faces up with an underlying glass plate. There will be *top* (*blue*) and *bottom* (*purple*) dye marks visible on the gel due to the Gel Loading Buffer II used for sample loading. Use a scalpel blade to cut out highly radioactive small gel pieces from the top region of the Decade Marker. Cut two slits at the *top right* and *bottom left* of the gel and fill the slits with the radioactive gel pieces. Encase the gel and the plate together with plastic wrap, and expose the gel to a Storage Phosphor screen in a cassette for ~30 min. Scan the Storage Phosphor screen and print the autoradiography image on a paper. Place the gel with the underlying glass over the printed image, such that the two radioactive gel pieces in the slits align with the corresponding signals on the image. Cut out the regions where the ^{32}P-labeled randomized oligos are found

3. Anti-FLAG mouse monoclonal antibody (Anti-DYKDDDDK tag, Wako).
4. Tube Rotator (MACSmix).
5. 1× PBS: 136.9 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4.
6. RIPA buffer: 1× PBS (pH 7.4), 0.1% SDS, 0.5% deoxycholate, 0.5% NP40.

2.5 Immunopurification of FLAG-AGO2 Complex

1. Tube Rotator (MACSmix).
2. 1× PBS: 136.9 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4.
3. RIPA buffer: 1× PBS (pH 7.4), 0.1% SDS, 0.5% deoxycholate, 0.5% NP40.

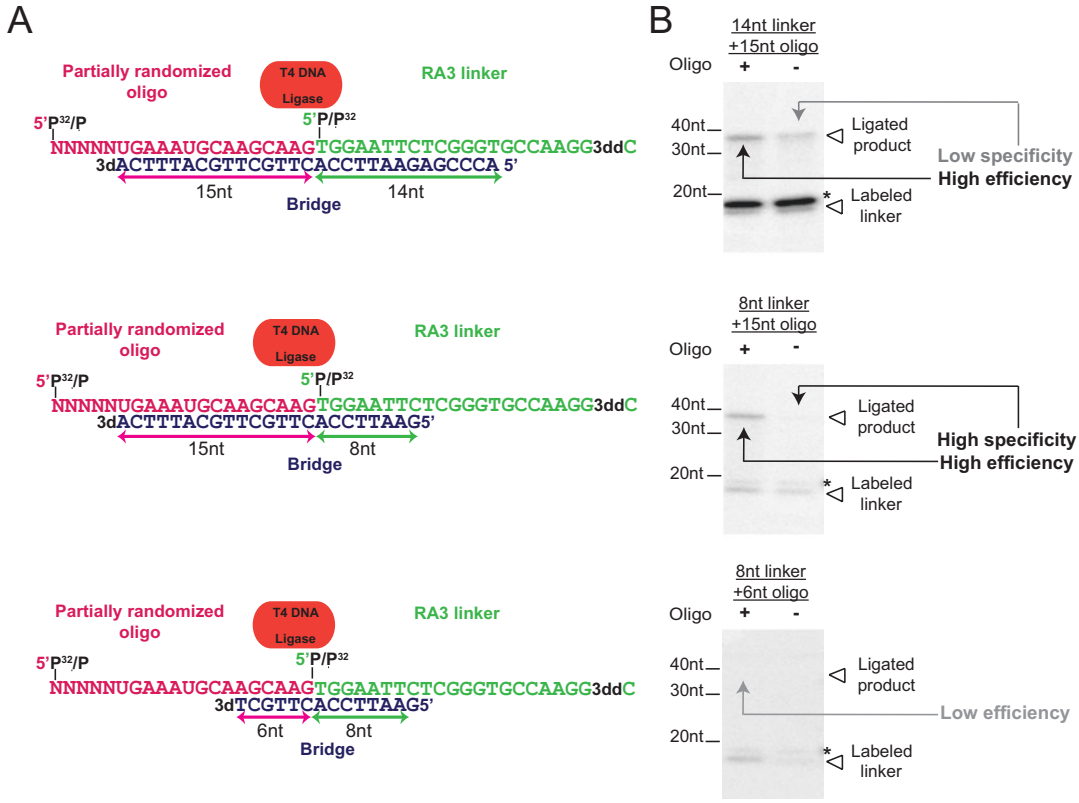


Fig. 4 Optimization of splinted-ligation. (**a**) Design of three different bridges. *Top*: the bridge is composed of a standard 14-nt segment base-paired with the RA3 linker (*green*) [15] and a 15-nt segment fully base-paired with the partially randomized oligo (*red*). *Middle*: the bridge is the same to the top bridge, except that its 5′ end (*green*) is shorter by 6 nt. *Bottom*: the bridge results from trimming the middle bridge further at its 3′ end by 9 nt. (**b**) The optimization outcome for the splinted-ligation using the three different bridges shown in Fig. 2a. All three gel images were adjusted to have the same degree of exposure as indicated by a marker (not shown). Ligated products are seen at ~40 nt. The bands at ~20 nt represent RA3 that was not ligated to the oligo but failed to dephosphorylate by the CIP treatment. A large difference in the intensity between the two bands at ~40 nt (with and without oligo) indicates a high specificity of the bridge. Although the "14 nt linker +15 nt oligo" bridge (*top*) has similar efficiency as that of "8 nt linker +15 nt oligo" bridge (*middle*), it has a much lower specificity. On the other hand, reducing the base pairing with the partially randomized oligo to six (*bottom*) resulted in extremely low efficiency, with no detectable ligated products. Since "8 nt linker +15 nt oligo" bridge gives both high specificity and high efficiency, it was used for library construction. Asterisk indicates artifacts

4. 0.4 M NaCl.
5. Phenol:Chloroform:Isoamyl Alcohol (PCI): Saturated with 10 mM Tris–HCl (pH 8.0), 1 mM EDTA. Mixed at 25:24:1.
6. Pellet Paint Co-Precipitant (Novagen).
7. 100% Ethanol.

2.6 Library Construction

1. T4 Polynucleotide Kinase (PNK) (NEB).
2. 10× PNK Buffer (NEB).

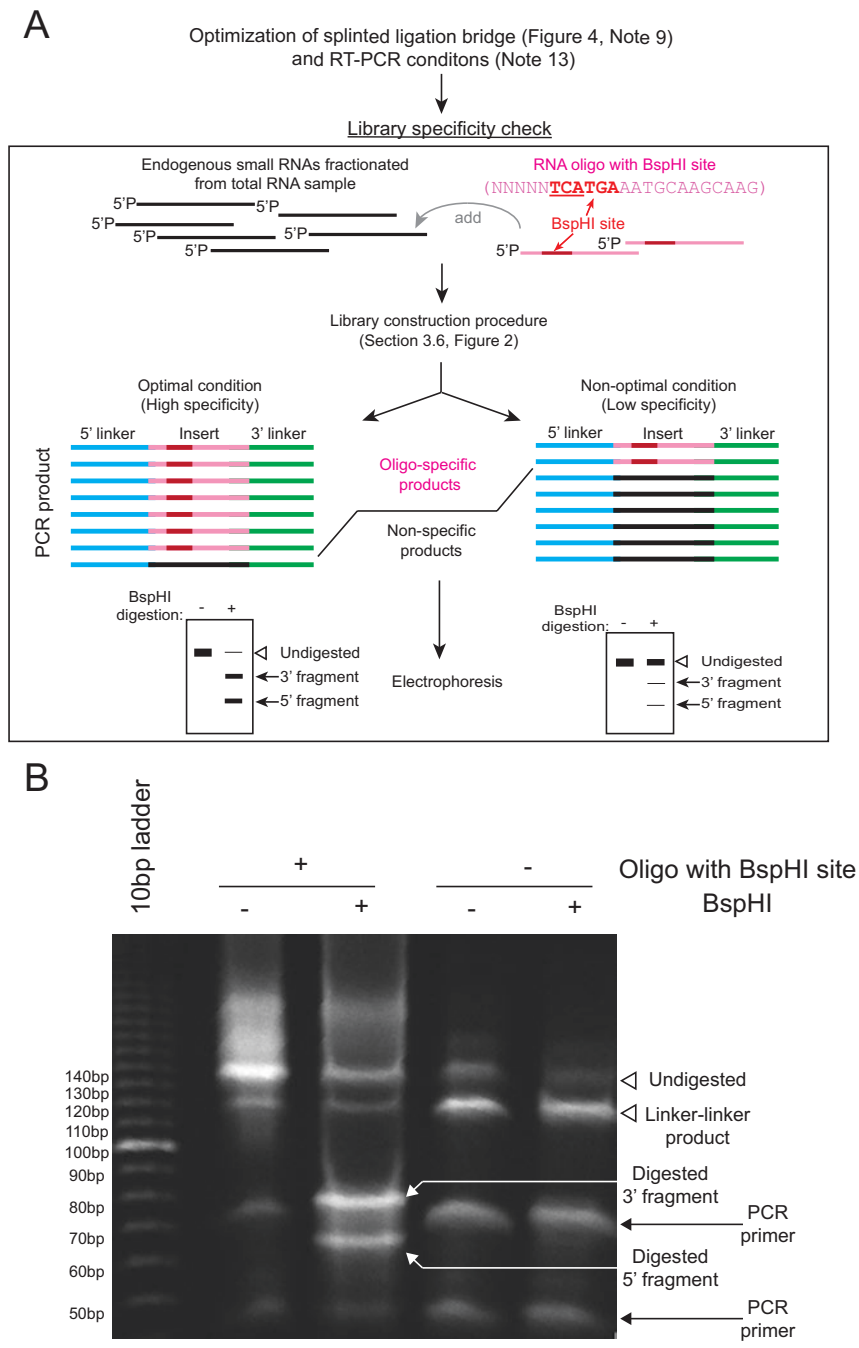


Fig. 5 Specificity check of small RNA library. (**a**) Procedure flow of library specificity check. After selecting the optimal bridge (Fig. 4, *see* **Note 9**) and optimal RT-PCR conditions (*see* **Note 13**), the anticipated library specificity can be checked by using randomized oligos (*pink*) containing a restriction site (*red*). Additional 3 nt (*red*, *underlined*) are inserted to the original oligo sequence to create the BspHI restriction site in this case. These restriction-site-harboring oligos are added to size-fractionated (20–30 nt) endogenous RNA (*black*), mimicking 1× "AGO2-loading-IP" sample (*see* **Note 9**). The RNA mixture is then subjected to library construction procedure as described in Subheading 3.6 and Fig. 2. The resulting PCR products are made up of insert (*pink*/*black*)

3. 6000 Ci/mmol 10 mCi/ml [γ-^{32}P]- ATP EasyTide (PerkinElmer).
4. Heat block.
5. MicroSpin G25 columns (GE Healthcare).
6. Scintillation cocktail Ultima Gold F (PerkinElmer).
7. 20 ml scintillator plastic vial.
8. Scintillator machine.
9. 100 μM DNA oligo bridge with 3′-deoxy base (3d-base) at the 3′ end (*see* **Note 1**), standard desalting (underlined: base-pairing with oligo sequence): GAATTCCACTTGCTTGCATTTC/3d-A/.
10. 10× Capture Buffer: 100 mM Tris–HCl (pH 7.5), 750 mM KCl.
11. 100 nM, 5′ phosphorylated illumina RA3 linker, HPLC: 5′P-T GGAATTCTCGGGTGCCAAGG/3ddC/.
12. PCR Strip Tubes.
13. PCR machine.
14. 2× Rapid Ligation Buffer (Promega).
15. T4 DNA Ligase (Promega).
16. 2× Gel Loading Buffer II (Ambion).
17. Vertical slab gel unit (Sturdier).
18. 50 ml Falcon Tube.
19. SequaGel Buffer (ULTRA-PURE-SEQUAGEL kit, National Diagnostics).
20. SequaGel Diluent (ULTRA-PURE-SEQUAGEL kit, National Diagnostics).
21. SequaGel Concentrate (ULTRA-PURE-SEQUAGEL kit, National Diagnostics).

Fig. 5 (continued) flanked by RP1 primer/5′ linker sequence (*blue*) and RPI/3′ linker sequence (*green*) and are restriction digested by BspHI. Majority of the PCR products will contain the oligo sequence with the BspHI site as the inserts (*pink* with *red*) if the bridge and RT-PCR conditions are highly specific. Therefore, libraries constructed with the optimal conditions will have ~60 (5′ fragment) and ~80 bp (3′ fragment) digested bands as major species after BspHI digestion followed by gel electrophoresis. On the other hand, non-optimal conditions would give little/no digested bands as non-specific products will not change its size (~140 bp) even after BspHI digestion. (**b**) Gel electrophoresis image of a highly oligo-specific library. The PCR product from a sample containing the BspHI-site harboring RNA oligo produced a clear signal at ~140 bp (*black arrowhead*) that was digested into the two bands of ~60 and ~80 bp (*white arrows*) upon BspH1 digestion (*left side*), as elaborated in Fig. 5a. On the other hand, the PCR product from a sample not containing the RNA oligo merely had the band of ~140 bp and was resistant to a digestion by BspH1 (*right side*). The band seen at ~110 bp in all lanes was the product resulting from the empty 5′–3′ linker ligation without inserts (*white arrowhead*). *Black arrows* indicate the PCR primers

22. 10% (w/v) Ammonium persulfate (APS).
23. Tetramethylethylenediamine (TEMED).
24. Diluted Decade Marker RNA: Add 2 μl of freshly ^{32}P-labeled RNA markers (Ambion) to 23 μl of 2× gel loading buffer.
25. Disposable Serological Pipette.
26. 0.5× TBE (pH 8.3): 44.6 mM Tris, 44.5 mM boric acid, 1 m M EDTA.
27. PowerPac.
28. Surgical Blade.
29. Plastic wrap.
30. Storage Phosphor screen (GE Healthcare).
31. Storage Phosphor cassette (GE Healthcare).
32. Typhoon phosphor-imager scanner (Amersham BioSciences).
33. 0.4 M NaCl.
34. MACSmix Tube Rotator (MACSmix).
35. Centrifuge Columns, 0.8 ml (ThermoFisher Scientific).
36. 100% Ethanol.
37. Pellet Paint Co-Precipitant (Novagen).
38. 70% Ethanol.
39. 100 μM, illumina RA5 5′linker (RNA), standard desalting: GUUCAGAGUUCUACAGUCCGACGAUCNN (N: randomized nucleotide, *see* **Note 2**).
40. 10 mM Adenosine 5′-triphosphate (ATP) disodium salt solution.
41. 10× T4 RNA Ligase Buffer (Ambion).
42. T4 RNA Ligase (Ambion).
43. 50% PEG8000 (NEB).
44. RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen).
45. 5 μM, gene-specific RT primer, standard desalting (8 nt base-pairing with linker, underlined: 6 nt base-pairing with oligo sequence): GAATTCCACTTGCT.
46. Thermoscript RNase H-Reverse Transcriptase (Invitrogen).
47. 100 mM DTT.
48. 5× cDNA synthesis buffer (Invitrogen).
49. 10 mM each dNTP Mix.
50. iProof HF Mastermix (Bio-Rad).
51. 10 μM illumina RP1, TOP purified: AATGATACGGCGAC CACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA.
52. 10 μM gene-specific illumina RPI, TOP purified: CAAGCAG AAGACGGCATACGAGATYYYYYYGTGACTGGAGTT

CCTTGGCACCCGAGAATTCCACTTGCTTGC (Y: 6 nt barcode, underlined: 9 nt base-pairing with oligo sequence).

53. Phenol:Chloroform:Isoamyl Alcohol (PCI): Saturated with 10 mM Tris–HCl (pH 8.0),1 mM EDTA. Mixed at 25:24:1.
54. Chloroform.
55. 5 mg/ml Glycogen.
56. 30% Acrylamide/Bis Solution (29:1).
57. 10× TBE (pH 8.3): 892 mM Tris, 890 mM boric acid, 20 mM EDTA.
58. 10 bp DNA ladder (Invitrogen).
59. Blue/Orange 6× Loading Dye (Promega).
60. 1× TBE (pH 8.3): 89.2 mM Tris, 89.0 mM boric acid, 2 mM EDTA.
61. 1× SYBR Safe in 0.5× TBE: Add 30 μl 10,000× SYBR Safe (Invitrogen) to 300 ml 0.5× TBE.
62. Shaker.
63. Gel Doc (BioRad).
64. Blue light emitter (Maestrogen).
65. Nanodrop (ThermoFisher Scientific).

3 Methods

To analyze the sequence specificity of AGO2, RNA species loaded to AGO2 in the in vitro reaction were cloned and sequenced. To analyze the enrichment/depletion of each of the ~1000 species (=4^5: all combinations of 4 nucleotides at 5 positions) in the loaded population compared to the population before the loading reaction, we made a "mock-loading" control library (Fig. 1b). For the "mock-loading" control, we mixed the RNA oligo with endogenous AGO2-associated small RNA species immuno-purified from the cells prior to library construction. Enrichment/depletion of each of the randomized RNA sequences in the AGO2-IP library was estimated using the abundance of the sequence in the mock-loading control library as the latter represents the initial oligo population (*see* **Note 3**).

Here, we will describe the methods of enriching RNA species preferentially loaded to AGOs in the lysate (Subheadings 3.1–3.5) and constructing deep-sequencing libraries (Subheading 3.6).

3.1 ^{32}P ssRNA Oligo Labeling for In Vitro Loading

1. Set up two of the following PNK reaction in eppendorf tubes (*see* **Note 4**):
2. Incubate the reaction tubes at 37 °C for 1 h.

1 μM partially randomized RNA oligo	4.5 μl
10× PNK buffer	2 μl
[γ-^{32}P]-ATP	3 μl
T4 PNK enzyme	1 μl
Sterile water	9.5 μl

3. Gently flick two of the MicroSpin G25 columns to mix the beads and place them on the microfuge tubes provided by the kit. Spin the columns at 800 rcf for 1 min and transfer the columns to eppendorf tubes. After transferring the labeling reactions to the columns, spin the tubes at 800 rcf for 1 min and discard the columns.
4. Heat the samples at 95 °C for 2 min and place them immediately on ice for another 2 min to inactivate PNK and denature the ssRNA oligo. Collect all the solutions into one single tube if the reactions were done in multiple tubes.
5. Add 1 μl of the ^{32}P-labeled oligo in 5 ml Scintillation cocktail in a scintillation vial and measure the radioactivity using a scintillation counter (*see* **Note 5**).

3.2 Cell Lysate Preparation for In Vitro Loading

1. Harvest S2R+/FLAG-HA-AGO2 cells from 12 confluent T150 flasks by spinning the culture in 12 falcon tubes at 800 rcf for 10 min. Discard the supernatant.
2. Add 1× cell pellet volume of hypotonic buffer to the cells and resuspend them thoroughly using a pipette. Make 500–800 μl aliquots of the suspended sample with 1.5 ml tubes.
3. Lyse the cells five times through a 25G needle affixed to a 1 ml syringe. Change the needle for every tube. Spin the tubes at 21,100 rcf at 4 °C for 15 min. Collect all the supernatants in a 15 ml falcon tube and mix it well.
4. Measure the total protein concentration using the Protein Assay Dye reagent.
5. Adjust the total protein concentration to ~10 mg/ml by adding the hypotonic buffer. Freeze lysate aliquots (160 μl each) quickly in liquid nitrogen and store them at −80 °C. Typically, ~40 lysate aliquots could be made from 12 of the T150 flasks of cells.

3.3 In Vitro Loading Reaction

In this section, the lysate prepared in Subheading 3.2 will be used for the in vitro loading of the ^{32}P-labeled, partially randomized oligos (Subheading 3.1). The samples should always be kept on ice unless otherwise specified.

1. Thaw ten tubes from Subheading 3.2, **step 5**. Add the following reagents into each of the ten tubes (*see* **Note 6**):

5 mM ATP	32 µl
1 M KOAc	32 µl
100 mM DTT	16 µl
RNAse out	0.8 µl
Sterile water	77.2 µl

2. Label five of the tubes as "AGO-loading," another five as "mock-loading" (*see* **Note 6**). Add 2×10^6 counts per minute (cpm) of ^{32}P–labeled oligo (Subheading 3.1, **step 4**) to each of the five "AGO-loading" tubes. In contrast, add an equivalent amount of water to "mock-loading" tubes. Mix the samples thoroughly with a 1000 µl pipette tip.
3. Incubate all samples at 25 °C on a heat block for 1 h (*see* **Note 7**).

3.4 Preparation of Antibody-Bound Beads

For immunoprecipitation of the ten reactions from Subheading 3.3, **step 3**, ten tubes of antibody-bound beads will be needed.

1. Add 120 µl of Dyna beads to each of ten new tubes.
2. Insert tubes into the magnetic rack and remove supernatant. Wash the beads thoroughly with citrate phosphate buffer three times. Pipette 800 µl of citrate phosphate buffer to each of the tubes.
3. Add 20 µg of FLAG antibody to each tube. Rotate the samples for 30 min at room temperature.
4. Spin down the tubes briefly, and place them on the magnetic rack. After removing all solution, wash the beads with citrate phosphate buffer three times and subsequently with RIPA buffer two times. Make sure that the buffer is completely removed after the last wash.

3.5 Immunopurification of FLAG-AGO2 Complex

1. Label five tubes from Subheading 3.4, **step 4** "AGO2-loading-IP" and the remaining five "mock-loading-IP." Add 700 µl of RIPA buffer to each of radioactive loading reactions from Subheading 3.3, **step 3** and mix them well. Transfer the loading reactions to the respective antibody-bound beads according to the labels.
2. Rotate the tubes at 4 °C for 1.5 h.
3. Spin down all samples briefly and place them on the magnetic rack. Discard all lysate solution.
4. Wash the beads thoroughly with RIPA buffer four times. Mix the beads well with 500 µl of RIPA buffer and transfer entire volumes (with beads) to new eppendorf tubes (*see* **Note 8**). Label them accordingly.

5. Place these tubes to the magnetic rack and remove the solution. Wash beads with PBS and remove solution completely.
6. Add 200 µl of 0.4 M NaCl followed by 200 µl of PCI to the beads and mix thoroughly. Spin the tubes at 21,100 rcf at 4 °C for 15 min.
7. For all samples, transfer 150 µl of the upper-phase to new pre-labeled tubes. Add 2 µl of Pellet Paint Co-Precipitant and 450 µl of 100% ethanol to each tube. Shake to mix. Store them at −20 °C overnight to ensure efficient precipitation.

3.6 Library Construction

This section describes the method for construction of small RNA libraries that enrich the target sequence (Fig. 2). Gel purification steps described here (Fig. 3) are important in enhancing the library specificity and recommended not to be skipped. Before constructing the actual library for deep-sequencing, the bridge sequence and ligation conditions need to be optimized to ensure efficient ligation (Fig. 4, *see* **Note 9**). We also strongly recommend checking the specificity of the library with the selected bridge oligo. A partially randomized oligo containing a restriction enzyme digestion site should be used (*see* **Note 10**) so that the specificity of resulting libraries can be examined by digestion of the PCR product (Fig. 5).

1. Centrifuge all IP samples from Subheading 3.5, **step 7** at 21,100 rcf at 4 °C for 30 min. Discard the supernatants leaving behind just the pinkish pellets (due to the Pellet Paint Co-Precipitant). Rinse the pellet with 70% ethanol and ensure no ethanol left in the tube. Air dry the RNA pellets for 1 min. Resuspend pellet in 5 µl of 2× gel loading buffer. Combine all the samples with same labels so that each IP sample has 25 µl volume of RNA. Add a proper amount (1100 cpm) of ^{32}P-labeled randomized oligo found in "AGO2-loading-IP" into the "mock-loading-IP." (The amount of RNA oligo that is present in the "AGO2-loading-IP" is estimated according to the method in *see* **Note 6**.)
2. Add 25 µl of 2× gel loading buffer to each sample. Heat all the samples and diluted Decade Markers at 95 °C for 2 min, and put them on ice for another 2 min.
3. Set up the Sturdier vertical slab gel unit(s) with 1.5 mm spacers for gel purification of the small RNA species (*see* **Note 11**).
4. In a 50 ml falcon tube, make 15% urea gel as follows and mix it well:

SequaGel diluent	15 ml
SequaGel concentrate	30 ml
SequaGel buffer	5 ml
10% APS	400 µl
TEMED	20 µl

5. Use a 25 ml serological pipette to fill up the vertical slab with the urea gel. Insert a 1.5 mm comb from above and stop ¾ way through. The gel fully solidifies in about 30 min.
6. After pre-running the urea gel in 0.5× TBE buffer, wash the wells with 0.5× TBE buffer again.
7. Load diluted Decade Markers and samples from Subheading 3.6, **step 2** into the wells and carry out electrophoresis at 300 V for 2.5 h.
8. After the electrophoresis, disassemble the gel set and leave the gel facing up on the underlying glass plate. Use a scalpel blade to cut two slits at the top left and bottom right of the gel. Cut out highly radioactive small gel pieces from the top area of the Decade marker lane and insert them into the slits. Secure the gel and the plate together with plastic wrap, and expose the gel to a Storage Phosphor screen in a cassette for ~30 min (Fig. 3).
9. Scan the screen using the Typhoon scanner. Print the scanned image and place the gel over the printed image, aligning radioactive signals from the two slits (Fig. 3). Cut out the ~20–30 nt region for each lane where the ^{32}P-labeled randomized oligos are found (Fig. 2c).
10. Place each gel piece into a new eppendorf tube. Add 400 µl of 0.4 M NaCl and grind the gel piece by a pipette tip. Rotate the tubes overnight to elute the small RNAs.
11. Transfer the entire suspensions to Pierce columns (columns sit on top of new eppendorf tubes) and spin them at 8000 rcf for 1 min. Discard the columns. Add 1000 µl of 100% ethanol and 2 µl of Pellet Paint Co-Precipitant to each tube, and precipitate small RNA samples at −80 °C for 45 min or −20 °C for 3 h.
12. Spin all the samples at 21,100 rcf at 4 °C for 30 min. Discard the supernatants. Rinse the RNA pellet with 70% ethanol and ensure there is no remaining ethanol. Air dry the RNA pellets for 1 min.
13. Resuspend each pellet with 8 µl sterile water.
14. Set up the following splinted-ligation reactions in PCR tubes:

Sample from Subheading 3.6, **step 13**	8 µl
100 nM bridge in capture buffer	1 µl
100 nM 5′P-RA3 linker (not ^{32}P-labeled)	1 µl

15. Place the tubes on a PCR machine with the settings below:

(a) 94 °C	1 min
(b) 65 °C	2 min
(c) 37 °C	10 min
(d) 37 °C	Forever
(e) 37 °C	60 min
(f) 70 °C	10 min

16. Right after step (c) (37 °C for 10 min), add 10 µl of 2× ligation buffer and 2 µl of T4 DNA ligase to each of the reactions. Mix the reactions by pipetting. Start step (e) to allow 1-h ligation at 37 °C. *Do not remove the tubes from the PCR machine in the process.*
17. Cast two 15% urea gels with 1.5 mm spacers (*see* Subheading 3.6, **steps 3** and **6**).
18. After adding 25 µl of 2× gel loading buffer to each sample, heat all the samples and the diluted Decade Markers at 95 °C for 2 min, and put them on ice for 2 min.
19. Load the samples from Subheading 3.6, **step 18** into the wells, and run the gel for ~3 h at 300 V.
20. Gel purify the ~40–50 nt ligated products as described in Subheading 3.6, **steps 8** and **12** (Fig. 2c) (*see* **Note 12**). Resuspend each RNA pellet with 3.5 µl of sterile water.
21. Denature the sample and the 5′ linker by heating the following mixture at 70 °C for 2 min.

Sample from Subheading 3.6, **step 20**	3.5 µl
100 µM RA5 linker	2 µl

22. Set up the ligation reaction and incubate at 37 °C for 2 h.

Denatured sample from Subheading 3.6, **step 21**	5.5 µl
10 mM ATP	2 µl
10× T4 RNA ligase buffer	2 µl
50% PEG 8000	8 µl
T4 RNA ligase 1	2 µl
RNAse OUT	0.5 µl

23. Set up a 12% urea gel and pre-run the gel (*see* Subheading 3.6, **steps 3** and **6**):

SequaGel diluent	21 ml
SequaGel concentrate	24 ml
SequaGel buffer	5 ml
10% APS	400 µl
TEMED	20 µl

24. After adding 25 µl of gel loading buffer to the samples from Subheading 3.6, **step 22**, heat the mixtures at 95 °C for 2 min and put them on ice for another 2 min. Load the samples into the well and run the gel for 3.5 h at 300 V.
25. Expose the gel and cut out the bands in the ~70–80 nt region according to Subheading 3.6, **steps 8** and **9** (Fig. 2d) (*see* **Note 12**).
26. Elute and precipitate RNAs as described in Subheading 3.6, **steps 10** and **12**. Resuspend each RNA pellet with 6 µl of sterile water, and transfer it to a PCR tube.
27. Reverse transcribe the RNA samples. Add 5 µl of 5 µM gene-specific RT primer to the samples from Subheading 3.6, **step 26** (*see* **Note 13**). Heat the mixture at 63 °C for 2 min and then cool them to 58 °C. *Without removing the tubes* from the PCR machine, add the following reagents:

5× cDNA synthesis buffer	4 µl
100 mM DTT	1 µl
10 mM dNTP mix	2 µl
RNaseOUT	1 µl
ThermoScipt	1 µl

28. Continue to incubate the reactions at 58 °C for 1 h. Terminate the reaction by heating the samples at 85 °C for 15 min.
29. Set up the following PCR reactions:

First-strand cDNA (from Subheading 3.6, **step 28**)	5 µl
2× iProof HF Master Mix	50 µl
10 µM RP1	8 µl
10 µM RPI (different for each sample) (*see* **Note 14**)	8 µl
Sterile water	29 µl

30. Run the PCR reaction with the following program:

(a) 98 °C	2 min
(b) 98 °C	10 s
(c) 66 °C	30 s
(d) 72 °C	15 s
Step (b) to (d): 25 cycles	
(e) 72 °C	10 min

31. After transferring the PCR samples to new eppendorf tubes, mix them well with 200 µl of 0.4 M NaCl and 300 µl of PCI. Centrifuge the mixture at 21,100 rcf for 5 min, transfer the aqueous layer to 200 µl chloroform, and mix it well. Centrifuge at 21,100 rcf for 5 min again. After transferring the aqueous layer to a new eppendorf tube, add 1 µl of glycogen and three volumes of 100% ethanol. Precipitate the samples at −80 °C for 45 min or −20 °C for 3 h.
32. Centrifuge all the samples at 21,100 rcf at 4 °C for 30 min. Discard the supernatants leaving behind just the white pellets (due to glycogen). Rinse the pellet with 70% ethanol and ensure no ethanol left. Air dry the RNA pellets for 1 min. Resuspend the pellet with 24 µl of sterile water.
33. Set up the Sturdier vertical slab gel unit(s) with 1.5 mm spacers.
34. In a 50 ml falcon tube, mix the following regents to make 6% native gel:

30% Acrylamide	10 ml
10× TBE	5 ml
Sterile water	35 ml
10% APS	300 µl
TEMED	23 µl

35. Use a 25 ml serological pipette to fill up the vertical slab with the native gel. Insert a 1.5 mm comb from above and stop ¾ way through. The gel fully solidifies in about 45 min.
36. Aliquot 6 µl of DNA loading buffer to all samples from Subheading 3.6, **step 32**. Mix 3 µl of 10 bp ladder with 6 µl of DNA loading buffer in a new eppendorf tube. *Do not heat the samples.* Load all the samples including the ladder into the wells. Start electrophoresis at 200 V for 1.5 h.

37. Stain the gel in 0.5× TBE containing 1× SYBR Safe, for 5–10 min on a shaker. After staining, place the gel on a plastic wrap with an underlying glass plate. Visualize bands under the blue light and cut out bands of 130–140 bp (*see* **Note 14**) (Fig. 2e). Transfer the gel pieces in an eppendorf tube and add 80 μl of sterile water to it. Grind the gel into tiny pieces for overnight elution at room temperature.
38. Transfer all the suspensions to Pierce columns (columns sit on the top of new eppendorf tubes) and centrifuge them at 8000 rcf for 1 min. Discard the columns. Check the concentration of the PCR products by nanodrop. The library is ready for quality check and sequencing.

4 Notes

1. The bridge oligo should have a 3′ end modification (*see* Subheading 2.6, **item 9**) to prevent it from acting as RT primer. However, the 3′ dideoxy modification is not essential for the preliminary tests of bridges (*see* **Note 9**).
2. The randomized bases at the 3′ end of the 5′ linker aid in improving the ligation efficiency [14].
3. In addition to the mock-loading sample, we also constructed two additional control libraries. One of the controls undergoes the same loading reaction as well as the AGO2-IP sample, except for using normal mouse IgG instead of FLAG-antibody. This will ensure that the enriched sequences in the AGO2-IP libraries are the species bound not to the beads but to AGO2. Another control ("Raw") is a library using the partially randomized RNA oligo without mixing it with the AGO2-associated endogenous small RNAs. "Raw" sample serves as a good control for differentiating between contamination/abnormal read counts versus real events of binding. Abnormal read counts could come from endogenous RNAs that share same backbone sequence as the randomized oligo. We exclude sequences that have abnormally high read counts in the mock-loading samples compared to those in "Raw" samples by assuming that those reads were derived from endogenous species. For example, in this protocol, we are randomizing the first five positions of the mir-317 loop sequence (NNNNNUGAAAUGCAAGCAAG) and hence the resulting libraries contain a high read count of the endogenous mir-317 loop sequence (UUGAAUGAAAUGCAAGCAAG). *This sequence could be removed by the filtering using the "Raw" library data.*
4. The scale of the ^{32}P-labeling reaction of randomized oligo should be determined according to the scale of loading reaction.

Each standard labeling reaction typically yields ~6 × 10^6 cpm labeled oligos and the five tubes of lysate used for "AGO2-loading-IP" (Subheading 3.3) would require 1 × 10^7 cpm in total. Therefore, two ^{32}P labeling reactions were carried out in this protocol.

5. The ^{32}P-labeled oligos can be stored at −20 °C.
6. It is important to estimate the scale of loading reaction per IP sample before preparing the library. We typically perform in vitro loading of the randomized oligo in 160 μl of lysate, and subsequently ~80 μl of the reaction is used for each of AGO2-IP and Mouse IgG mock IP. For library construction, a proper scale yields ~1100 cpm of oligos in the AGO2-IP sample, which corresponded to 10× of this standard reaction. Therefore, 800 μl of the lysate is required per IP sample (equivalent to 5 tubes of 160 μl lysate).
7. The incubation temperature during the loading reaction is important because the degradation rates of the introduced ssRNAs can vary significantly depending on the temperature and would affect the loading efficiency. The temperature needs to be kept constant throughout the reaction by placing the tubes on a heat block and keeping samples away from any heat sources. The loading reaction shown here is optimized at 25 °C.
8. This step is essential to avoid carryover of oligonucleotides non-specifically bound to the walls of eppendorf tubes.
9. It is highly recommended to optimize the bridge design used in the 3′ linker splinted-ligation step. As specificity generally increases as the length of bridge decreases, we aim to reduce the number of base-pairings with the 3′ linker and oligo as much as possible. In addition, determination of the minimum base-pairings to oligo aids in deciding the maximum number of randomized bases for study.

   We used three bridges with different lengths for optimization (Fig. 4a). The bridge sequences for this test are shown below (100 nM in 1× Capture buffer, standard desalting; underlined: base-pairing with the constant sequence of the partially randomized RNA oligo):

   - ACCCGAGAATTCCA<u>CTTGCTTGCATTTCA</u> (14-nt linker +15-nt oligo).
   - GAATTCCA<u>CTTGCTTGCATTTCA</u> (8-nt linker +15-nt oligo).
   - GAATTCCA<u>CTTGCT</u> (8-nt linker +6-nt oligo).

   The splinted-ligation assay was carried out according to a published protocol [15]. The optimal bridge is the one that gives the highest ratio of ~40 nt ligated product between the lane with oligo and the negative control lane (without oligo) (Fig. 4b).

For optimization of splinted-ligation reaction, a large amount of "AGO2-loading-IP" samples will be required. Alternatively, one can mix appropriate amounts of non ^{32}P-labeled RNA oligo and size-fractionated (20–30 nt) small RNA that is gel-purified from the lysate to mimic the "AGO2-loading-IP" samples, instead of using precious AGO-IP RNA samples. Based on our Western blotting analysis (not shown), the AGO2-IP efficiency is ~10% and the in vitro loading followed by AGO2-IP typically yield 0.8 fmol RNA oligo (~1100 cpm). Therefore, 0.8 fmol 5′-monophosphorylated (non-radioactive) RNA oligo was added to an RNA sample gel purified from 80 μl of the lysate (10% of 800 μl lysate used per IP sample).

10. To estimate the specificity of library construction, instead of using the RNA oligo for actual library construction we used the following RNA oligo (underlined: BspHI restriction site: NNNNN<u>UCAUGA</u>AAUGCAAGCAAG). The PCR products containing the correct target sequence retain the diagnostic BspHI site (Fig. 5a), whereas non-target species derived from endogenous sources do not. The cleavage percentage reflects the specificity of the library (Fig. 5b). For this test, we prepared two samples: One of the samples contains the randomized oligo harboring the BspHI site while another contains no oligo (Fig. 5b). In this way, one could deduce if the digested products are due to the specific amplification of oligo sequence or background contributed by endogenous RNAs. It is useful to have a undigested sample as a negative control in the case where unexpected bands emerge in the "no oligo" sample or no smaller sized bands are seen in the "with oligo" sample. Having non-template RT and/or non-template PCR as a control helps detect template contamination. Do not overexpose the agarose gel to the UV light so as to ensure linearity of quantification. This protocol uses the GAATTCCA<u>CTTGCTT GCATTTCA</u> (8-nt linker +15-nt oligo) bridge and gives 80–90% specificity.
11. To minimize contamination of shorter or longer RNA species, the AGO2-IP samples were gel purified.
12. Note that 3′- and 5′-end linker ligation efficiencies are not 100%. Therefore, residual free radiolabeled oligos may remain in the regions of 20–30 and 40–50 nt. Run the gel long enough to clearly separate ligated and non-ligated RNA products for subsequent easy excision. After the 5′ ligation, the ligated (~70–80 nt) and non-ligated (~40–50 nt) RNA products should be separated well.
13. It is desirable to optimize the gene-specific RT and RPI primers as well as the corresponding annealing temperature. As a general guideline, the specificity is determined primarily by the

number of the complementary bases between the oligo and the RT primer, and the degree of the complementarity between the oligo and the RPI primer alters the specificity less strongly. In addition, it is essential to make sure that at least ~6 nt of the oligo sequence (right after the randomized region) are not base-paired by the RT or RPI primer as they are needed for identification of the specific species in the sequencing data. If the splinted-ligation is sufficiently specific, it is optional to use gene-specific primers or high annealing temperature. In this protocol, the RT and RPI primers have 6- and 9-nt base-pairings with the oligo, respectively (*see* Subheading 2.6, **items 45** and **52**). RT and PCR annealing temperatures are set at 58 and 66 °C, respectively.

14. Do not include any of ~110 bp PCR products because they are linker-linker ligation products.

Acknowledgements

We thank Gregory Hannon for sharing the tagged AGO2 plasmid. Research in K.O.'s group was supported by the National Research Foundation, Prime Minister's Office, Singapore under its NRF Fellowship Programme (NRF2011NRF-NRFF001-042).

References

1. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
2. Kobayashi H, Tomari Y (2016) RISC assembly: coordination between small RNAs and Argonaute proteins. Biochim Biophys Acta 1859(1):71–81. doi:10.1016/j.bbagrm.2015.08.007
3. Suzuki HI, Katsura A, Yasuda T, Ueno T, Mano H, Sugimoto K, Miyazono K (2015) Small-RNA asymmetry is directly driven by mammalian Argonautes. Nat Struct Mol Biol 22(7):512–521. doi:10.1038/nsmb.3050
4. Chak LL, Okamura K (2014) Argonaute-dependent small RNAs derived from single-stranded, non-structured precursors. Front Genet 5:172. doi:10.3389/fgene.2014.00172
5. Okamura K, Ladewig E, Zhou L, Lai EC (2013) Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals. Genes Dev 27(7):778–792. doi:10.1101/gad.211698.112
6. Winter J, Link S, Witzigmann D, Hildenbrand C, Previti C, Diederichs S (2013) Loop-miRs: active microRNAs generated from single-stranded loop regions. Nucleic Acids Res 41(10):5503–5512. doi:10.1093/nar/gkt251
7. Lima WF, Prakash TP, Murray HM, Kinberger GA, Li W, Chappell AE, Li CS, Murray SF, Gaus H, Seth PP, Swayze EE, Crooke ST (2012) Single-stranded siRNAs activate RNAi in animals. Cell 150(5):883–894. doi:10.1016/j.cell.2012.08.014
8. Chorn G, Klein-McDowell M, Zhao L, Saunders MA, Flanagan WM, Willingham AT, Lim LP (2012) Single-stranded microRNA mimics. RNA 18(10):1796–1804. doi:10.1261/rna.031278.111
9. Liu J, Yu D, Aiba Y, Pendergraff H, Swayze EE, Lima WF, Hu J, Prakash TP, Corey DR (2013) ss-siRNAs allele selectively inhibit ataxin-3 expression: multiple mechanisms for an alternative gene silencing strategy. Nucleic Acids Res 41(20):9570–9583. doi:10.1093/nar/gkt693
10. Hu J, Liu J, Narayanannair KJ, Lackey JG, Kuchimanchi S, Rajeev KG, Manoharan M, Swayze EE, Lima WF, Prakash TP, Xiang Q, Martinez C, Corey DR (2014) Allele-selective inhibition of mutant atrophin-1 expression by

duplex and single-stranded RNAs. Biochemistry 53(28):4510–4518. doi:10.1021/bi500610r

11. Matsui M, Corey DR (2016) Non-coding RNAs as drug targets. Nat Rev Drug Discov. doi:10.1038/nrd.2016.117
12. Okamura K, Liu N, Lai EC (2009) Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. Mol Cell 36(3):431–444. doi: S1097-2765(09)00687-X [pii] 10.1016/j.molcel.2009.09.027
13. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel J, Sachidanandam R, Hannon G, Brennecke J (2008) An endogenous siRNA pathway in Drosophila. Nature 453:798–802
14. Zhang Z, Lee JE, Riemondy K, Anderson EM, Yi R (2013) High-efficiency RNA cloning enables accurate quantification of miRNA expression by deep sequencing. Genome Biol 14(10):R109. doi:10.1186/gb-2013-14-10-r109
15. Maroney PA, Chamnongpol S, Souret F, Nilsen TW (2008) Direct detection of small RNAs using splinted ligation. Nat Protoc 3(2):279–287. doi:10.1038/nprot.2007.530

Chapter 4

Dumbbell-PCR for Discriminative Quantification of a Small RNA Variant

Megumi Shigematsu, Shozo Honda, and Yohei Kirino

Abstract

Cellular RNAs are often expressed as multiple isoforms of complex heterogeneity in both length and terminal sequences. IsomiRs, the isoforms of microRNAs, are such an example. Distinct quantification of each RNA variant is necessary to unravel the biogenesis mechanism and biological significance of heterogenetic RNA expression. Here we describe Dumbbell-PCR (Db-PCR), a TaqMan RT-PCR-based method that distinctively quantifies a specific small RNA variant with single-nucleotide resolution at terminal sequences. Db-PCR enables the quantitative analysis of RNA terminal heterogeneity without performing Next-Generation Sequencing.

Key words Dumbbell-PCR, Db-PCR, TaqMan qRT-PCR, T4 RNA Ligase 2, Small regulatory RNA, miRNA, isomiR, tRNA fragment

1 Introduction

Recent advances in Next-Generation Sequencing (NGS) technologies have revealed that RNAs are not always expressed as single entities with fixed terminal sequences. Instead, their expression allows for multiple isoforms bearing complex heterogeneity in both length and terminal sequences. For example, identical microRNA (miRNA) genes encode mature isoforms, termed isomiRs, that vary in size by one or more nucleotides (nt) at the 5′- and/or 3′-end of the miRNA [1]. The isomiRs are differentially expressed across different cell and tissue types at different developmental stages and during disease [2–7]. To unravel the complexities of RNA heterogeneity and their biological significance, the accurate and convenient quantification of individual small RNA variants is imperative. A conventional TaqMan RT-PCR using stem-loop primers [8], microarray analysis, and Northern blot analysis are all insufficient to distinguish miRNAs from their corresponding isomiRs [9, 10]. NGS can capture the entire repertoire of miRNAs and their corresponding isomiRs. However, the

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_4,

required cost, time, and bioinformatic analyses could preclude the use of NGS for examining only focused isomiR species.

We recently developed Dumbbell-PCR (Db-PCR), an efficient and convenient TaqMan RT-PCR-based method, that can specifically quantify individual small RNA variants with single-nucleotide resolution at terminal sequences [10]. Db-PCR has three variations: 3′-Db-PCR, 5′-Db-PCR, and Db-PCR (Fig. 1). Of these, the first two methods have been designed for selective quantification of 3′-variants and 5′-variants of RNAs, respectively, whereas Db-PCR enables simultaneous discrimination of target RNAs from their corresponding 5′- and 3′-variants. In each of the three methods, a stem-loop adapter is specifically ligated to target RNA variants by using T4 RNA ligase 2 (Rnl2) which catalyzes RNA ligation at a 3′-OH/5′-P nick in a double-stranded RNA (dsRNA) or an RNA-DNA hybrid [11–13]. For the 3′- or 5′-Db-PCR method, a 3′-Dumbbell adapter (3′-Db-adapter; containing a protruding 3′-end) or a 5′-Dumbbell adapter (5′-Db-adapter; containing a protruding 5′-end) is hybridized and ligated to the 3′- or 5′-end of the target RNA in the total RNA, respectively (Fig. 1a, b). Rnl2 ligation efficiency is severely reduced when double-stranded nucleotides of the substrate contain gaps or overlaps [12], suggesting that Rnl2 ligation achieves high specificity toward target RNA. Db-PCR has been designed by combining 3′- and 5′-Db-PCRs, and this Db-PCR procedure involves hybridization and Rnl2 ligation of the 5′-Dumbbell adapter containing a spacer (5′-Dbs-adapter) to the 5′-end of target RNAs, followed by 3′-Db-adapter addition and ligation to the 3′-end (Fig. 1c).

For the 3′- and 5′-Db-PCR methods, the ligation product with a one-sided “dumbbell-like” secondary structure is amplified and quantified by a TaqMan RT-PCR. This involves the TaqMan probe targeting the boundary of the adapter and target RNA. Because the TaqMan probe has the ability to discriminate a single nucleotide difference [14], the design results in highly specific detection of target RNA which does not cross-react with its terminal variants. In addition to the 3′-terminal discrimination by the TaqMan probe, Db-PCR allows the forward primer to discriminate the 5′-end of the target RNA by recognizing 5′-terminal nucleotides of the target RNA. Db-PCR has broad applicability for quantifying various small RNAs in different cell types, and the quantification results are consistent with those from Northern blot analyses [10]. Thus, Db-PCR is an efficient and convenient technique for specific detection and differential expression analysis of RNA variants, such as isomiRs (Fig. 2a). In addition, it can also shed light on the biological significance of small RNAs coexisting with abundant precursor RNAs in cells (e.g., tRNA-derived RNAs), allowing for their specific detection and quantification which require discrimination from their precursor RNAs (Fig. 2b).

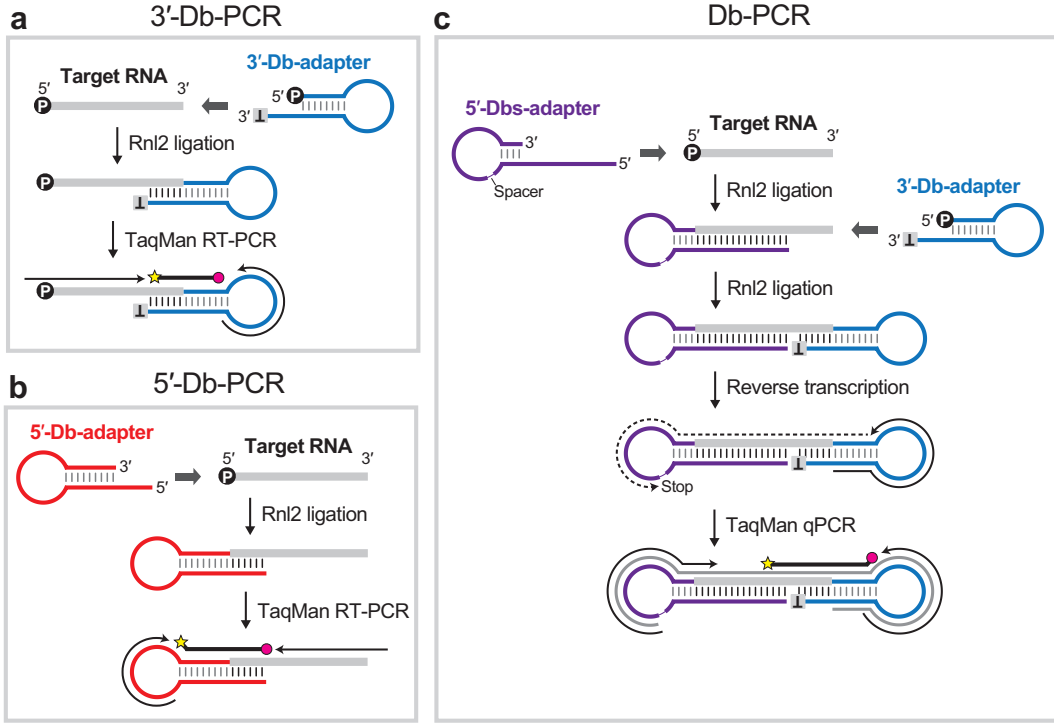


Fig. 1 Schematic representation of 3′-Db-PCR (**a**), 5′-Db-PCR (**b**), and Db-PCR (**c**)

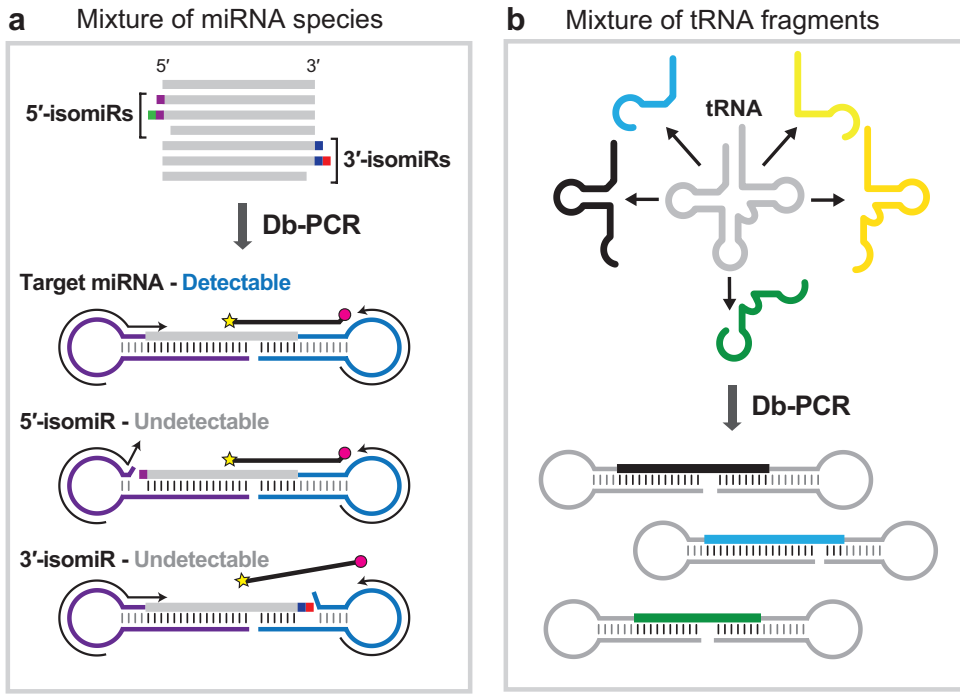


Fig. 2 Utilization of Db-PCR for specific quantification of isomiRs (**a**) and tRNA fragments (**b**)

2 Materials

1. Total cellular RNA extracted by an acid guanidinium-phenol reagent, such as TRIsure (Bioline; *see* **Note 1**).
2. 10× Annealing buffer: 50 mM Tris–HCl (pH 8.0) and 100 mM $MgCl_2$.
3. T4 RNA Ligase 2 [Rnl2; New England Biolabs (NEB)].
4. 10× T4 Rnl2 reaction buffer (supplied with Rnl2, NEB).
5. dNTPs, 10 mM of each (Thermofisher Scientific).
6. RNasin Ribonuclease Inhibitor (Promega).
7. SuperScript III Reverse Transcriptase (Invitrogen).
8. 5× First-strand buffer (supplied with SuperScript III, Invitrogen).
9. 100 mM DTT (supplied with SuperScript III, Invitrogen).
10. Premix Ex Taq (Probe qPCR), ROX Plus (TaKaRa).
11. MultiGen OptiMax Thermal Cycler (Labnet).
12. StepOne Plus Real-Time PCR system (Applied Biosystems).
13. Synthetic adapters and primers [synthesized by Integrated DNA Technologies (IDT)] (Table 1).

3 Methods

The outline of the procedures is shown in Fig. 1. The described conditions are for discriminative quantification of miR-16 and its variant whose representative results are shown in Fig. 3. Unless otherwise indicated, all solutions and materials should be kept on ice during the procedure. Also, RNase-free solutions, tubes, and pipettes should be used throughout.

3.1 3′-Db-PCR

3.1.1 3′-Db-Adapter Ligation

1. Mix 1 μg of cellular total RNA with 20 pmol of the 3′-Db-adapter (9-μL volume) in a 1.5 mL centrifuge tube.
2. Incubate at 90 °C for 3 min, and immediately add 1 μL of 10× Annealing buffer.
3. Incubate at 37 °C for 20 min to allow annealing and hybridization to occur.
4. Add 2 μL of 10× Ligase buffer, 1 U of Rnl2, and RNase-free water to make 20-μL volume solution.
5. Incubate at 37 °C for 1 h, followed by overnight incubation at 4 °C.

3.1.2 Reverse Transcription

1. Mix 1 μL of the ligated RNA with 0.5 μL of dNTP (10 mM each), 1 μL of 5 μM RT primer, and 4.5 μL of RNase-free water (total 7 μL) in a 0.2 mL PCR tube.

Table 1
Sequences of adaptors and primers for Db-PCR methods to quantify miR-16 and its terminal variant miR-16-[5′+U/3′–G]

Method	Primer/adapter	Sequence (5′–3′; for detection of miR-16)	Sequence (5′–3′; for detection of miR-16-[5′+U/3′–G])
3′-Db-PCR	3′-Db-adapter	/5Phos/CTCAGTGC**AGGGTCCGAGGTATT** CGCACTGAGCGCCAA/3InvdT/ (*see* **Note 2**)	/5Phos/CTCAGTGC**AGGGTCCGAGGTATT**CGCA CTGAGGCCAA/3InvdT/
	RT primer	CTCAGTGCGAATACCTCGGACCCT	Same as left
	Forward primer	GCGCTAGCAGCACGTAAATAT	GCGCTAGCAGCACGTAAAT
	Reverse primer	CGAATACCTCGGACC	Same as left
	TaqMan probe	/56-FAM/TGGCGCTCA/ZEN/GTG/3IABkFQ/	/56-FAM/TTGGCCTCA/ZEN/GTG/3IABkFQ/
5′-Db-PCR	5′-Db-adapter	CTGCTACTCAGTGCGTGG**GAGGGTGTGTGG TCTTGCTTGGTG**TGCACTGrArG (*see* **Note 3**)	CTGCTAACTCAGTGCATGG**GAGGGTGTGTGGT CTTGCTTGGTGTG**CACTGrArG
	RT primer	GCGACGCCAATATTTACGTG	Same as left
	Forward primer	GAGGGTGTGTGGTCTT	Same as left
	Reverse primer	GCGACGCCAATATTTACGTG	Same as left
	TaqMan probe	/56-FAM/TGTGCACTG/ZEN/AGTAGCAG/3IABkFQ/	/56-FAM/TGCACTG/ZEN/AGTTAGCAG/3IABkFQ/
Db-PCR	5′-Dbs-adapter	TATTTACGTGCTGCTACGTC**G/idSp/ TGGAGTGTGTGCTTT**GArCrG (*see* **Note 4**)	ATTTACGTGCTGCTAACGTC**G/idSp/ TGGAGTGTGTGCTTT**GArCrG
	3′-Db-adapter[a]	/5Phos/CTCAGTGC**AGGGTCCGAGGTATTC** GCACTGAGCGCCAA/3InvdT/	/5Phos/CTCAGTGC**AGGGTCCGAGGTATTC**GC ACTGAGGCCAA/3InvdT/
	RT primer[a]	CTCAGTGCGAATACCTCGGACCCT	Same as left
	Forward primer	TGGAGTGTGTGCTTTGACGTAGC	TGGAGTGTGTGCTTTGACGTTA
	Reverse primer[a]	CGAATACCTCGGACC	Same as left
	TaqMan probe	/56-FAM/AAATATTGG/ZEN/CGCTCAGTGCA/3IABkFQ/	/56-FAM/AATATTGG/ZEN/CCTCAGTGCA/3IABkFQ/

A, G, C, and T designate DNA, whereas rA and rG designate RNA. Bolded letters indicate the nucleotides of a loop region, while underlined letters indicate the nucleotides of a protruding region that hybridizes to target RNA

[a]These adapter/primers for Db-PCR are the same as those used for 3′-Db-PCR

2. Incubate at 90 °C for 2 min using a thermal cycler, and then immediately cool on ice.
3. Add 0.5 μL of 100 mM DTT, 2 μL of 5× First-strand buffer, 0.25 μL (10 U) of RNase inhibitor, and 0.25 μL (50 U) of SuperScript III Reverse Transcriptase (total 10 μL).
4. Using a thermal cycler, incubate at 55 °C for 60 min, followed by incubation at 70 °C for 15 min to deactivate the enzyme.

3.1.3 TaqMan PCR

1. Dilute the cDNA solution to 1:5. Mix 1.5 μL of the diluted cDNA with 2 pmol each of the forward and reverse primers, 100 nM of TaqMan probe (*see* **Note 5**), 5 μL of 2× Premix Ex Taq reaction solution, and RNase-free water (total 10 μL).
2. Using a real-time PCR machine, incubate the mixture at 95 °C for 20 s, followed by 30–40 cycles of 95 °C for 1 s and 60 °C (*see* **Note 5**) for 20 s.
3. Estimate the quantity of target RNAs based on Ct values.

3.2 5′-Db-PCR

3.2.1 5′-Db-Adapter Ligation

1. Mix 1 μg of cellular total RNA with 20 pmol of the 5′-Db-adapter (9-μL volume) in a 1.5 mL centrifuge tube.
2. Perform the ligation procedures as described in the **steps 2–5** of Subheading 3.1.1.

3.2.2 Reverse Transcription and TaqMan PCR

1. Perform the RT procedures as described in the **steps 1–4** of Subheading 3.1.2.
2. Perform the PCR procedures as described in the **steps 1–3** of Subheading 3.1.3.

3.3 Db-PCR

3.3.1 5′-Dbs-Adapter Ligation, Followed by 3′-Db-Adapter Ligation

1. Mix 1 μg of cellular total RNA with 20 pmol of the 5′-Dbs-adapter (9-μL volume) in a 1.5 mL centrifuge tube.
2. Perform the annealing procedures as described in the **steps 2–4** of Subheading 3.1.1.
3. Incubate at 37 °C for 30 min.
4. Add 20 pmol of 3′-Db-adapter (1 μL), and then incubate the total 21-μL volume reaction mixture at 16 °C for 30 min, followed by overnight incubation at 4 °C.

3.3.2 Reverse Transcription and TaqMan PCR

1. Perform the RT procedures as described in the **steps 1–4** of Subheading 3.1.2.
2. Perform the PCR procedures as described in the **steps 1–3** of Subheading 3.1.3.

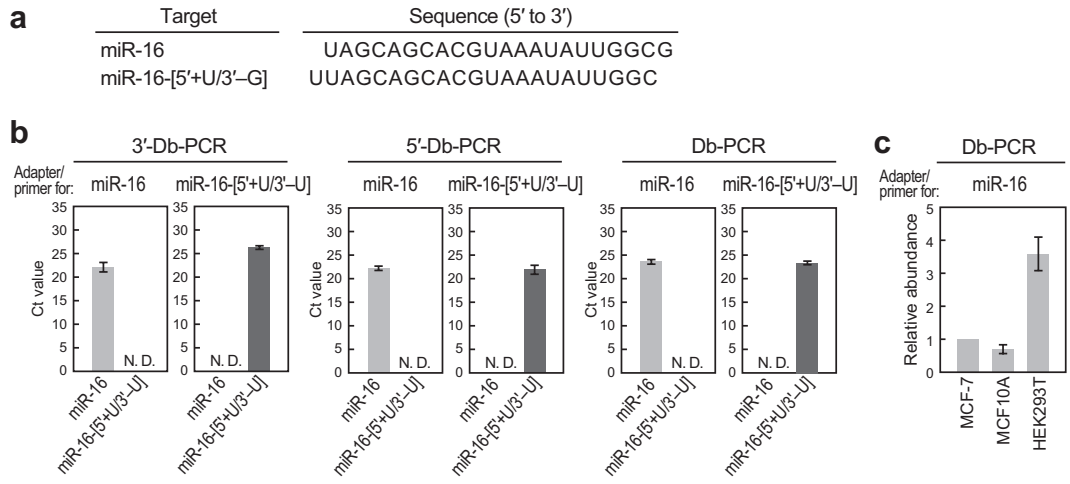


Fig. 3 Db-PCR to specifically quantify miR-16 and its variant. (**a**) The Db-PCR schemes were evaluated by targeting human miR-16, a widely expressed miRNA, and its 5′-/3′-variant, miR-16-[5′+U/3′–G], that contains an additional U at the 5′-end while lacking a G at the 3′-end. (**b**) 3′-Db-PCR (*left*), 5′-Db-PCR (*middle*), and Db-PCR (*right*) were employed for the detection of synthetic miR-16 and its variant (20 fmol). The data represent the average Ct values from three independent experiments, with bars showing the SD. miR-16, but not its variant, was specifically detected by all the Dumbbell methods in which the adapter/primer/probe set for miR-16 was utilized. When the adapter/primer/probe set for the variant was utilized, the variant, but not miR-16, was specifically detected. (**c**) Db-PCR was employed for the detection of miR-16 and its variant expressed in MCF-7, MCF10A, and HEK293T cells. U6 snRNA expression was also quantified for use as an internal control. This utilized SsoFast EvaGreen Supermix (BioRad) and the forward (5′-TCGCTTCGGCAGCACATATAC-3′) and reverse (5′-CGAATTTGCGTGTCATCCTTG-3′) primers. The abundance in MCF-7 cells was defined as 1, and the average relative abundance from three independent experiments are shown with SD bars. miR-16 was differentially expressed in the three cells, while its variant did not give a detectable signal in all the examined cells

4 Notes

1. For Rnl2-catalyzed ligation of Db-adapters, the target RNA should contain a 5′-P (for 5′-Db-PCR and Db-PCR) and/or 3′-OH ends (for 3′-Db-PCR and Db-PCR) (Fig. 1). If the target RNA is expected to contain a different terminal structure, the total RNA should be subjected to dephosphorylation/phosphorylation treatment using T4 Polynucleotide Kinase (T4 PNK) in the presence of ATP prior to adapter ligation.
2. The 3′-Db-adapter comprises DNA forming a stem-loop structure with the loop sequences identical to those of the stem-loop primer for the miRNA quantification method [8]. The protruding 3′-terminus comprises 5–7 nucleotides complementary to and thereby hybridize with the 3′-terminal nucleotides of the target RNA [10]. After designing the adapter for the desired target RNA, the overall secondary structure of the adapter should be checked using mfold [15] or

other programs to confirm the formation of the stem-loop structure. The adapter contains 5′-P for Rnl2 ligation and 3′-inverted dT to prevent undesired nucleotide synthesis from the 3′-end of the adapter during RT-PCR.

3. The 5′-Db-adapter forms a stem-loop structure with 6–7 nucleotides protruding from the 5′-end to hybridize the 5′-terminal nucleotides of the target RNA [10]. The Db-adapter contains both 5′- and 3′-OH ends and comprises DNA, with the exception of the last two 3′-terminal nucleotides that are designed as RNA. The hybridization of the 5′-Db-adapter with the target RNA generates double-stranded DNA/RNA hybrids containing a nick of "RNA-OH-3′/5′-P-RNA" between the 3′-end of the adapter and 5′-end of the target. This is an efficient substrate for Rnl2 ligation [11–13].
4. The 5′-Dbs-adapter has a similar structure to the 5′-Db-adapter but contains a base-lacking 1′,2′-dideoxyribose spacer in the loop region, which terminates reverse transcription at its preceding nucleotide and thereby prevents generation of highly structured cDNAs for efficient PCR reaction. The length of the protruding 5′-end of the adapter affects detection efficiency; the longer the protruding 5′-end, the more efficient the detection is [10]. This probably occurs because of the high stability of the adapter–target RNA duplex. The 5′-Dbs-adapter with a 16-nt protruding end was used in this study and in our previous study [10].
5. The indicated conditions are for discriminative quantification of miR-16 and its variant whose results are shown in Fig. 3. Detection efficiency and specificity are dependent on the design and concentration of the TaqMan probe as well as the annealing/polymerization temperature in real-time PCR reactions. Longer length/higher concentration of the TaqMan probe and lower annealing/polymerization temperature result in more efficient detection with lower Ct values. However, these have the consequence of a higher probability of undesirable cross-reactions with non-targeted RNA containing similar sequences. The typical length of TaqMan probe is 12–20 nt. Preliminary experiments are required in advance to determine the appropriate concentration to use for the TaqMan probe (typically between 100–400 nM) and the annealing/polymerization temperature (typically between 50–65 °C) for efficient detection and specific discrimination of target RNAs.

Acknowledgments

This study was supported by NIH grant (GM106047 to YK).

References

1. Neilsen CT, Goodall GJ, Bracken CP (2012) IsomiRs–the overlooked repertoire in the dynamic microRNAome. Trends Genet 28(11): 544–549. doi:10.1016/j.tig.2012.07.005. S0168-9525(12)00112-6 [pii]
2. Somel M, Guo S, Fu N, Yan Z, Hu HY, Xu Y, Yuan Y, Ning Z, Hu Y, Menzel C, Hu H, Lachmann M, Zeng R, Chen W, Khaitovich P (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. Genome Res 20(9):1207–1218. doi:10.1101/gr.106849.110. gr.106849.110 [pii]
3. Fernandez-Valverde SL, Taft RJ, Mattick JS (2010) Dynamic isomiR regulation in drosophila development. RNA 16(10):1881–1888. doi:10.1261/rna.2379610. rna.2379610 [pii]
4. Bizuayehu TT, Lanes CF, Furmanek T, Karlsen BO, Fernandes JM, Johansen SD, Babiak I (2012) Differential expression patterns of conserved miRNAs and isomiRs during Atlantic halibut development. BMC Genomics 13:11. doi:10.1186/1471-2164-13-11. 1471-2164-13-11 [pii]
5. Tan GC, Chan E, Molnar A, Sarkar R, Alexieva D, Isa IM, Robinson S, Zhang S, Ellis P, Langford CF, Guillot PV, Chandrashekran A, Fisk NM, Castellano L, Meister G, Winston RM, Cui W, Baulcombe D, Dibb NJ (2014) 5′ isomiR variation is of functional and evolutionary importance. Nucleic Acids Res 42(14): 9424–9435. doi:10.1093/nar/gku656. gku656 [pii]
6. Loher P, Londin ER, Rigoutsos I (2014) IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies. Oncotarget 5(18):8790–8802. doi:10.18632/oncotarget.2405. 2405 [pii]
7. Telonis AG, Loher P, Jing Y, Londin E, Rigoutsos I (2015) Beyond the one-locus-one-miRNA paradigm: microRNA isoforms enable deeper insights into breast cancer heterogeneity. Nucleic Acids Res 43(19):9158–9175. doi:10.1093/nar/gkv922. gkv922 [pii]
8. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 33(20):e179. doi:10.1093/nar/gni178. 33/20/e179 [pii]
9. Schamberger A, Orban TI (2014) 3′ IsomiR species and DNA contamination influence reliable quantification of microRNAs by stem-loop quantitative PCR. PLoS One 9(8):e106315. doi:10.1371/journal.pone.0106315. PONE-D-13-49830 [pii]
10. Honda S, Kirino Y (2015) Dumbbell-PCR: a method to quantify specific small RNA variants with a single nucleotide resolution at terminal sequences. Nucleic Acids Res 43(12):e77. doi:10.1093/nar/gkv218. gkv218 [pii]
11. Bullard DR, Bowater RP (2006) Direct comparison of nick-joining activity of the nucleic acid ligases from bacteriophage T4. Biochem J 398(1):135–144. doi:10.1042/BJ20060313. BJ20060313 [pii]
12. Nandakumar J, Ho CK, Lima CD, Shuman S (2004) RNA substrate specificity and structure-guided mutational analysis of bacteriophage T4 RNA ligase 2. J Biol Chem 279(30):31337–31347. doi:10.1074/jbc.M402394200. M402394200 [pii]
13. Nandakumar J, Shuman S (2005) Dual mechanisms whereby a broken RNA end assists the catalysis of its repair by T4 RNA ligase 2. J Biol Chem 280(25):23484–23489. doi:10.1074/jbc.M500831200. M500831200 [pii]
14. Ranade K, Chang MS, Ting CT, Pei D, Hsiao CF, Olivier M, Pesich R, Hebert J, Chen YD, Dzau VJ, Curb D, Olshen R, Risch N, Cox DR, Botstein D (2001) High-throughput genotyping with single nucleotide polymorphisms. Genome Res 11(7):1262–1268. doi:10.1101/gr.157801
15. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–3415

Chapter 5

MicroRNA Detection by Whole-Mount In Situ Hybridization in *C. elegans*

Yoshiki Andachi and Yuji Kohara

Abstract

MicroRNAs (miRNAs) loaded on argonaute proteins guide RNA-induced silencing complexes to target mRNAs. An excellent method to decipher the spatiotemporal expression patterns of miRNAs is whole-mount in situ hybridization (WISH), which has been successfully used in vertebrate embryos but still remains unavailable for many animal species. Here, we describe a WISH method for miRNA detection in *Caenorhabditis elegans* at both embryonic and post-embryonic stages. Strategies devised for detection include fixation of animals with carbodiimide at a high temperature and subsequent partial digestion of the fixed animals with an extremely high concentration of proteinase. WISH signals are visualized by staining with a chromogenic substrate or a fluorescent dye.

Key words microRNA, In situ hybridization, *C. elegans*, Fixation, Chromogenic staining, Fluorescent staining

1 Introduction

MicroRNAs (miRNAs) are RNA components of the RNA-induced silencing complex (RISC) that contains argonaute proteins, and play a central role in targeting the RISC to mRNAs bearing a sequence complementary to the miRNAs. Expression analysis of miRNAs is indispensable to elucidate their functions. A technique for detecting cells that express miRNAs is in situ hybridization. Whole-mount in situ hybridization (WISH), which investigates the whole-body distribution of miRNAs, is especially effective for examining their spatiotemporal expression patterns. A WISH method for miRNA detection was initially developed in zebrafish embryos [1], and the method was quickly adapted for embryos of mouse, chicken, medaka, and fruit fly [2–5]. In the WISH method, oligonucleotide probes containing locked nucleic acids (LNAs) are used to acquire hybridization signals with high specificity [1]. LNA is a nucleotide analog in which the 2′-oxygen and 4′-carbon of the ribose moiety are connected via a methylene bridge that locks the

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_5,

ribose in the 3′-endo conformation, which enhances base stacking and backbone pre-organization [6]. In addition, the intensity of hybridization signals is increased by fixation of sectioned mouse tissues with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) at a high temperature [7, 8]. EDC reacts with phosphates at the 5′ end of miRNAs to introduce phosphoramidate linkages, leaving all miRNA nucleotides accessible to probes for hybridization [7]. Cross-linking with EDC, however, acts not only on miRNAs but also on proteins to form peptide bonds [8, 9]. Fixed samples are usually subjected to partial digestion with proteinase to enable WISH reagents, such as probes, to penetrate inward. The epidermis of *Caenorhabditis elegans* is surrounded by cuticle layers composed of collagens, which start to be produced at the late stage of embryogenesis. We discovered that fixation of *C. elegans* with EDC at a high temperature required digestion of the fixed animals with an extremely high concentration of proteinase to reveal WISH signals [10]. WISH methods for mRNA detection have already been established in *C. elegans* at the embryo, larval, and adult stages [11–13]. Based on these WISH methods, we developed a new WISH method for miRNA detection by incorporating these procedures, i.e., fixation with EDC at a high temperature, partial digestion with an extremely high concentration of proteinase, and hybridization with LNA probes [10]. A hundred or more animals stuck on a slide are analyzed all at once by WISH to show hybridization signals visualized by chromogenic staining (Fig. 1a, b) or

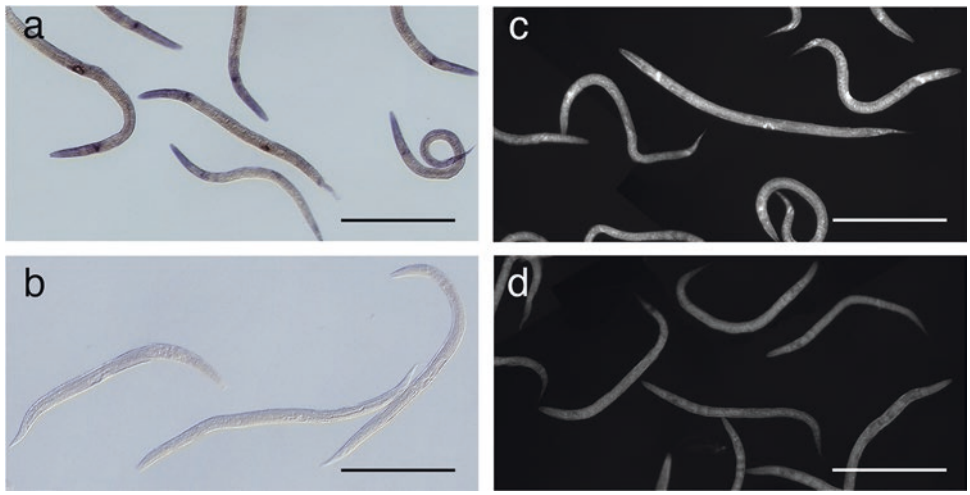


Fig. 1 WISH analysis for the detection of *mir-48* miRNA at the fourth larval stage. The analysis was conducted on a strain that contains the wild-type *mir-48* gene (**a**, **c**), and on one that is homozygous for the deletion of the *mir-48* gene (**b**, **d**). Hybridization signals were visualized by chromogenic staining (**a**, **b**) and fluorescent staining (**c**, **d**) Scale bars are 500 μm. The sequence of the probe used for the detection of *mir-48* miRNAs is 5′-(DIG)TCG**C**AT**C**TA**C**TG**A**GC**C**TA**C**C**T**CA-3′, in which DIG shows 5′-labeled digoxigenin, and LNA nucleotides are indicated in bold. The hybridization temperature in using the probe is 58 °C

fluorescent staining (Fig. 1c, d). Subsequently, another in situ hybridization method has been reported in which gonads and embryos are released from adults and then freeze-cracked on dry ice to increase the accessibility of miRNAs [14].

2 Materials

2.1 Culture

1. LB broth: Dissolve 10 g of tryptone, 5 g of yeast extract, 10 g NaCl in 1 L of deionized H_2O. Adjust pH to 7.0 with 4 M NaOH. Sterilize by autoclaving.
2. S basal: 50 mM potassium phosphate, pH 6.0, 100 mM NaCl, 5 mg/L cholesterol. Sterilize by autoclaving.
3. NGM agar: Dissolve 2.5 g of peptone, 3 g of NaCl, 5 mg of cholesterol, 17 g of agar in 1 L of deionized H_2O by autoclaving. Autoclave the following components separately, and add them to the above solution using sterile technique: 25 mL of 1 M potassium phosphate, pH 6.0, 1 mL of 1 M $CaCl_2$, 1 mL of 1 M $MgSO_4$.
4. Trace metals solution: 5 mM EDTA, 2.5 mM $FeSO_4$, 1 mM $MnCl_2$, 1 mM $ZnSO_4$, 0.1 mM $CuSO_4$. Sterilize by autoclaving. Store in the dark.
5. S medium: Autoclave the following components separately, and add them to 1 L of S basal using sterile technique: 10 mL of 1 M potassium citrate, pH 6.0, 10 mL of trace metals solution, 3 mL of 1 M $CaCl_2$, 3 mL of 1 M $MgSO_4$.
6. M9 buffer: 86 mM NaCl, 42 mM Na_2HPO_4, 22 mM KH_2PO_4, 1 mM $MgSO_4$. Sterilize by autoclaving.
7. Alkaline hypochlorite solution: 250 mM KOH, 12% sodium hypochlorite solution.
8. Sucrose solution: Dissolve 600 g of sucrose in 400 mL of deionized H_2O.

2.2 Fixation

1. Phosphate-free fixative: 80 mM HEPES, pH 6.9, 1.6 mM $MgSO_4$, 0.8 mM EGTA, 100 mM NaCl, 3.7% formaldehyde.
2. EDC buffer: 130 mM 1-methylimidazole, 300 mM NaCl. Adjust pH to 8.0 with 6 M HCl.
3. EDC fixative: EDC buffer containing 160 mM EDC (*see* **Note 1**).
4. BO_3 buffer: 25 mM sodium borate, pH 9.0.
5. PBS: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4.
6. PBT: PBS containing 0.1% Tween-20.
7. Proteinase K solution: Dilute conc. proteinase K stock solution in PBT containing 1% β-mercaptoethanol.

8. Fixative: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 80 mM HEPES, pH 6.9, 1.6 mM $MgSO_4$, 0.8 mM EGTA, 3.7% formaldehyde.
9. Poly-L-lysine solution: Dilute 0.1% poly-L-lysine stock solution ten times with deionized H_2O.
10. Slide: Load 100 μL of poly-L-lysine solution into each well of a 3-well teflon printed diagnostic slide, and keep at room temperature for 1 h (*see* **Note 2**). After removal of the solution, heat the slide at 60 °C for 1 h.
11. Basal embryo handling buffer: 300 mM mannitol, 50 mM HEPES, pH 7.2, 10 mM NaCl, 10 mM $MgCl_2$.
12. Embryo handling buffer: basal embryo handling buffer containing 1 mM EGTA, 2 mM NH_4NO_3, 0.1% gelatin.
13. Yatalase solution: Dissolve powdered yatalase in embryo handling buffer at a working concentration of 15 mg/mL (*see* **Note 3**). Dispense into small aliquots and store at −20 °C.
14. TBS: 10 mM Tris, pH 7.5, 100 mM NaCl.

2.3 Hybridization

1. Hybridization buffer: 50% formamide, 75 mM sodium citrate, pH 7.0, 750 mM NaCl, 100 μg/mL heparin, 2% blocking reagent (*see* **Note 4**).
2. Probe mixture: Hybridization buffer containing 10 nM digoxigenin-labeled LNA probe. Heat the mixture at 95 °C for 2 min, and then cool on ice.
3. SSC dilution buffer: 3 mM sodium citrate, pH 7.0, 30 mM NaCl.
4. TN: 100 mM Tris, pH 7.5, 150 mM NaCl.
5. TNT: TN containing 0.5% Tween-20.
6. TNB: TN containing 0.5% blocking reagent (*see* **Note 5**).
7. Anti-digoxigenin-phosphatase solution: Dilute conc. anti-digoxigenin-AP, Fab fragments solution 2500 times with TNB (*see* **Note 6**).
8. Staining buffer: 100 mM Tris, pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 1 mM levamisole, 0.1% Tween-20.
9. NBT/BCIP reaction mixture: Dilute 50× conc. nitroblue tetrazolium chloride (NBT)/5-bromo-4-chloro-3-indolyl-phosphate (BCIP) stock solution 50 times with staining buffer (*see* **Note 7**).
10. Mounting medium: 90% glycerol, 10 mM Tris, pH 7.4, 100 mM NaCl, 1% propyl gallate, 1 mM levamisole.
11. Anti-digoxigenin-peroxidase solution: Dilute conc. anti-digoxigenin-POD, Fab fragments solution 500 times with TNB (*see* **Note 8**).
12. TSA-Cy3 reaction mixture: Dilute cyanine 3 amplification reagent 50 times with 1× plus amplification diluent (*see* **Note 9**).

3 Methods

3.1 Synchronized Population of C. elegans

1. Cultivate *Escherichia coli* strain OP50, which is used as food for *C. elegans*, in a 5-L conical flask containing 1 L of LB broth at 37 °C on a rotary shaker until stationary phase.
2. After cooling on ice, collect the bacteria by centrifugation at 5000 × *g* for 15 min at 4 °C, wash the pellet twice by resuspension in an appropriate amount of S basal at 4 °C, and finally suspend the pellet in 100 mL of S basal at 4 °C (*see* **Note 10**).
3. Grow *C. elegans* strain N2 on an NGM agar plate with a lawn of *E. coli* strain OP50 at 20 °C until the food supply is exhausted.
4. Collect animals by washing the plate with 1 mL of S basal. Grow them in a 500-mL conical flask containing 80 mL of S medium and 20 mL of the bacterial suspension at 20 °C on a rotary shaker until many animals grow up to be gravid adults.
5. Collect animals in liquid culture by centrifugation at 800 × *g* for 1 min, and suspend them in 5 mL of M9 buffer.
6. Mix the animal suspension with 10 mL of alkaline hypochlorite solution, and leave to stand with occasional agitation for approximately 7 min until the animals are dissolved (*see* **Note 11**).
7. Collect remaining embryos by centrifugation at 800 × *g* for 1 min, wash them three times by resuspension in an appropriate amount of M9 buffer, and finally suspend them in 1 mL of S basal.
8. Grow the embryos in liquid culture as in **step 4** until animals reach the developmental stage of interest (*see* **Note 12**).
9. Collect animals by centrifugation, suspend them in 25 mL of M9 buffer chilled at 4 °C, and transfer the animal suspension to a new 50-mL centrifuge tube (*see* **Note 13**).
10. Mix the animal suspension with 25 mL of sucrose solution chilled at 4 °C, and immediately centrifuge the mixture at 800 × *g* for 5 min at 4 °C.
11. Transfer animals floating on the top of the mixture to a 500-mL conical flask containing 200 mL of M9 buffer, and incubate for 20 min at 20 °C on a rotary shaker.
12. Collect animals by centrifugation, and wash them three times by resuspension in an appropriate amount of M9 buffer.

3.2 Fixation of Larvae and Adults

1. Transfer approximately 300 μL of animals to a new 2-mL siliconized microcentrifuge tube (*see* **Note 14**).
2. In the following procedures, incubation is carried out in 1 mL of a specified solution at room temperature for 2 min with continuous rotation of the tube, and is terminated by centrifugation at 400 × *g* for 1 min followed by removal of the supernatant fluid by pipetting, unless otherwise indicated.

3. 80% methanol, 20% dimethyl sulfoxide chilled at −20 °C. Incubate at 4 °C for 10 min.
4. Methanol chilled at 4 °C. Incubate at 4 °C.
5. 90% methanol, 10% phosphate-free fixative chilled at 4 °C. Incubate at 4 °C (*see* **Note 15**).
6. 70% methanol, 30% phosphate-free fixative chilled at 4 °C. Incubate at 4 °C.
7. 50% methanol, 50% phosphate-free fixative chilled at 4 °C. Incubate at 4 °C.
8. 30% methanol, 70% phosphate-free fixative chilled at 4 °C. Incubate at 4 °C.
9. Phosphate-free fixative chilled at 4 °C. Incubate at 4 °C.
10. Phosphate-free fixative chilled at 4 °C. Incubate at 4 °C for 20 min.
11. EDC buffer. Repeat three times.
12. EDC fixative. Incubate at 60 °C for 1 h (*see* **Note 16**).
13. BO_3 buffer. Repeat three times.
14. BO_3 buffer containing 10 mM dithiothreitol. Incubate for 20 min.
15. BO_3 buffer. Repeat three times.
16. BO_3 buffer containing 0.6% H_2O_2. Incubate for 20 min.
17. PBT. Repeat three times.
18. Proteinase K solution at a concentration of 2 mg/mL (*see* **Note 16**). Incubate at 37 °C for 20 min.
19. PBS. Repeat three times.
20. Fixative. Incubate for 20 min.
21. PBS. Repeat three times.
22. Suspend the fixed animals in 1 mL of PBS.
23. After loading 200 μL of PBS into each poly-L-lysine-coated well of a Teflon printed diagnostic slide, add 20 μL of the animal suspension to the PBS (*see* **Note 17**). Leave to stand for 5 min, and then remove most of the supernatant fluid by pipetting.
24. In the following procedures, animals stuck on the slide are immersed in a specified solution at room temperature for 5 min, unless otherwise indicated.
25. 30% ethanol, 70% PBS.
26. 50% ethanol, 50% PBS.
27. 70% ethanol, 30% PBS.
28. 90% ethanol.
29. Ethanol. Repeat twice (*see* **Note 18**).

3.3 Hybridization of Larvae and Adults

1. In the following procedures, fixed animals stuck on the slide are immersed in a specified solution at room temperature for 5 min, unless otherwise indicated.
2. Ethanol containing 24 mM HCl.
3. 90% ethanol.
4. 70% ethanol, 30% PBS.
5. 50% ethanol, 50% PBS.
6. 30% ethanol, 70% PBS.
7. PBT. Repeat twice.
8. Wipe off excess fluid on the slide (*see* **Note 19**). Load 100 μL of proteinase K solution at a concentration of 0.5 mg/mL into each well of the slide (*see* **Note 16**). Incubate at 37 °C for 20 min in a moist chamber (*see* **Note 20**).
9. PBS. Repeat twice.
10. Fixative. Incubate for 20 min.
11. PBS. Repeat twice.
12. PBS containing 27 mM glycine.
13. PBS.
14. 50% hybridization buffer, 50% PBS. Incubate for 10 min.
15. Hybridization buffer. Incubate for 10 min.
16. Wipe off excess fluid on the slide (*see* **Note 19**). Load 100 μL of hybridization buffer into each well of the slide, and place the slide in a moist chamber (*see* **Note 21**). Incubate for 1 h at a hybridization temperature suitable for a probe (*see* **Note 22**).
17. Exchange the buffer with 100 μL of probe mixture. Incubate for 2 h in a moist chamber at the hybridization temperature (*see* **Note 21**).
18. Hybridization buffer heated at the hybridization temperature. Incubate for 10 min at the hybridization temperature with continuous agitation. Repeat twice.
19. SSC dilution buffer heated at the hybridization temperature. Incubate for 10 min at the hybridization temperature with continuous agitation. Repeat twice.
20. TNT. Repeat twice.
21. Wipe off excess fluid on the slide (*see* **Note 19**). Load 100 μL of TNB into each well of the slide. Incubate for 30 min in a moist chamber (*see* **Note 20**).
22. For staining with a chromogenic substrate, proceed from **steps 23** to **29**. For staining with a fluorescent dye, proceed from **steps 30** to **36**.
23. Exchange the TNB with 100 μL of anti-digoxigenin-phosphatase solution. Incubate for 2 h in a moist chamber (*see* **Note 20**).

24. TNT. Incubate for 10 min with continuous agitation. Repeat four times.
25. Staining buffer. Repeat twice.
26. Wipe off excess fluid on the slide (*see* **Note 19**). Load 100 μL of the NBT/BCIP reaction mixture into each well of the slide. Incubate in a moist chamber in the dark for a few hours to a few days until chromogenic staining appears in the animals (*see* **Note 20**).
27. TN containing 20 mM EDTA. Repeat twice.
28. TN containing 0.1 mM EDTA. Repeat twice.
29. Wipe off excess fluid on the slide. After loading 20 μL of mounting medium, place a cover glass on the slide (*see* **Note 23**).
30. Exchange the TNB with 100 μL of anti-digoxigenin-peroxidase solution. Incubate for 1 h in a moist chamber (*see* **Note 20**).
31. TNT. Incubate for 10 min with continuous agitation. Repeat four times.
32. Wipe off excess fluid on the slide (*see* **Note 19**). Load 50 μL of the TSA-Cy3 reaction mixture into each well of the slide. Incubate for 5 min.
33. TNT. Repeat twice.
34. TN containing 0.02 μg/mL 4′,6-diamidino-2-phenylindole (DAPI).
35. TN. Repeat twice.
36. Wipe off excess fluid on the slide. After loading 20 μL of mounting medium, place a cover glass on the slide (*see* **Note 23**).

3.4 Fixation of Embryos

1. Transfer approximately 100 μL of embryos to a new 1.5-mL siliconized microcentrifuge tube (*see* **Note 24**).
2. Add 100 μL of yatalase solution. Incubate for 75 s with vortexing.
3. Suspend the embryos in 1 mL of embryo handling buffer. Centrifuge the embryo suspension at 800×*g* for 1 min, and then remove the supernatant fluid by pipetting. Repeat three times.
4. Suspend the embryos in 1 mL of basal embryo handling buffer. Centrifuge the embryo suspension at 800×*g* for 1 min, and then remove the supernatant fluid by pipetting.
5. Suspend the embryos in 1 mL of basal embryo handling buffer (*see* **Note 25**).
6. After loading 200 μL of basal embryo handling buffer into each poly-L-lysine-coated well of a Teflon printed diagnostic

slide, add 20 μL of the embryo suspension to the buffer (*see* **Note 17**). Leave to stand for 10 min, and then remove most of the supernatant fluid by pipetting.

7. In the following procedures, embryos stuck on the slide are immersed in a specified solution at room temperature for 5 min, unless otherwise indicated.
8. Methanol chilled at −20 °C. Incubate at 4 °C for 10 min.
9. Methanol.
10. 90% methanol, 10% phosphate-free fixative (*see* **Note 15**).
11. 70% methanol, 30% phosphate-free fixative.
12. 50% methanol, 50% phosphate-free fixative.
13. 30% methanol, 70% phosphate-free fixative.
14. Phosphate-free fixative.
15. Phosphate-free fixative. Incubate for 20 min.
16. 30% ethanol, 70% TBS.
17. 50% ethanol, 50% TBS.
18. 70% ethanol, 30% TBS.
19. 90% ethanol, 10% TBS.
20. Ethanol. Repeat twice (*see* **Note 18**).

3.5 Hybridization of Embryos

1. In the following procedures, embryos stuck on the slide are immersed in a specified solution at room temperature for 5 min, unless otherwise indicated.
2. Ethanol containing 0.6% H_2O_2.
3. 90% ethanol, 10% TN.
4. 70% ethanol, 30% TN.
5. 50% ethanol, 50% TN.
6. 30% ethanol, 70% TN.
7. TN.
8. EDC buffer. Repeat twice.
9. EDC fixative heated at 60 °C. Incubate at 60 °C for 20 min (*see* **Note 16**).
10. PBT. Repeat twice.
11. Wipe off excess fluid on the slide (*see* **Note 19**). Load 100 μL of proteinase K solution at a concentration of 2 mg/mL into each well of the slide (*see* **Note 16**). Incubate at 37 °C for 20 min in a moist chamber (*see* **Note 20**).
12. After this step, carry out the procedures from **step 9** in Subheading 3.3.

4 Notes

1. EDC fixative should be prepared just before use.
2. Three-well Teflon printed diagnostic slides should be washed with laboratory cleaning solution before use in order to ensure hydrophilic surface conditions of the wells.
3. Yatalase degrades the chitin layer of the eggshell. We use yatalase (TakaraBio, Shiga, Japan) for this purpose.
4. Formamide should be stirred with anion- and cation-exchange resin for 30 min to deionize, and the resin should be removed by paper filtration. We use Blocking Reagent (Roche Diagnostics GmbH, Mannheim, Germany) for this purpose.
5. We use Blocking Reagent (PerkinElmer, Boston, MA, USA) for this purpose.
6. We use Anti-digoxigenin-AP, Fab fragments (Roche Diagnostics GmbH) for this purpose.
7. We use DIG Nucleic Acid Detection Kit (Roche Diagnostics GmbH) for this purpose.
8. We use Anti-Digoxigenin-POD, Fab fragments (Roche Diagnostics GmbH) for this purpose.
9. We use TSA Plus Cyanine three System (PerkinElmer) for this purpose.
10. Vigorous shaking is required to suspend the bacteria pellet. The bacterial suspension can be stored at 4 °C for at least 1 month.
11. Incubation for longer than the specified time leads to a reduction in the viability of embryos.
12. If the amount of animals is not sufficient for use, the animals are allowed to grow up to be gravid adults, and the procedures from **steps 5** to **8** in Subheading 3.1 are repeated so as to prepare animals at the next generation.
13. Unless the animals are cooled enough, the amount of animals floating at **step 11** in Subheading 3.1 decreases.
14. The amount of animals required for fixation indicates the volume at the time when they are precipitated by centrifugation. If too few animals are used, most of the animals are lost during the following procedures. On the other hand, too many animals may prevent the fixation reaction from proceeding enough.
15. The use of solutions containing phosphate ions should be avoided because the ions interfere with the cross-linking reaction.
16. The conditions of fixation with EDC and digestion with proteinase K should be optimized according to experimental

design. Fixation at a higher temperature or for a longer time introduces cross-linking into more miRNAs, resulting in their retention in fixed animals. Nevertheless, the intensity of hybridization signals decreases, possibly due to a reduction of the permeability of cuticle layers. Digestion with proteinase K at a higher concentration, at a higher temperature up to 60 °C or for a longer time increases the permeability, resulting in intense hybridization signals. However, too much digestion causes poor morphology and leads to high background signals. The first choice for optimization is to conduct the digestion at different concentrations of proteinase K in each well on a slide.

17. To evenly distribute fixed animals in the wells, the micropipette holding the animal suspension is held almost horizontally, and the suspension should be pushed from the side into the buffer in the wells.
18. The fixed animals stuck on the slide can be stored at −20 °C for at least 1 month.
19. After excess fluid is wiped off, the surface of Teflon printed diagnostic slides is water repellent.
20. There is no need to overlay a lid on the wells during the incubation. As a moist chamber, we use a plastic box with snap-tight lid in which paper towels wetted with deionized H_2O are laid on the bottom, and a slide holder is put on the paper towels.
21. There is no need to overlay a lid on the wells during the incubation. As a moist chamber, we use a plastic box with snap-tight lid in which paper towels wetted with a buffer composed of 50% deionized formamide, 75 mM sodium citrate, pH 7.0, and 750 mM NaCl are laid on the bottom, and a slide holder is put on the paper towels.
22. The hybridization temperature for a probe is experimentally determined as follows. The hybridization temperature is tentatively set at 20–25 °C below the melting temperature of the probe that is calculated by the probe manufacturer. WISH experiments are carried out at the hybridization temperature and at slightly different (both higher and lower) hybridization temperatures. A hybridization temperature is chosen at which tissue-specific signals are detected in wild-type animals but are lost in their deletion mutants.
23. For permanent storage, four sides of cover glass need to be sealed with nail polish to prevent drying out.
24. The amount of embryos required for fixation indicates the volume at the time when they are precipitated by centrifugation.
25. The efficacy of the degradation of chitin layers is verified by the presence of a few percent of late-stage embryos that have escaped from the eggshell.

Acknowledgement

This work was supported by JSPS KAKENHI Grant number 22241047 and 25250025 to Y.K.

References

1. Wienholds E, Kloosterman WP, Miska E et al (2005) MicroRNA expression in zebrafish embryonic development. Science 309:310–311
2. Sokol NS, Ambros V (2005) Mesodermally expressed drosophila microRNA-1 is regulated by twist and is required in muscles during larval growth. Genes Dev 19:2343–2354
3. Ason B, Darnell DK, Wittbrodt B et al (2006) Differences in vertebrate microRNA expression. Proc Natl Acad Sci U S A 103:14385–14389
4. Darnell DK, Kaur S, Stanislaw S et al (2006) MicroRNA expression during chick embryo development. Dev Dyn 235:3156–3165
5. Kloosterman WP, Wienholds E, de Bruijn E et al (2006) In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. Nat Methods 3:27–29
6. Petersen M, Wengel J (2003) LNA: a versatile tool for therapeuticsand genomics. Trends Biotechnol 21:74–81
7. Pena JT, Sohn-Lee C, Rouhanifard SH et al (2009) miRNA in situ hybridization in formaldehyde and EDC-fixed tissues. Nat Methods 6:139–141
8. Renwick N, Cekan P, Masry PA et al (2013) Multicolor microRNA FISH effectively differentiates tumor types. J Clin Invest 123:2694–2702
9. Sung HW, Chang WH, Ma CY et al (2003) Crosslinking of biological tissues using genipin and/or carbodiimide. J Biomed Mater Res 64A:427–438
10. Andachi Y, Kohara Y (2016) A whole-mount in situ hybridization method for microRNA detection in Caenorhabditis elegans. RNA 22:1099–1106
11. Tabara H, Motohashi T, Kohara Y (1996) A multi-well version of in situ hybridization on whole mount embryos of Caenorhabditis elegans. Nucleic Acids Res 24:2119–2124
12. Motohashi T, Tabara H, Kohara Y (2006) Protocols for large scale in situ hybridization on C. elegans larvae. In: The C. elegans research community (ed) WormBook. http://www.wormbook.org. Accessed 24 Jul 2006
13. Motohashi T, Hirono K, Kohara Y (2015) In situ hybridization on whole mount embryos of C. elegans. In: The C. elegans research community (ed) WormBook. http://www.wormbook.org. Accessed 19 May 2015
14. McEwen TJ, Yao Q, Yun S et al (2016) Small RNA in situ hybridization in Caenorhabditis elegans, combined with RNA-seq, identifies germline-enriched microRNAs. Dev Biol 418: 248–257

Chapter 6

cCLIP-Seq: Retrieval of Chimeric Reads from HITS-CLIP (CLIP-Seq) Libraries

Panagiotis Alexiou, Manolis Maragkakis, Zissimos Mourelatos, and Anastassios Vourekas

Abstract

HITS-CLIP (High-Throughput Sequencing after in vivo Crosslinking and Immunoprecipitation, CLIP-Seq) libraries contain fragments of the RNA sequences bound in vivo by an RNA binding protein (RBP). Such fragments, especially if they represent RNA duplexes bound in vivo by the RBP, can occasionally be ligated together to form *chimeric* CLIP tags. Chimeric CLIP tags from Argonaute CLIP libraries can provide the exact base pairing profiles of small RNAs with their target RNA sequences, thus solving a critical problem in the field of post-transcriptional regulation. We recently reported an analysis of chimeric reads from the *Drosophila* Piwi protein Aubergine, which revealed a novel mechanism for mRNA entrapment within germ RNP granules. We term this novel approach chimeric CLIP (cCLIP) and present here the main steps that a researcher can take after the acquisition of the deep sequencing data, for the identification of candidate chimeric reads in Piwi CLIP libraries. Extending the scope beyond small-RNA binding proteins, we believe that cCLIP can be utilized to elucidate the in vivo functions of RNA-binding proteins in general, and especially those that modulate RNA secondary structures. We, therefore, also describe aspects of the generalized chimeric read identification problem, which can find use in the analysis of the CLIP libraries of any RNA-binding protein.

Key words HITS-CLIP, CLIP-Seq, RNA-IP, Argonaute, Piwi, Next generation sequencing, Illumina, cDNA, Immunoprecipitation, Chimeric reads, RNA-binding protein, Ribonucleoprotein complexes, In vivo, Transcriptomic analysis, Base-paired RNA, miRNA target sequences, piRNA target sequences, Posttranscriptional RNA processing, Gene silencing

1 Introduction

High-Throughput Sequencing after Crosslinking and Immunoprecipitation (HITS-CLIP, CLIP-Seq) has become the leading method to probe the target specificity of RNA-binding proteins (RBPs) in general and Argonaute family proteins in particular. UV crosslinking of freshly harvested live tissue or live cells induces

Panagiotis Alexiou and Anastassios Vourekas contributed equally to this work.

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_6,

the formation of covalent bonds between RNA and proteins that are in complex, and this allows the immunopurification of the crosslinked complexes under very stringent conditions [1–3]. CLIP libraries contain fragments of the RNA sequence that were directly bound by the studied protein in vivo, called CLIP tags. Recently, our group and others have discovered CLIP tags that contain fragments of RNA sequences ligated together [4–9]. These reads were called chimeric (also referred to as hybrid), and it was determined that they represent base-paired RNA sequences, bound in vivo by the RNA-binding protein that is being probed by the CLIP approach, and ligated together by virtue of their proximity while bound by and crosslinked to the RBP (Fig. 1). Chimeric reads between small RNAs and their complementary RNA targets have been discovered in Argonaute (Ago) and Piwi CLIP libraries [4–6, 8, 9], and have been analyzed to provide the small RNA (miRNA or piRNA) target sequence in vivo. This discovery reduces the need for prediction approaches for target sequence identification in an Argonaute CLIP experiment, approaches that are blind to the availability of the target sequence in vivo, and have known limitations concerning the specificity and sensitivity of target sequence detection. Similar to the chimeric reads are the hybrid reads identified by the hiCLIP method, which revealed secondary structures bound by Staufen 1 [7].

Chimeric reads can be formed by endogenous activities, in addition to being an objective of a specific step of the CLIP experimental protocol [6, 9]. Our group has adapted and optimized the HITS-CLIP protocol [2] for Piwi proteins, and has previously published a detailed description of this protocol [10]. Although we did not include a specific step to promote the generation of chimeric reads, numerous chimeras were identified in mouse (unpublished data) and *Drosophila* Piwi protein CLIP libraries [5], in agreement with previous observations [6, 9]. It was suggested that an endogenous ligase activity present in the lysates of crosslinked cells can ligate RNAs cleaved by endonuclease RNase T1 (provided exogenously) to produce 5′ -OH and 2′-3′ cyclic phosphate ends

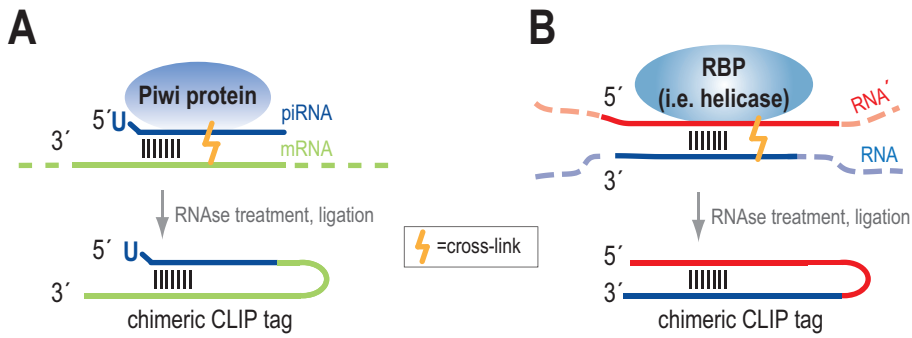


Fig. 1 Formation of chimeric reads in CLIP libraries. (**a**) Chimeric reads in Piwi CLIP libraries. (**b**) Hypothesis on how chimeric reads can be formed in CLIP libraries of any RNA binding protein that binds RNA duplexes

(these are slowly resolved into 3′-P) [6]. These ends can be substrates for the endogenous RtcB tRNA ligase activity, which joins the spliced tRNA halves after the removal of introns [11], therefore RtcB can play the role of the suggested activity that generates chimeric reads in the absence of exogenously provided ligase. We have observed the telltale sequence bias of RNase A cleavage on CLIP tags from mouse testis [12] and *Drosophila* early embryo (0–2 h) libraries. RNase A, like RNase T1, generates 5′ -OH ends and 2′-3′ cyclic phosphate ends, therefore both activities required for the generation of chimeric reads can be endogenously present in the cell lysate. We have verified that the chimeric reads generated by the endogenous activities accurately capture authentic, in vivo piRNA base pairing with mRNAs in the context of a Piwi-piRNA-mRNA tripartite complex [5]. It is also possible that a small degree of other nucleolytic and ligase activities can promote the generation of some chimeras. The subsequent steps of our protocol should prevent any other RNA ligase byproducts; therefore, our data are compatible with fortuitous chimeric read generation (possibly during the immunoprecipitation step) by endogenous nucleolytic and ligase activities [5]. We have noticed that such activities are apparently not always present in the cell and tissue lysates at sufficient levels to promote the formation of chimeras; therefore, the researcher should employ one of the relevant published protocols, if chimeras are an experimental objective.

We have observed the presence of chimeric reads in libraries of various RNA-binding proteins (unpublished data). These concern both *in cis* and *in trans* RNA duplex elements. We are certain that the analysis of these chimeras can reveal a trove of information on the type of structures that different RBPs recognize, and how the activities of certain RBPs can potentially modulate these structures.

This method paper aims to provide details on how to bioinformatically retrieve and analyze chimeric reads from CLIP libraries (Fig. 2), assuming that the researcher has employed one of the relevant methods that have been described adequately elsewhere. We use as example CLIP libraries of the Piwi protein Aubergine, and we try to set the framework for generalizing the chimeric read identification approach for any RNA-binding protein that binds structured RNAs.

Candidate chimeric reads are found in the pool of the unmapped CLIP tags (*see* **Note 1**). We filter reads in which an intact piRNA sequence can be found either at the 5′ or 3′ end of an unmapped read. We have observed the occurrence of chimeras containing piRNAs at both the 5′ and 3′ ends at comparable rates, and at least in Aubergine CLIP libraries the properties of such chimeras are similar. Therefore, we include both types of chimeras in downstream analyses. We then map the remaining fragment of the chimera against the genome, and for the purposes of the analysis of the entrapment of germ plasm localized mRNAs, we keep only chimeras that contain both a piRNA and a sequence fragment that can be uniquely aligned to an mRNA.

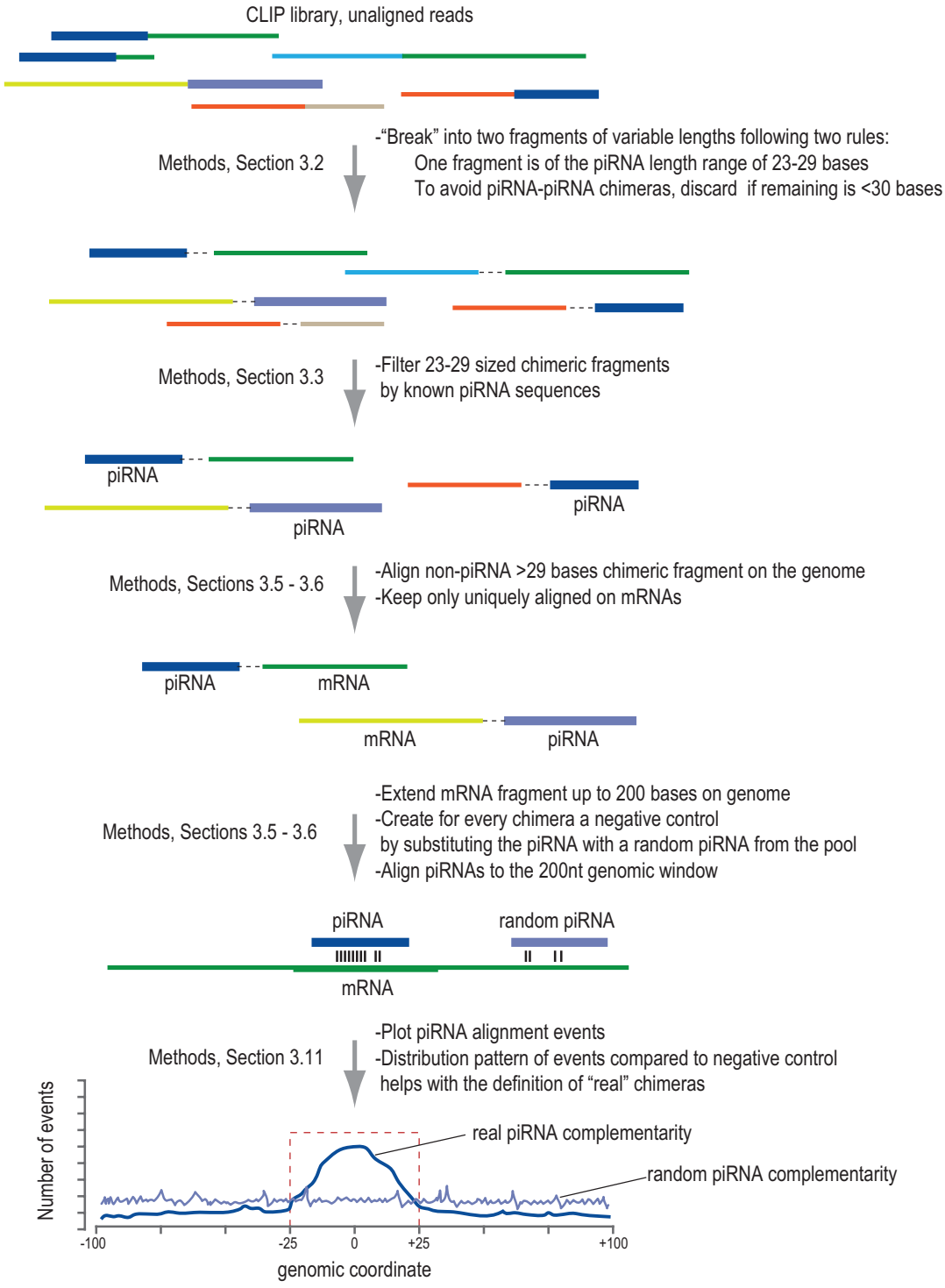


Fig. 2 Graphic outline of the retrieval of chimeric reads from a Piwi CLIP library. Throughout the text, we provide additional guidelines for applying this methodology to address the general problem of retrieving chimeric reads from a CLIP library of any RNA binding protein

Critical for the successful retrieval and subsequent analysis of chimeric reads from CLIP libraries is the approach followed for uncovering the base pairing patterns between the two RNA fragments found within the chimeric reads. We note that the use of RNA folding algorithms for determining candidate chimeras has several weaknesses, and in our opinion should be used with caution. First, the chimeric read is expected to fold back into a hairpin, but since this structure may have never existed in vivo, considering the thermodynamic score of this structure as a proxy for the legitimacy of the chimera can be misleading. Protein binding can modify the thermodynamic stability of an RNA duplex, and in fact many RNA-binding proteins such as RNA helicases in the presence of ATP can force the formation of RNA structures or duplexes that are not favored thermodynamically [13]. Further, due to the randomness of the nucleolytic trimming of the crosslinked RNAs, part of the base-paired sequences might be missing, or parts of the mRNA target sequence might be able to fold back on itself.

Having the above in mind, we followed a straightforward approach that reduces the assumptions about the base pairing profiles of the piRNA and mRNA fragments of chimeric reads. We employed a local alignment strategy that maps the piRNA sequence to a genomic window, extended −/+100 bases from the midpoint of the genomic alignment of the mRNA fragment. This allows the uncovering of the full base-paired sequence in case of missing nucleotides, and also provides a reliable estimate of the signal-to-noise ratio (here as noise we consider the random complementarity that a piRNA can show to genomic sequences neighboring the captured complementary site). Our pipeline delivers the chimeras and the complementarity profiles of the pairs of RNA fragments found in them (in this example piRNAs and mRNA fragments). We utilize the plot of the complementarity events relative to the mRNA fragment midpoint as a way to evaluate the authenticity of the complementarity events and therefore the chimeras, by comparing the distribution of these events with the negative control (random piRNA complementarity), and the events occurring by chance in sequences neighboring the mRNA fragment (Figs. 2 and 3). The researcher can use filters present in our pipeline to select any subset of the chimeras based on the complementarity score and on the position of the complementarity event relative to the mRNA fragment midpoint.

2 Required Programs and Data

2.1 Programs

1. Perl (https://www.perl.org/).
   (a) GenOO [14] (https://github.com/genoo/GenOO/).
   (b) Modern::Perl (http://search.cpan.org/~chromatic/Modern-Perl-1.20161005/lib/Modern/Perl.pm).

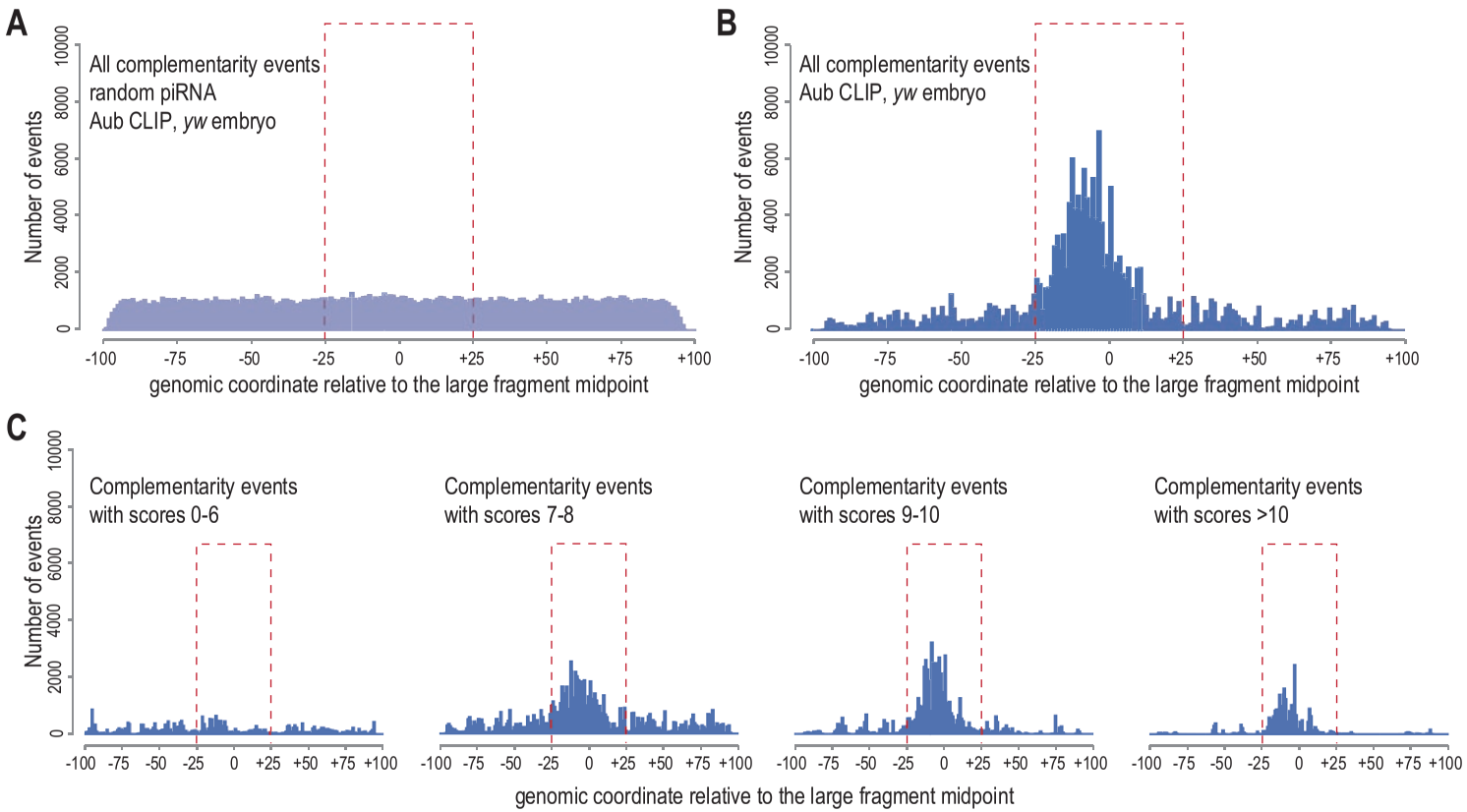


Fig. 3 piRNA–mRNA complementarity events (counts) within −/+100 bases from the midpoint of the mRNA fragment of chimeras found in CLIP libraries of *Drosophila* Piwi protein Aubergine. (**a**) Negative control (random piRNA) complementarity events for one Aubergine CLIP library. (**b**) All complementarity events for the same library as in (**a**). (**c**) Complementarity events for the same library as in (**a**) and (**b**), plotted separately for four alignment score groups. The *dashed red* line demarcates the subset of the events that occur within −/+25 bases from the midpoint of the mRNA fragment. Visual observation reveals that these events are substantially more abundant compared to negative control and to events occurring further away from the midpoint of the mRNA fragment, for scores >6. Our pipeline provides the option to extract any subset of chimeric reads by the complementarity score or the relative position from the mRNA fragment

   (c) Getopt::Long (http://search.cpan.org/~jv/Getopt-Long-2.49.1/lib/Getopt/Long.pm).
   (d) File::Path (http://search.cpan.org/~riche/File-Path-2.12/lib/File/Path.pm).
2. R (for plots only) (https://cran.r-project.org/).
   (a) Optparse (https://cran.r-project.org/web/packages/optparse/).
3. STAR aligner (https://github.com/alexdobin/STAR).
4. Samtools (http://www.htslib.org/).
5. Bedtools (http://bedtools.readthedocs.io/).

2.2 Data

1. Chromosome sequences for reference genome (FASTA format).
2. Gene locations (.gtf format file).

3 Methods

This is a step-by-step description of the bioinformatic pipeline used to retrieve the chimeric reads from a Piwi CLIP library. The segments of the functional bash script that can run this pipeline can be found in every section. A functional, complete bash script that can be used to run this analysis, along with all the required analysis scripts, are also deposited online (https://github.com/palexiou/CLIP-chimaeric/).

3.1 Preprocessing and Alignment

The input CLIP reads have been prepared by trimming adaptor sequences followed by alignment to the genome using the STAR aligner (*see* **Note 2**).

3.2 Split Unmapped Reads

Script: split_unmapped_reads.pl

Unmapped reads are split into "small" and "large" fragments and converted into .fasta format files. The flags *--min-small 23* and *--max-small 29* limit the length of small sequence output. The flag *--min-large 30* only allows large fragments 30 nucleotides or longer to be produced. Chimeras are split at every point that allows for pairs of fragments within the given length constraints to be produced (*see* **Notes 3–5**). For example, a 55 nt long read would be broken in pairs of (23,32), (24,31), (25,30), (30,25), (31,24) and (32,22) nucleotides, within the constraints used here.

```
perl dev/split_unmapped_reads.pl \
    --ifile $UnmappedReadsFile.fastq \
    --min-small 23 \
    --max-small 29 \
```

```
        --min-large 30 \
        --out-small results/$samplename/frag-
    ments.small.fa \
        --out-large results/$samplename/frag-
ments.large.fa
```

3.3 Filter Small Fragments with Library of Known Small Reads (piRNAs)

Script: filter_fa_by_sam_sequence.pl

We want to retain only possible chimeric pairs that contain a real piRNA sequence. To this end, the small fragments are filtered based on the sequences found in the aligned reads of a piRNA CLIP library. Other sequence filtering steps can be performed here.

```
    cat $Aligned.SmallReadsFile.bam \
        | samtools-1.2 view \
            -h \
            - \
        | perl dev/filter_fa_by_sam_sequence.pl \
            --ifile results/$samplename/frag-
ments.small.fa \
            --ffile - \
        > results/$samplename/fragments.small.
filtered.fa
```

3.4 Filter Large Fragments with Passed Small Fragments

Script: filter_by_name.pl

After the "small fragment file" has been filtered, we remove from the "large fragment file" all large fragments that were paired with small fragments that did not pass the filter (*see* **Note 6**).

```
    perl dev/filter_by_name.pl \
        --ifile results/$samplename/fragments.
large.fa \
        --ref-file results/$samplename/frag-
ments.small.filtered.fa \
        > results/$samplename/fragments.large.
filtered.fa
```

3.5 Align Fragments to the Genome

Tool: STAR 2.5.2b aligner

We align the filtered large and small fragment files separately to the reference genome using the STAR aligner. The attached bash script contains the STAR settings used in this step. Neither alignment allows introns based on known genes. Again, small fragments are allowed to map in up to 50 genomic locations with equally good score.

```
### ALIGN LARGE FRAGMENTS
    STAR-2.5.2b \
```

```
        --genomeDir data/genome/dm3/ \
        --readFilesIn results/$samplename/
fragments.large.filtered.fa \
        --readFilesCommand cat \
        --runThreadN 12 \
        --outSAMtype SAM \
        --outSAMattributes All \
        --outFilterMultimapScoreRange 0 \
        --alignIntronMax 1 \
        --outFilterMatchNmin 8 \
        --outFilterMatchNminOverLread 0.9 \
        --outFileNamePrefix
results/$samplename/fragments.large.filtered.

    ### ALIGN SMALL FRAGMENTS
    STAR-2.5.2b \
        --genomeDir data/genome/dm3/ \
        --readFilesIn results/$samplename/
fragments.small.filtered.fa \
        --readFilesCommand cat \
        --runThreadN 12 \
        --outSAMtype SAM \
        --outSAMattributes All \
        --outFilterMultimapScoreRange 0 \
        --alignIntronMax 1 \
        --outFilterMatchNmin 8 \
        --outFilterMatchNminOverLread 0.9 \
        --outFileNamePrefix
results/$samplename/fragments.small.filtered. \
        --outFilterMultimapNmax 50
```

3.6 Post-Alignment Filters

Tool: bedtools intersect

We use *bedtools intersect* and a General Transfer/Feature Format (.gtf/.gff) file of gene location to filter large fragments that map on genes. The output of this process should be two .sam files—one containing the filtered and aligned small fragments, and one containing the filtered and aligned large fragments (*see* **Notes 7** and **8**).

```
    cat results/$samplename/fragments.large.
filtered.Aligned.out.sam \
        | grep 'NH:i:1' \
        | samtools-1.2 view \
            -b \
            -T $GenomeFile.fa \
            - \
        > results/$samplename/fragments.large.
filtered.Aligned.bam
```

```
    bedtools intersect \
        -a results/$samplename/fragments.
large.filtered.Aligned.bam \
        -b $GenesFile.gtf \
        -split \
        -u \
    | samtools-1.2 view \
        - \
        > results/$samplename/fragments.large.
filtered.Aligned.OnGenes.sam
```

3.7 Keep Unique Pair

Script: filter_best_pairs.pl

At this step, we still possibly have multiple pairs of fragments per original chimeric read. Considering the genomic alignment of both the short and long fragments, we choose the pair of fragments with the highest sum of matching nucleotides. If there are multiple pairs with the same total alignment score, a random pair is selected.

```
    perl dev/filter_best_pairs.pl \
        --ifile1 results/$samplename/fragments.
large.filtered.Aligned.OnGenes.sam \
        --ifile2 results/$samplename/fragments.
small.filtered.Aligned.out.sam \
        --ofile1 results/$samplename/fragments.
large.filtered.Aligned.OnGenes.paired.sam \
        --ofile2 results/$samplename/fragments.
small.filtered.Aligned.paired.sam
```

3.8 Get Fragment Sequences

Script: sam-to-fasta.pl

We extend the large reads to a 200-nucleotide (nt) window using the *--out-length 200* flag and convert to .fasta file format. We also convert small reads to .fasta file format without extending their length (*see* **Note 9**).

```
    perl dev/sam-to-fasta.pl \
        --sam results/$samplename/fragments.
large.filtered.Aligned.OnGenes.paired.sam \
        --chr_dir data/$species/chromosomes/ \
        --out-length 200 \
        > results/$samplename/fragments.large.
filtered.Aligned.OnGenes.paired.200nt.fa
```

```
    perl dev/sam-to-fasta.pl \
        --sam results/$samplename/fragments.
small.filtered.Aligned.paired.sam \
        --chr_dir data/$species/chromosomes/ \
        --exact \
        > results/$samplename/fragments.small.
filtered.Aligned.paired.fa
```

3.9 Make Pairs Table

Script: make_pairs_table.pl

This step combines the two .fasta sequence files produced in the previous step to create a tab delimited pairs table. We run the same command a second time with the *--shuffle* option to make the same number of random pairs (*see* **Note 10**).

```
    perl dev/make_pairs_table.pl \
        --ifile1 results/$samplename/fragments.
small.filtered.Aligned.paired.fa \
        --ifile2 results/$samplename/fragments.
large.filtered.Aligned.OnGenes.paired.200nt.fa \
        > results/$samplename/pairs.tab
```

```
    perl dev/make_pairs_table.pl \
        --ifile1 results/$samplename/fragments.
small.filtered.Aligned.paired.fa \
        --ifile2 results/$samplename/fragments.
large.filtered.Aligned.OnGenes.paired.200nt.fa \
        --shuffle \
        > results/$samplename/pairs.shuffled.tab
```

3.10 Align

Script: align.pl

We align the small fragments to the extended 200-nt fragments (for both real and random pairs). We use a slightly modified Smith–Waterman [15] alignment method (weights: match = +1, mismatch = −1, gap = −2). Differences of our alignment versus Smith–Waterman: (a) No penalties are given to non-matching nucleotides on the edges of the alignment. (b) If there are multiple optimal alignment scores, one is picked randomly. (c) Alignments in which part of one sequence is outside the boundaries of the other sequence are not considered. The fields of the output alignment file are as follows: "chimeric read name", "sequence small fragment", "sequence large fragment (200 nt)", "length small fragment", "length large fragment", "alignment score", "midpoint location on large fragment", "binding vector large fragment", "binding vector small fragment reverse complement", and "reverse complement sequence small fragment" (*see* **Notes 11** and **12**).

```
    perl dev/align.pl \
        --ifile results/$samplename/pairs.tab \
        > results/$samplename/pairs.aligned.tab
```

```
    perl dev/align.pl \
        --ifile results/$samplename/pairs.shuf-
fled.tab \
        > results/$samplename/pairs.shuffled.
aligned.tab
```

3.11 Plots

Script: plots.R

Using the alignment files as input (real and random alignments) this R script produces several plots that help understand the alignment and therefore provide valuable insight into the attributes of the complementarity between the piRNAs and their target sequences. Plot-1 is a scatter plot of real versus random alignment scores. Plot-2 is a histogram of real and random alignment scores. Plot-3 is a histogram of the distribution of real and random alignment midpoint locations on the 200-nt extended large fragment.

```
Rscript dev/plots.R \
    --ifile results/$samplename/pairs.
aligned.tab \
    --sfile results/$samplename/pairs.shuf-
fled.aligned.tab \
    --ofile results/$samplename/pairs.
aligned.pdf
```

Script: filter_alignment.pl

Allows user defined filtering of the alignment file. Flags --min-score, --max-score, and --min-midpoint, --max-midpoint, give the constraints for output alignments. These settings allow the user to filter a subset of the chimeras based on their alignment characteristics for further analysis downstream.

4 Notes

1. Gapped read alignment can interfere with the retrieval of the candidate chimeras, especially for CLIP libraries in which local secondary structures are expected to be captured by chimeric reads. In this case, the researcher should opt for non-gapped alignment of the CLIP libraries, and also develop a rationale for separating chimeric reads from standard CLIP tags that may have small deletions in their sequence.
2. Using the flag –outReadsUnmapped Fastx will output unmapped reads into a separate file. Using the flag –outFilterMultimapNmax 50 allows up to 50 equal score mappings per read instead of the default ten. This is recommended when working with repetitive sequences such as piRNAs. Any aligner that produces a standard SAM/BAM file can be used at this step.
3. In order to avoid piRNA:piRNA chimeras, the remaining sequence after generating the piRNA-sized small fragment is required to be >29-nt long.
4. Pairs with the small fragment both on the 5′ end and the 3′ end of the chimeric read can be output.
5. Extending the method to any size is a matter of tweaking the –min-small, --max-small, --min-large, --max-large flags to the

desired lengths. If these flags are left undefined then pairs of fragments of all potential sizes will be produced. The researcher should be aware that many sequence fragments with a size of ≤20 bases cannot be unambiguously aligned to the genome. This ambiguity increases, as the size gets smaller. We chose to process only chimeric reads that contain sequence fragments that are unambiguously identified within a piRNA library or within mRNA exons. This requirement reduces uncertainty in downstream analyses such as complementarity between the two chimeric read fragments.

6. This is an optional step, but as the next step will be a resource consuming alignment to the genome, it is more efficient to filter unwanted reads before aligning rather than after.
7. Using the –split option we take into consideration only the exonic regions of genes.
8. At this point, any additional filters can be applied to the aligned files.
9. The –chr_dir flag denotes a folder containing a file per chromosome named $ChrName.fa.
10. The random pairs consist of a random small sequence (piRNA) of the same length as the original small fragment of chimeric read, paired with the same long fragment sequence. In this way, we avoid biases of sequence length distribution in the alignment scores.
11. The binding vector shows "1" for binding nucleotides within the binding area, "0" for non-binding nucleotides within the binding area, "g" for nucleotides that have a gap across from them in the alignment (insertions), "s" for soft clipped non-matching nucleotides (outside the binding area and not matching in the extension), "m" for soft clipped matching nucleotides (outside the binding area and matching in the extension). The binding area is the area in which binding nucleotides count toward the final alignment score—the area of local alignment.
12. Areas of the small fragment outside the core alignment area that spans from the beginning of the trackback cell to the first encountered cell of score zero are expanded in an ungapped alignment and do not count toward the alignment score or midpoint. In other words, soft-clipped nucleotides are not penalized and do not affect the alignment score but are reported.

Acknowledgments

We thank members of the Mourelatos laboratory for discussions. This research was supported by NIH grant GM072777 and a grant from ALS Therapy Alliance to Z.M.

References

1. Licatalosi DD et al (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464–469
2. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460:479–486
3. Ule J et al (2003) CLIP identifies nova-regulated RNA networks in the brain. Science 302(5648):1212–1215
4. Kameswaran V et al (2013) Epigenetic regulation of the DLK1-MEG3 MicroRNA cluster in human type 2 diabetic islets. Cell Metab 19:135–145
5. Vourekas A, Alexiou P, Vrettos N, Maragkakis M, Mourelatos Z (2016) Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. Nature 531:390–394
6. Grosswendt S et al (2014) Unambiguous identification of miRNA: target site interactions by different types of ligation reactions. Mol Cell 54:1042–1054
7. Sugimoto Y et al (2015) {hiCLIP} reveals the in vivo atlas of {mRNA} secondary structures recognized by Staufen 1. Nature 519:491–494
8. Helwak A, Kudla G, Dudnakova T, Tollervey D (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 153:654–665
9. Moore MJ et al (2015) miRNA-target chimeras reveal miRNA 3′-end pairing as a major determinant of Argonaute target specificity. Nat Commun 6:8864
10. Vourekas A, Mourelatos Z (2014) HITS-CLIP (CLIP-Seq) for mouse piwi proteins. Methods Mol Biol 1093:73–95
11. Popow J et al (2011) HSPC117 is the essential subunit of a human tRNA splicing ligase complex. Science 331:760–764
12. Vourekas A et al (2015) The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. Genes Dev 29:617–629
13. Yang Q, Fairman ME, Jankowsky E (2007) DEAD-box-protein-assisted RNA structure conversion towards and against thermodynamic equilibrium values. J Mol Biol 368:1087–1100
14. Maragkakis M, Alexiou P, Mourelatos Z (2015) GenOO: a modern perl framework for high throughput sequencing analysis. bioRxiv. doi:10.1101/019265
15. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197

Chapter 7

Kinetic Analysis of Small Silencing RNA Production by Human and *Drosophila* Dicer Enzymes In Vitro

Susan E. Liao and Ryuya Fukunaga

Abstract

Dicer enzymes produce small silencing RNAs such as microRNAs (miRNAs) and small interfering RNAs (siRNAs), which then are loaded into Argonaute proteins and act as sequence-specific guides. A powerful tool to understand the molecular mechanism of small silencing RNA production by Dicers is an in vitro RNA processing assay using recombinant Dicer proteins. Such biochemical analyses have elucidated the substrate specificities and kinetics of Dicers, the mechanism by which the length of small RNAs produced by Dicers is determined, and the effects of Dicer-partner proteins and endogenous small molecules such as ATP and inorganic phosphate on small RNA production by Dicers, among others. Here, we describe methods for in vitro small RNA production assay using recombinant human and *Drosophila* Dicer proteins.

Key words Dicer, miRNA, siRNA, RNA silencing, Kinetics

1 Introduction

The ribonuclease (RNase) enzyme Dicer is central to the production of small silencing RNAs such as microRNAs (miRNAs) and small interfering RNAs (siRNAs) [1–4]. In canonical miRNA biogenesis, Dicer processes precursor miRNAs (pre-miRNAs) into miRNA duplexes, while some miRNAs are produced in a Dicer-independent manner. Dicer also cleaves long dsRNA substrates to produce siRNAs. Endogenous sources of long dsRNA include viral RNA genomes or replication intermediates, convergent mRNA transcription, transposon transcripts, and partially self-complementary hairpin RNAs. Exogenous long dsRNA can be introduced experimentally to trigger RNA interference (RNAi). After their production, miRNA duplexes and siRNA duplexes are loaded into Argonaute proteins. Only one of the strand of the duplexes (the guide strand) is retained in Argonaute proteins while the other strand (the passenger strand) is released to solution. The guide-strand and Argonaute complex is called

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_7,

an RNA-induced silencing complex (RISC). miRNA-Argonaute RISC complexes can repress mRNA translation and promote mRNA turnover, whereas siRNA-Argonaute RISC complexes cleave their RNA targets.

An in vitro RNA processing assay using recombinant Dicer proteins can be used to investigate the functions of Dicer in small RNA biogenesis. Such in vitro RNA processing assays directly proved that Dicer processes precursor substrate RNAs into miRNAs and siRNAs, characterized biochemical and structural elements required for Dicer activity, determined kinetic parameters in small RNA production by Dicer, and elucidated how Dicer partner proteins and small molecules modulate small RNA biogenesis [5–19].

Here, we describe RNA processing assays which use recombinant human or *Drosophila* Dicer to process various RNA substrates in test tubes. It is important to note that the number of Dicer genes and their respective functions in small RNA biogenesis varies between these species. A single human Dicer enzyme produces both miRNAs and siRNAs. In contrast, *Drosophila* possesses two Dicer genes: Dicer-1, which produces ~22–24 nt miRNAs from pre-miRNAs, and Dicer-2, which precisely produces 21 nt siRNAs from long dsRNAs [20]. These Dicer proteins share common domain architecture: N-terminal helicase domain, dsRNA-binding domain, platform domain, PAZ (Piwi-Argonaute-Zwille) domain, connector helix, two RNase III domains, and C-terminal dsRNA-binding domain (Fig. 1). The crystal structure of Dicer from the unicellular *Giardia lamblia*, which lacks the helicase and dsRNA-binding domains and produces ~25–27 nt siRNAs, revealed that it adopts a hatchet-shaped structure; the RNase III domains form the blade, and the PAZ domain makes up the base of the handle [6]. Electron microscope studies revealed that human Dicer and *Drosophila* Dicer-2 adopt an L-shape structure where the platform, PAZ, and RNase III domains constitute the longer axis while the helicase domain composes the shorter axis of the L-shape [21–24]. Unlike human Dicer and *Drosophila* Dicer-2, *Drosophila* Dicer-1 lacks the DExD/H motif in the helicase domain. The ATP hydrolysis activity of the DExD/H motif of the *Drosophila* Dicer-2 helicase domain has been shown to facilitate Dicer-2 translocation along dsRNA substrates [9, 11].

In this chapter, we describe an in vitro assay for characterizing Dicer RNA processing activity, including purification of recombinant human and *Drosophila* Dicer proteins, preparation of various RNA substrates, performance of RNA processing reactions in test tubes, and analysis of kinetic data. The methods described here can aid studies developing quantitative understanding of Dicer function, mechanism, and regulation in small RNA processing.

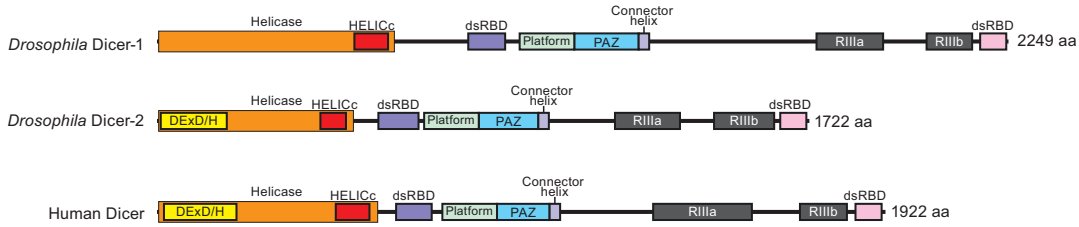


Fig. 1 *Drosophila* and human Dicer domain architecture. DExD/H, DExD/DExH box helicase domain; HELICc, helicase conserved C-terminal domain; dsRBD, dsRNA-binding domain; PAZ, PAZ domain; RIIIa and RIIIb, Ribonuclease III domain

2 Materials

2.1 Purification of Recombinant Dicer Proteins

1. Bac-to-Bac® N-His TOPO® Cloning Kit (Thermo Fisher Scientific).
2. MAX Efficiency® DH10Bac™ Competent Cells (Thermo Fisher Scientific).
3. FuGENE HD Transfection Reagent (Promega).
4. Sf9 cells in Sf-900™ II SFM (Thermo Fisher Scientific).
5. Sf-900™ II SFM (Thermo Fisher Scientific).
6. Type B Dounce glass homogenizer.
7. Ultracentrifuge.
8. Glass Econo-Column chromatography column (Biorad).
9. Ni-Sepharose 6 Fast Flow (GE Healthcare).
10. HiTrap Q HP (GE Healthcare).
11. AKTA protein purification system such as AKTA Pure (GE Healthcare).
12. Ni-lysis buffer: 50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 40 mM imidazole, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM tris(2-carboxyethyl) phosphine (TCEP), 1 tablet/50 mL buffer of cOmplete™ Mini EDTA-free Protease Inhibitor cocktail (Roche).
13. Ni-wash buffer: 50 mM HEPES-NaOH (pH 7.5), 500 mM NaCl, 40 mM imidazole, 0.5 mM PMSF, 1 mM TCEP.
14. Ni-elution buffer: 50 mM HEPES-NaOH (pH 7.5), 500 mM NaCl, 500 mM imidazole, 1 mM TCEP.
15. Dialysis buffer: 20 mM HEPES-NaOH (pH 8.0), 20 mM NaCl, 1 mM EDTA, 1 mM TCEP.
16. HiTrapQ buffer A: 20 mM HEPES-NaOH (pH 8.0), 20 mM NaCl, 1 mM TCEP.
17. HiTrapQ buffer B: 20 mM HEPES-NaOH (pH 8.0), 1 M NaCl, 1 mM TCEP.

18. Slide-A-Lyzer™ Dialysis Cassette, MWCO 10,000 (Thermo Fisher Scientific).
19. Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane (Millipore).
20. Concentration buffer: 30 mM HEPES-KOH (pH 7.4), 100 mM KOAc, 1 mM TCEP.
21. 80% glycerol buffer: 80% (v/v) glycerol, 30 mM HEPES-KOH (pH 7.4), 100 mM KOAc, 1 mM TCEP.
22. Liquid nitrogen.
23. SDS-PAGE gels, SDS-PAGE sample buffer, SDS-PAGE running buffer, and running apparatus.
24. UV spectrophotometer.

2.2 Preparation of Cold Pre-miRNA

1. Pre-miRNA oligos (Custom synthesis).
   Drosophila pre-*let-7:* 5′-UGA GGU AGU AGG UUG UAU AGU AGU AAU UAC ACA UCA UAC UAU ACA AUG UGC UAG CUU UCU-3′.
   Human pre-miR-590: 5′- GAG CUU AUU CAU AAA AGU GCA GUA UGG UGA AGU CAA UCU GUA AUU UUA UGU AUA AGC UAG U-3′.
2. 10 units/μL T4 Polynucleotide Kinase (NEB).
3. 10× T4 PNK buffer (NEB).
4. 100 mM ATP.
5. 2× formamide RNA loading buffer without xylene cyanol: 98% (v/v) formamide, 10 mM EDTA (pH 8.0), and 0.025% (w/v) bromophenol blue.
6. 0.3 M NaCl.
7. Acrylamide/Bis 19:1, 40% (w/v) solution.
8. 5× TBE buffer: 1.1 M Tris, 900 mM Borate, 25 mM EDTA (pH 8.3).
9. 10% (w/v) ammonium persulfate (APS).
10. Urea-PAGE gel solution [1× TBE, 7 M Urea, Acrylamide/Bis 19:1 8% (w/v)].
11. 150 mL, 0.22 μm Tube Top Vacuum Filter.
12. Vertical gel electrophoresis apparatus (C.B.S. Scientific).
13. Glass Plate Set Notched, 33 × 22 cm. (C.B.S. Scientific).
14. Sequencing Spacer Set 1.5 mm Thick × 22 cm(l) (C.B.S. Scientific).
15. 1.5 mm × 8 Well Comb for 33 cm wide gel (C.B.S. Scientific).
16. Aluminum Heat Dispersion Plate, 33(w) × 6.8 cm(h) (C.B.S. Scientific).
17. Rain-X original glass treatment, and Rain-X anti-fog solutions.
18. Handheld UV lamp.

19. Fluor-coated TLC plate.
20. Razor blade.

2.3 Preparation of 5′-Radiolabeled Pre-miRNA

1. γ-^{32}P ATP (6000 Ci/mmol, 150 mCi/mL. PerkinElmer).
2. Illustra MicroSpin G-25 Columns (GE Healthcare).
3. 2× formamide RNA loading buffer: 98% (v/v) formamide, 10 mM EDTA (pH 8.0), 0.025% (w/v) xylene cyanol, and 0.025% (w/v) bromophenol blue.
4. 10 mg/mL glycogen.
5. Spin-X Plastic Centrifuge Tube Filters (Corning).
6. Storage Phosphor Screen (GE Healthcare).
7. Typhoon FLA-9500 (GE Healthcare).
8. Scintillation counter.
9. Decade Marker (Ambion).

2.4 Preparation of Cold Long dsRNAs

1. pEGFP-N3 plasmid (Clontech).
2. Forward DNA primer for the 106 nt sense strand template: 5′-GTACTTAATACGACTCACTATAGCCAGGGCACGGGCAGCTTGCCG-3′. (The sequence of T7 promoter is underlined).
3. Reverse DNA primer for the 106 nt sense strand template: 5′-mTmTGGGCCACAAGTTCAGCGTGTCC-3′ (m, 2′-O-methyl ribose). (Incorporation of two bases with a 2′-O-methyl-modification at the 3′-end can cause the transcription to stop without addition of non-templated nucleotides [25]).
4. Forward DNA primer for the 106 nt antisense strand template: 5′-GTACTTAATACGACTCACTATAGGGCCAGGGCACGGGCAGCTTGCCG-3′.
5. Reverse DNA primer for the 106 nt antisense strand template: 5′-mGmGGCCACAAGTTCAGCGTGTCC-3′ (m, 2′-O-methyl ribose).
6. 2000 units/mL Phusion® High-Fidelity DNA Polymerase (New England Biolabs).
7. 5× Phusion® HF Buffer (New England Biolabs).
8. dNTPs (10 mM each).
9. Thermal cycler.
10. Phenol/Chloroform/Isoamyl Alcohol (25:24:1 Mixture, pH 6.7).
11. 3 M sodium acetate (pH 5.2).
12. ATP, CTP, GTP, UTP, and GMP stocks (100 mM each, pH 7.0).

13. 100 mM dithiothreitol (DTT).
14. 10× T7 transcription buffer: 400 mM Tris–HCl (pH 8.0), 260 mM $MgCl_2$, 25 mM Spermidine, 0.1% (v/v) Triton-X100.
15. Inorganic pyrophosphatase (IPP) in storage buffer: Resuspend the IPP (Sigma) in storage buffer [50 mM Hepes-KOH (pH 7.4), 1 mM DTT, 50% (v/v) glycerol] at a concentration of 0.5 units/μL.
16. 40 units/μL RNasin RNase inhibitor (Promega).
17. 1 unit/μL RQ1 RNase-free DNase (Promega).
18. T7 RNA polymerase, purchased commercially or prepared in the laboratory.
19. 10 unit/μL Alkaline Phosphatase, Calf Intestinal (CIP) (New England Biolabs).
20. 10× CIP buffer (500 mM Potassium Acetate, 200 mM Tris-acetate, 100 mM Magnesium Acetate, 1 mg/mL BSA, pH 7.9). (Same as 10× CutSmart buffer (New England Biolabs)).

2.5 Preparation of Internally Radiolabeled Long dsRNAs

1. α-^{32}P UTP (800 Ci/mmol, 10 mCi/mL).

2.6 Performing In Vitro RNA Processing Reactions

1. 1 M HEPES-KOH buffer (pH 7.4).
2. 1 M Potassium acetate.
3. 1 M Magnesium acetate.
4. Glycerol.
5. 10 mg/mL bovine serum albumin (BSA).
6. 0.4 mm × 34 Well Comb for 33 cm wide gel (C.B.S. Scientific).
7. Sequencing Spacer Set 0.4 mm Thickness × 22 cm(l) (C.B.S. Scientific).
8. Whatman™ 3MM Chr Chromatography Paper (Fisher Scientific).
9. Thermal Laminating Film (GBC).
10. HydroTech™Vacuum Gel Drying System (Biorad).

2.7 Analyzing Enzyme Kinetics

1. Igor Pro software (WaveMetrics).

3 Methods

3.1 Purification of Recombinant Proteins

1. PCR-amplify the coding sequences of *Drosophila* Dicer-1 and Dicer-2 and human Dicer (NCBI IDs: NM_079729, NM_079054, and NM_177438, respectively) and clone them into either the pFastBacHT or the pFastBac/NT-TOPO

plasmid for N-terminally 6× His-tagged recombinant Dicer protein expression in Sf9 cells. The cDNA clone plasmids for *Drosophila* Dicer-2 and human Dicer are available from Drosophila Genomics Resource Center (FI15132) and Addgene (#19873), respectively.

2. Express the desired Dicer protein in Sf9 cells at 25–27 °C using the prepared plasmids and Bac-to-Bac® Baculovirus Expression System following the manufacturer's manual.
3. Harvest the Sf9 cells expressing Dicer protein 48–72 h after baculovirus transfection by centrifugation at 1500 × *g* for 15 min at 4 °C. Flash-freeze pelleted cells in liquid nitrogen and store pellets in −80 °C until use.
4. Perform all purification procedures here and below at 4 °C or on ice. Resuspend the Sf9 cell pellet in 40 mL Ni-lysis buffer. Homogenize cells using a type B Dounce glass homogenizer. Apply ~40 strokes.
5. Ultracentrifuge the homogenate at 38,000 × *g* for 60 min at 4 °C.
6. Carefully transfer the supernatant to a new 50 mL conical tube.
7. Mix supernatant with ~5 mL of Ni-sepharose beads pre-equilibrated with Ni-wash buffer. Incubate mixture on a gentle tube rotator for 30 min at 4 °C.
8. Centrifuge the beads at 3000 × *g* for 5 min at 4 °C in a swing bucket rotor. Discard supernatant.
9. Wash beads by mixing with ~40 mL Ni-wash buffer. Centrifuge beads at 3000 × *g* for 5 min at 4 °C and discard supernatant. Repeat wash two more times.
10. Resuspend beads in ~10 mL of Ni-wash buffer and pour beads in an Econo-column. Wash the beads in the column with ~200 mL of Ni-wash buffer.
11. Slowly elute the proteins by pouring Ni-elution buffer into the Econo-column, adjusting the flow rate as needed with the stopcock. Elute the proteins into three 25 mL fractions.
12. Examine the amount and purity of the proteins in the eluted fractions by SDS-PAGE. Usually, the first fraction contains the highest yield of the target protein.
13. After identifying fraction(s) containing the highest yield of eluted proteins, dialyze fractions against Dialysis buffer at 4 °C overnight using a Slide-A-Lyzer™ Dialysis Cassette (MWCO 10,000) (*see* **Note 1**).
14. Purify the proteins using a HiTrapQ column on AKTA HPLC system. Pre-equilibrate the column with HiTrapQ buffer A, load the protein sample onto the column, and wash with HiTrapQ buffer A. Elute the proteins with a linear gradient of HiTrapQ buffer A and HiTrapQ buffer B.

15. Analyze the elution fractions using SDS-PAGE.
16. Pool the desired elution fractions. Concentrate fractions using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane. Perform buffer exchange with Concentration buffer and then concentrate to desired concentration.
17. Measure the UV spectrum of the concentrated proteins using the concentration buffer as a blank solution with a UV spectrophotometer. Determine protein concentrations based on the absorbance at 280 nm and the calculated extinction coefficients. The extinction coefficients of untagged intact *Drosophila* Dicer-1 and Dicer-2 and human Dicer are 226,220, 180,840, and 211,670 M^{-1} cm^{-1}, respectively. For example, 1 O.D. at 280 nm for *Drosophila* Dicer-1 = 4.42 μM = 1.13 mg/mL of *Drosophila* Dicer-1 protein.
18. (Optional). Run 1 μL from concentrated protein sample on an SDS-PAGE gel to check purity. To determine relative protein amounts, cut out visible protein bands, perform amino acid analysis, and calculate concentrations of protein samples.
19. Add 80% glycerol buffer to the protein solutions to final 50% (v/v) glycerol concentration. Aliquot protein solutions into small volumes (ex. 10–20 μL) in tubes to avoid freeze–thaw. Flash-freeze them in liquid nitrogen and store at −80 °C.

3.2 Preparation of Cold Pre-miRNAs

Phosphorylate 5′ OH pre-miRNA oligos with cold ATP and T4 Polynucleotide Kinase to make 5′ monophosphorylated pre-miRNAs. Incubate the following reaction at 37 °C for 1 h (*see* **Note 2**):

- 357 μL water.
- 50 μL 10× T4 PNK buffer.
- 33 μL 300 μM 5′ OH pre-miRNAs.
- 10 μL 100 mM ATP.
- 50 μL 10 U/μL T4 Polynucleotide Kinase (PNK).

1. During the PNK reaction, prepare a 1.5 mm thick 8% acrylamide PAGE gel (1× TBE, 7 M Urea). Coat notched glass plates with hydrophobic glass treatment (e.g., Rain-X original glass treatment) and back plates with hydrophilic glass treatment (e.g., Rain-X anti-fog) to facilitate gel removal after running. Assemble the glass plates and spacers and pour gel.
2. Pre-run the gel in 0.5× TBE as a running buffer in the vertical gel electrophoresis apparatus. Use the aluminum plate to disperse the heat generated during the gel running. The gel should be pre-run for about 30 min at a constant power of ~15 W before loading RNA samples.

3. After the PNK reaction, mix the 5′ monophosphorylated RNA solutions with an equal amount of 2× formamide RNA loading buffer without xylene cyanol (*see* **Note 3**). Incubate at 95 °C for 2 min.
4. Load RNA solutions onto the pre-run gel.
5. Run the gel at ~15 W until the faster bromophenol blue dye almost runs off the gel.
6. After electrophoresis, remove the notched glass plate and leave the gel on the back plate. Cover the gel with plastic wrap. Place the gel, plastic wrap-side-down, on the removed glass plate. Remove the other glass plate and cover the gel with a second sheet of plastic wrap.
7. Place a fluor-coated TLC plate between the plastic-wrapped gel and the bottom glass plate.
8. Visualize RNA bands by using a hand-held UV lamp (254 nm; shortwave) on the surface of the gel and mark the band to be cut out. RNA will appear as dark shadow bands while the TLC plate will appear green. Minimize the time of UV exposure to RNA in order to avoid damaging RNA with UV.
9. Cut out pre-miRNA bands using a razor blade, and place them in a 50 mL tube. Elute RNAs from the gel pieces in 0.3 M NaCl on a tube rotator at 4 °C overnight.
10. Filter eluted RNA solutions through a 150 mL, 0.22 μm Tube Top Vacuum Filter to remove any gel pieces. Then perform ethanol precipitation by adding three volumes of ethanol.
11. Precipitate the RNA pellet by centrifugation at 21,000 × *g* for 30 min at 4 °C, wash the pellets with 80% (v/v) ethanol, and air dry. Resuspend RNAs in water.
12. Measure the UV spectrum using a UV spectrophotometer and determine the concentration based on the absorbance at 260 nm. 1 O.D. at 260 nm for RNA molecules = 40 ng/μL of RNA. Store in −80 °C freezer.

3.3 Preparation of 5′-Radiolabeled Pre-miRNAs

1. Phosphorylate 5′ OH pre-miRNA oligos with γ-^{32} ATP by incubating the reaction containing the following components at 37 °C for 30 min (*see* **Note 4**).
   - 11 μL water.
   - 3 μL 10× T4 Polynucleotide Kinase buffer.
   - 3 μL 10 μM 5′ OH pre-miRNA.
   - 3 μL 10 units/μL T4 Polynucleotide Kinase.
   - 10 μL γ-^{32}P ATP (6000 Ci/mmol, 150 mCi/mL).
2. During the PNK reaction, prepare and pre-run a 1.5 mm thick 8% acrylamide gel. Follow instructions as described in **steps 2** and **3** in Subheading 3.2, Preparation of cold pre-miRNAs.

3. After the PNK reaction, remove free γ-^{32}P ATP using Illustra MicroSpin G-25 Columns following the manufacturer's instructions.
4. Mix with an equal amount of 2× formamide RNA loading buffer. Incubate at 95 °C for 2 min. Load samples onto gel. Keep the microcentrifuge tube that contained radiolabeled RNA to prepare radioactive position markers (*see* **Step 7**).
5. Run the gel until the faster bromophenol blue dye almost runs off the gel and the slower xylene cyanol dye is about halfway down the gel.
6. After electrophoresis, remove one of the glass plates and cover the gel with plastic wrap.
7. Prepare radioactive position markers to mark location on the gel. Using the microcentrifuge tube which contained radiolabeled RNA from **step 4**, add ~20 μL of 2× formamide RNA loading buffer to the microcentrifuge tube. This ~20 μL solution will be radioactive and will be visible when scanning on the Phosphorimager. Cut Whatman 3MM filter paper into small pieces and attach one piece to Scotch tape. Repeat to make four position markers. Apply ~5 μL of radioactive RNA formamide loading buffer solution from the tube in **step 8** onto the small piece of the filter paper attached to Scotch tape.
8. Attach the radioactive position markers to the plastic wrap covering the gel, placing one at each of the four corners of the gel. These radioactive paper pieces will be used as position markers to identify location of the radioactive RNAs on the gel to facilitate gel excision after imaging the gel.
9. Cover the gel with another sheet of plastic wrap, sandwiching the radioactive filter paper pieces between the plastic wrap layers.
10. Expose the gel to a Storage Phosphor Screen for ~10 s and visualize the RNA bands and the position marker pieces by scanning the screen on a Phosphorimager.
11. Print out the scanned image at actual size. Insert the print out under the bottom gel glass plate and align it based on the position marker signals. Mark the RNA bands.
12. Cut out the radioactive pre-miRNA bands with a razor blade and put the cut out gel pieces in a microcentrifuge tube. Add 1 mL 0.3 M NaCl and elute RNA from the cutout gel pieces by rotating the tube on a tube rotator at 4 °C overnight (*see* **Note 5**).
13. Filtrate eluted solutions using Spin-X Centrifuge Tube Filters in order to remove gel pieces. Perform ethanol precipitation to further purify RNA. Add 1 μL of 10 mg/mL glycogen and three volumes of ethanol to each solution. Split the solutions

into two or three microcentrifuge tubes as needed. Store at −80 °C for at least 1 h.

14. Centrifuge samples to pellet RNA at 21,000 × *g* for 30 min at 4 °C. Remove supernatant and rinse pellets with 80% (v/v) ethanol. Then air dry.
15. Resuspend RNA pellets in 100 μL water. Measure the radioactivity (CPM/μL; CPM: Counts Per Minute) using a scintillation counter. Store at −80 °C.

3.4 Preparation of Cold Long dsRNAs

1. Using a high-fidelity PCR enzyme that makes blunt ends, such as Phusion DNA polymerase and a pEGFP-N3 plasmid as PCR template, PCR-amplify the T7 RNA polymerase template dsDNA containing the T7 promoter for the sense RNA strand using the forward and reverse primers for the sense strand template in a 400 μL scale. The reaction contains the following components at the following temperature cycles (*see* **Note 6**). The PCR product is referred as "T7 template for sense RNA strand" below in this protocol.
   - 296 μL water.
   - 80 μL 5× Phusion® HF Buffer.
   - 8 μL dNTP mix (10 mM each).
   - 4 μL 20 ng/μL pEGFP-N3 plasmid.
   - 4 μL 100 μM Forward DNA primer for the 106 nt sense strand template.
   - 4 μL 100 μM Reverse DNA primer for the 106 nt sense strand template.
   - 4 μL 2000 units/mL Phusion® High-Fidelity DNA Polymerase.

   PCR step 1: 98 °C for 30 s.

   PCR step 2: 98 °C for 10 s.

   PCR step 3: 55 °C for 20 s.

   PCR step 4: 72 °C for 15 s.

   PCR step 5: go back to step 2 for 34 cycles.

   PCR step 6: 72 °C for 60 s.

   PCR step 7: pause at 4 °C.
2. Similarly, PCR-amplify the T7 RNA polymerase template dsDNA containing the T7 promoter for the antisense RNA strand using the forward and reverse primers for the antisense strand template in a 400 μL scale, as in **step 1**. The PCR product is referred as "T7 template for antisense RNA strand" below in this protocol.
3. After the PCR reaction, add an equal amount of phenol/chloroform/isoamyl alcohol (25:24:1 mixture, pH 6.7). vortex, centrifuge, and transfer aqueous phase to a new tube.

4. Perform ethanol precipitation by adding 1/10 volume of 3 M sodium-acetate (pH 5.2) and 2.5 volumes of ethanol. Precipitate the pellet by centrifugation at 21,000 × *g* for 20 min at 4 °C, wash the pellets with 70% (v/v) ethanol, and air dry. Resuspend pellets in 400 μL water.
5. Combine the following in a 50 mL tube in the order listed (add T7 RNA polymerase enzyme last). Keep the components for the reaction on ice, but add them to a reaction tube at room temperature. Perform the reaction for sense and antisense strands separately. Incubate at 37 °C for 2–4 h.
   - 500 μL 10× T7 transcription buffer.
   - 2250 μL water.
   - 50 μL 100 mM DTT.
   - 200 μL 100 mM ATP.
   - 200 μL 100 mM CTP.
   - 200 μL 100 mM GTP.
   - 200 μL 100 mM UTP.
   - 800 μL 100 mM GMP.
   - 100 μL 40 unit/μL RNasin RNase inhibitor.
   - 100 μL 0.5 unit/μL IPP.
   - 200 μL template DNA from the previous step (T7 template for sense RNA strand or T7 template for antisense RNA strand).
   - 200 μL T7 RNA polymerase.
6. Add 50 μL 1 unit/μL RQ1 RNase-free DNase and incubate at 37 °C for 30 min.
7. Perform phenol/chloroform extraction. Add an equal amount (5 mL) of phenol/chloroform, vortex, centrifuge, and transfer supernatant to a new tube.
8. Ethanol precipitate RNA by adding 1/10 volume (500 μL) of 3 M sodium acetate (pH 5.2) and 2.5 volumes (12.5 mL) of ethanol. Vortex, centrifuge, and wash the RNA pellets with 80% (v/v) ethanol. After air-drying the RNA pellet, resuspend in 400 μL water.
9. Dephosphorylate RNA by adding 46 μL of 10× CIP buffer and 20 μL of Alkaline Phosphatase, Calf Intestinal. Incubate at 37 °C for 60 min (*see* **Note 7**).
10. Add an equal amount (466 μL) of phenol/chloroform/isoamyl alcohol (25:24:1 mixture, pH 6.7), vortex, centrifuge, and transfer the aqueous phase to a new tube.
11. Perform ethanol precipitation by adding 1/10 volume of 3 M Sodium acetate (pH 5.2) and 2.5 volumes of ethanol. Vortex,

centrifuge at 21,000 × *g* for 30 min at 4 °C, and wash RNA pellets with 80% (v/v) ethanol. After air-drying the RNA pellets, resuspend in 400 µL water.

12. Phosphorylate RNAs by adding the following reagents and incubating the reaction at 37 °C for 60 min.
    - 48 µL 10× T4 Polynucleotide Kinase buffer.
    - 8 µL 100 mM ATP.
    - 20 µL 10 units/µL T4 Polynucleotide Kinase.
13. Add an equal amount (500 µL) of 2× formamide RNA loading buffer. Incubate at 95 °C for 2 min. Gel purify RNAs as in Subheading 3.2.

3.5 Preparation of Internally Radiolabeled Long dsRNAs

1. Set up transcription reaction for only the sense strand. Combine the following in a microcentrifuge tube in the order listed (add the T7 RNA polymerase enzyme last). Keep the components for the reaction on ice, but add to reaction tube at room temperature and incubate the sample at 37 °C for 2–4 h (*see* **Note 8**).
    - 5 µL 10× T7 transcription buffer.
    - 2.5 µL water.
    - 0.5 µL 100 mM DTT.
    - 2 µL 100 mM ATP.
    - 2 µL 100 mM CTP.
    - 2 µL 100 mM GTP.
    - 2 µL 1 mM UTP.
    - 8 µL 100 mM GMP.
    - 20 µL α-^{32}P UTP (800 Ci/mmol, 10 mCi/mL).
    - 1 µL 40 unit/µL RNasin RNase inhibitor.
    - 1 µL 0.5 unit/µL IPP.
    - 2 µL template DNA from Subheading 3.4, **step 4** (T7 template for sense RNA strand).
    - 2 µL T7 RNA polymerase.
2. Add 1 µL of 1 unit/µL RQ1 RNase-free DNase and incubate at 37 °C for 30 min.
3. Remove free α-^{32}P UTP using G-25 Spin Columns following the manufacturer's protocol.
4. Add 350 µL water and 400 µL phenol/chloroform, vortex, centrifuge, and transfer supernatant to a new tube.
5. Precipitate RNA by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ethanol. Vortex, centrifuge at 21,000 × *g* for 30 min at 4 °C, and wash RNA pellet with 80% (v/v) ethanol. After air-drying, resuspend RNA in 43 µL water.

6. Dephosphorylate RNA by adding 5 μL 10× CIP buffer and 2 μL Alkaline Phosphatase, Calf Intestinal. Incubate at 37 °C for 60 min (*see* **Note 7**).
7. Add 350 μL water and 400 μL phenol/chloroform to the RNA solution. Vortex, centrifuge, and transfer supernatant to a new tube.
8. Precipitate RNA by adding 1/10 volume 3 M sodium acetate (pH 5.2) and 2.5 volumes ethanol. Vortex, centrifuge at 21,000 × *g* for 30 min at 4 °C, and wash the RNA pellet with 80% (v/v) ethanol. After air-drying the RNA pellet, resuspend in 42 μL water.
9. Set up 5′ monophosphorylation reaction. Add the following to the RNA and incubate the reaction at 37 °C for 60 min:
   - 5 μL 10× T4 polynucleotide kinase buffer.
   - 1 μL 100 mM ATP.
   - 2 μL 10 units/μL T4 polynucleotide kinase.
10. Add an equal volume of 2× formamide RNA loading buffer. Incubate at 95 °C for 2 min. Gel purify the RNA as in Subheading 3.3.

3.6 Performing In Vitro RNA Processing Reactions

1. Radiolabel the Decade maker using γ-^{32}P ATP (6000 Ci/mmol 150 mCi/mL), following the manufacturer's protocol. Prepare an aliquot of 1/100 dilution as working concentration. Store both the diluted and undiluted marker solutions in −80 °C.
2. Prepare radioactive RNA substrates at an eightfold concentration of final concentration to be tested in the processing assays. For example, if the final concentration of RNA in the processing assay is 100 nM and 25,000 CPM/μL RNA substrate, prepare 800 nM and 200,000 CPM/μL RNA by mixing appropriate amounts of cold and radioactive RNA stocks and water. For dsRNA, assemble the duplex by adding the radioactive sense strand stock, the cold sense strand stock, the cold antisense strand stock, and water, to achieve 1:1.1 molar ratio between the sense and antisense strands at a desired concentration and radioactivity. Antisense RNA is added slightly in excess so that all radioactive sense strand is in duplex form.
3. Incubate at 80 °C for 2 min and cool gradually to room temperature.
4. Prepare the following reaction in a microcentrifuge tube. The concentration shown is the final concentration in the reaction. Incubate PNK reaction at 37 °C for 60 min. (Before adding the RNA substrate at the last step to start the processing reaction, pre-incubate the mixture at 25 °C for a few minutes). The reaction is typically performed in final 8 μL volume. Start the reaction by adding 1 μL of radioactive RNA substrate to 7 μL of

pre-incubated reaction pre-mixture. As a negative control, make the reaction sample without any Dicer enzyme. Take time zero sample for this negative control (*see* **Notes 9–11**).

- 18 mM HEPES-KOH (pH 7.4).
- 100 mM potassium acetate.
- 3 mM magnesium acetate.
- 1 mM ATP.
- 1 mM DTT.
- 1.5% (v/v) glycerol.
- 0.1 mg/mL BSA.
- 0.2 unit/μL RNasin RNase inhibitor.
- Dicer enzyme.
- Radioactive RNA.

5. Take an aliquot (typically 1 μL) and add 25-fold volume of 2× formamide RNA loading buffer (typically 25 μL) at a defined time points. The processing reaction is stopped in the 2× formamide RNA loading buffer.
6. Coat the notched glass plates with hydrophobic glass treatment (e.g., Rain-X original glass treatment), and coat the back plates with hydrophilic glass treatment (e.g., Rain-X anti-fog).
7. Make 0.4 mm thick, 1× TBE, 7 M Urea, 10% acrylamide PAGE gel. Pre-run the gel using 0.5× TBE as a running buffer in the vertical gel electrophoresis apparatus. Use the aluminum plate to disperse the heat generated during the gel running. Gel should be pre-run for about 30 min at a constant power of ~35 W before loading RNA samples.
8. Heat samples at 95 °C for 5 min. Load the samples onto gel. Load the radiolabeled decade marker as well.
9. Run gel at ~35 W until the faster bromophenol blue dye and the slower xylene cyanol dye are about 2/3 and 1/3, respectively, down the gel.
10. Remove one of the glass plates. Cover the gels with Whatman™ 3MM Chr Chromatography Paper. Flip the gels and then remove the other glass plate. Cover the gels with a thermal laminating film.
11. Dry gels using a gel drier system.
12. Expose dried gels to Storage Phosphor Screen.
13. Scan exposed screens on a phosphorimager (Fig. 2).

3.7 Analyzing Enzyme Kinetics

1. Quantify the unprocessed substrate bands, produced small RNA bands in each lane, and a background signal (Fig. 3).
2. Subtract the background signal from each of the quantitated signals.

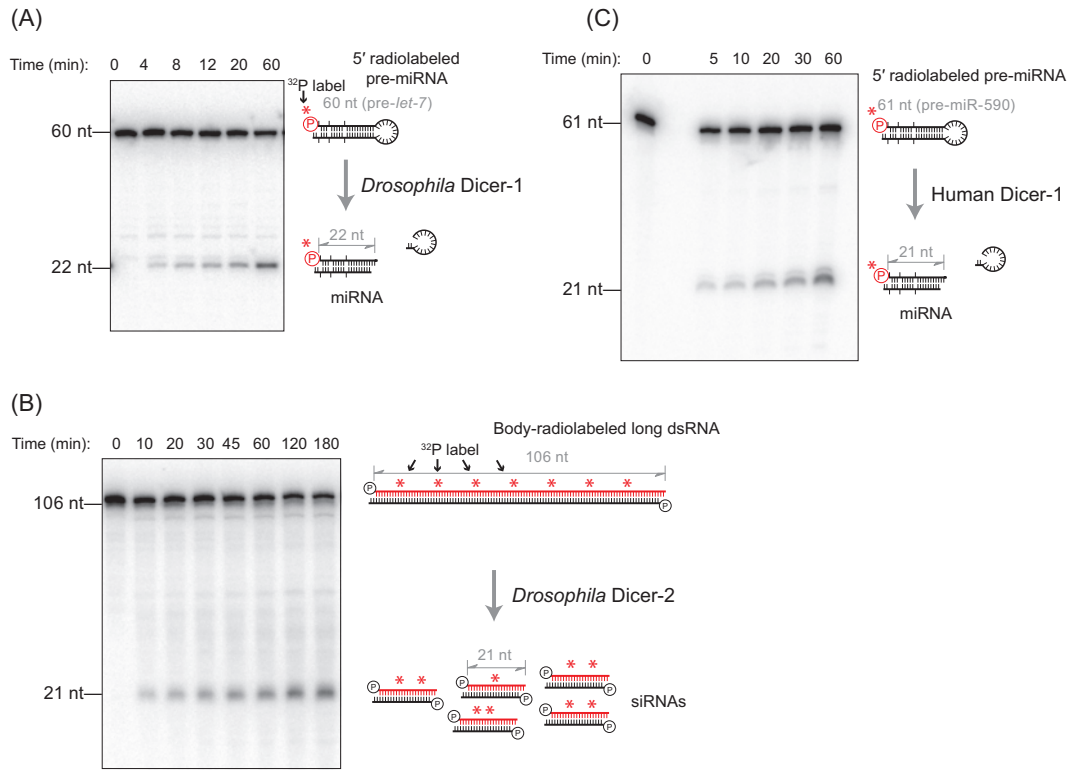
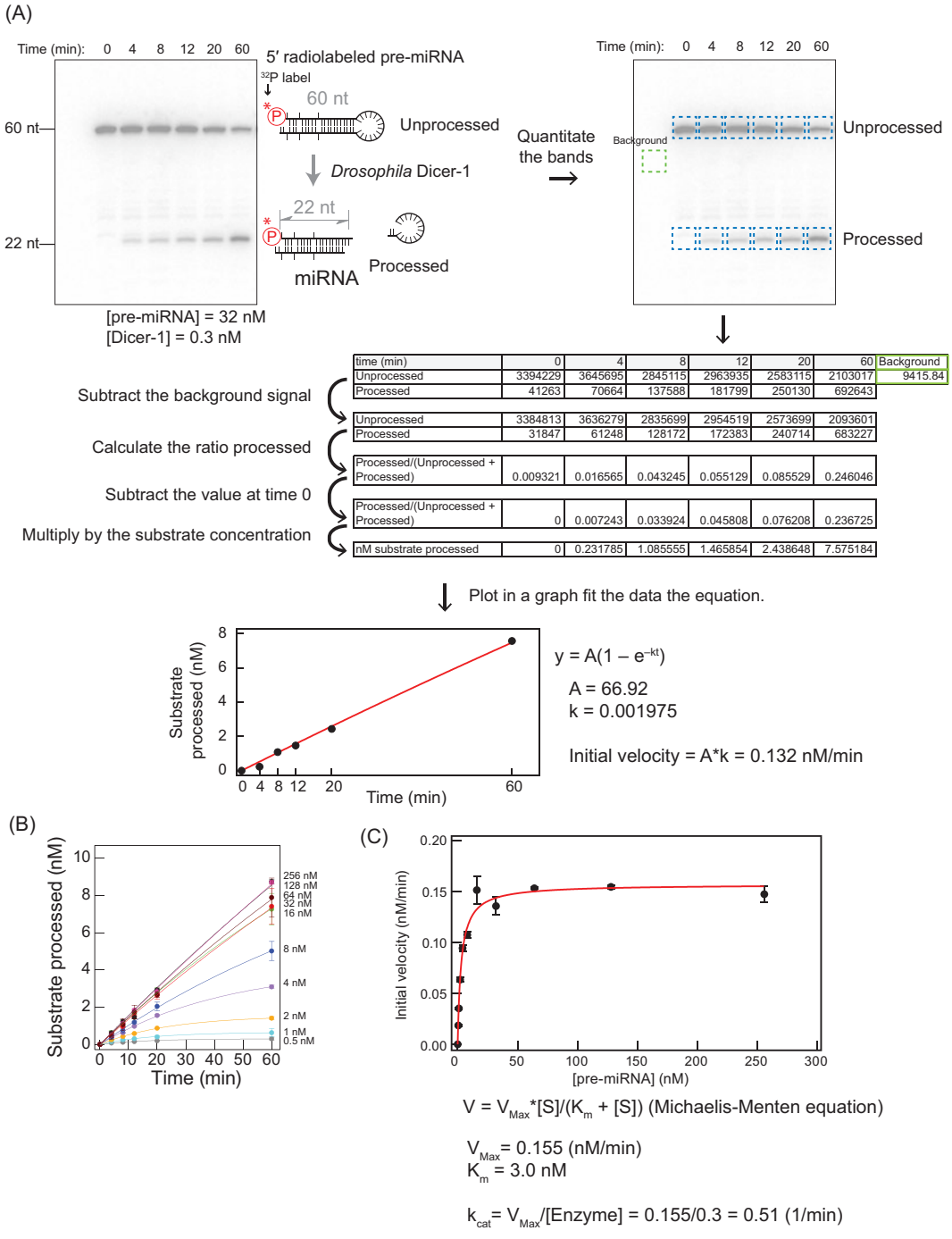


Fig. 2 Gel images of the Dicer in vitro RNA processing assay. (**a**) Assay for the miRNA production from 5′ radiolabeled pre-miRNA by *Drosophila* Dicer-1 [17]. 0.3 nM Dicer enzyme and 32 nM pre-*let-7* were used in this particular image. (**b**) Assay for the siRNA production from internally radiolabeled 104 bp long dsRNA by Dicer-2 [15]. 8 nM Dicer-2 and 100 nM long dsRNA were used in this particular image. (**c**) Assay for the miRNA production from 5′ radiolabeled pre-miRNA by human Dicer [18]. 1 nM human Dicer enzyme and 100 nM pre-miR-590 were used in this particular image. Asterisks indicate the ^{32}P radiolabels

3. Calculate the fraction of the substrate processed = small RNA product/(unprocessed substrate + small RNA product), in each lane.
4. Subtract the fraction of the substrate processed at time 0 (in the no enzyme condition) from each of the reaction time points.
5. Multiply fractions of the substrate processed with the substrate concentration used, and calculate the concentration of the substrate processed.
6. Plot the time course data in a graph. Fit the data to $y = A(1 - e^{-kt})$, where $dy/dt = Ake^{-kt}$. When $t = 0$, $dy/dt = Ak$; Ak gives the initial velocity of the reaction.
7. For Michaelis–Menten analysis, test multiple substrate concentrations and determine the initial rate for each of the conditions (Fig. 3b). Fit the initial rates in each substrate concentration condition to the Michaelis–Menten scheme $V = V_{Max} \times [S]/(K_m + [S])$ in Igor Pro (Fig. 3c). Calculate k_{cat} by dividing V_{max} with the enzyme concentration used.



time (min)	0	4	8	12	20	60	Background
Unprocessed	3394229	3645695	2845115	2963935	2583115	2103017	9415.84
Processed	41263	70664	137588	181799	250130	692643	
Unprocessed	3384813	3636279	2835699	2954519	2573699	2093601	
Processed	31847	61248	128172	172383	240714	683227	
Processed/(Unprocessed + Processed)	0.009321	0.016565	0.043245	0.055129	0.085529	0.246046	
Processed/(Unprocessed + Processed)	0	0.007243	0.033924	0.045808	0.076208	0.236725	
nM substrate processed	0	0.231785	1.085555	1.465854	2.438648	7.575184	

Fig. 3 Kinetic analysis of the in vitro RNA processing by Dicer. (**a**) An example of the kinetic analysis of 16 nM 5′ radiolabeled pre-*let-7* processing by 0.3 nM *Drosophila* Dicer-1. (**b**) Reaction time courses for each substrate concentrations of pre-*let-7* processing by 0.3 nM *Drosophila* Dicer-1. (**c**) The initial velocities calculated from the data in (**b**) were fit to the Michaelis-Menten scheme $V = V_{Max} \times [S]/(K_m + [S])$ and K_m, V_{max}, k_{cat} were determined [17]

4 Notes

1. The N-terminal His-tag can be cleaved off using a TEV protease, if preferred.
2. Alternatively, 5′ monophosphorylated pre-miRNA can be ordered as custom synthesis. In this case, 5′ OH pre-miRNA required for radioactive oligo preparation in Subheading 3.3 can be prepared by treating the 5′ monophosphorylated pre-miRNAs with Alkaline Phosphatase, Calf Intestinal.
3. Pre-miRNAs are usually found around halfway down the gel (migrate half the distance compared with the bromophenol blue dye). This position would be close to the xylene cyanol bands if xylene cyanol is included in the loading buffer. Xylene cyanol bands make shadow in the UV shadowing method and they may overlap with the RNA shadows, making it difficult to distinguish them. To avoid this potential problem, the loading buffer without xylene cyanol is preferred.
4. The protocol involving the radioactive materials has to be performed by authorized personnel and according to the local/institutional rules and regulations for work with radioactive substances. Use pipette tips provided with an aerosol filter to avoid radioactive contamination of pipetmans.
5. Confirm whether RNAs were excised from appropriate gel position by imaging the remaining gel using Phosphor Screen.
6. Each strand of the dsRNAs shown in this protocol is 106 nt long. The dsRNA is 104 bp long with 5′ monophosphorylated, 3′ two-nucleotide overhang at both ends.
7. Since the T7 RNA transcription reaction contained GMP, most RNA transcripts contain 5′ monophosphate while some may contain 5′ triphosphate. In order to make homogeneous 5′ monophosphate RNA, all phosphates (5′ monophosphate and 5′ triphosphate) will be removed first, and then 5′ monophosphorylation reaction will be performed.
8. Note that 1/100 fold of cold UTP is used here compared with the cold RNA preparation in Subheading 3.2. α-^{32}P ATP, α-^{32}P CTP, or α-^{32}P GTP, can also be used instead of α-^{32}P UTP, and the concentration of the respective cold nucleotide should be reduced by 100 times.
9. The concentrations of Dicer enzyme and RNA substrate as well as reaction time points may need to be optimized after an initial preliminary trial so that enough processing activities are observed but without reaction being saturated. Typically, 1–10 nM of Dicer enzyme and 30–100 nM RNA substrate are used. The first time point is taken at 4–10 min and several time points until the final time point at 60–180 min. For Michaelis-Menten

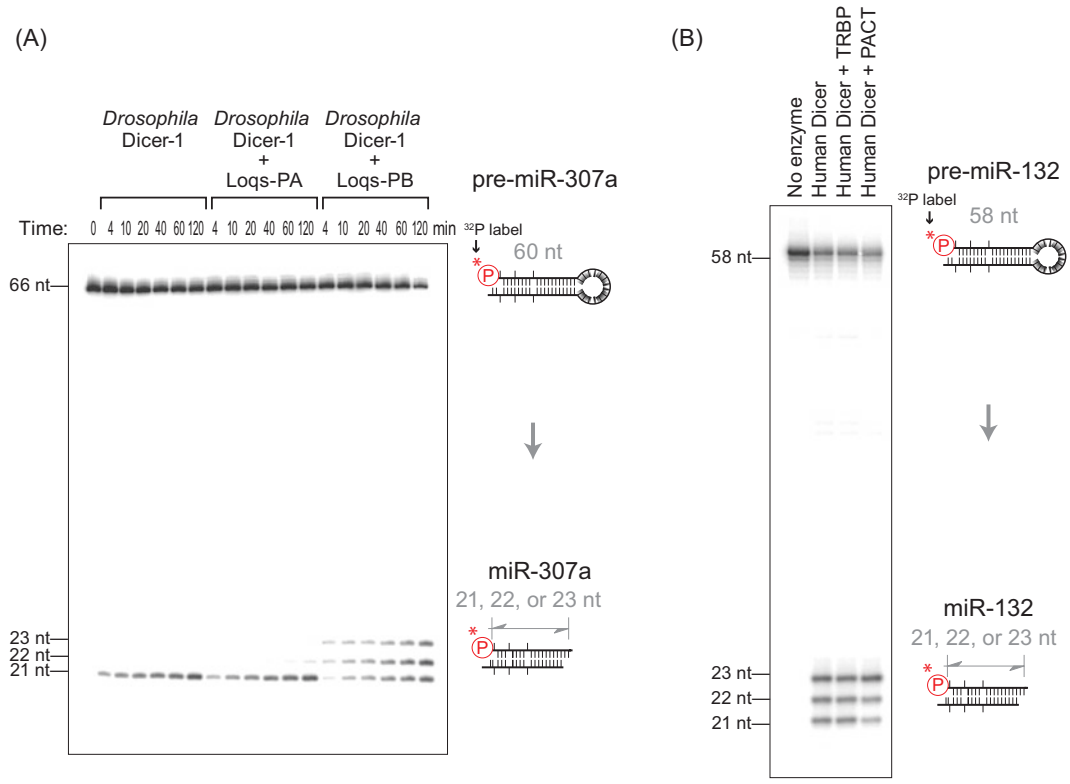


Fig. 4 Analysis of the effects of the Dicer partner proteins on the length of the miRNAs produced by Dicer. (**a**) 30 nM 5′ radiolabeled pre-miR-307a processing by 8 nM *Drosophila* Dicer-1 alone or Dicer-1 supplemented with purified, recombinant Loqs-PA or Loqs-PB (8 nM) [17]. (**b**) 100 nM 5′ radiolabeled pre-miR-132 processing by 10 nM human Dicer alone or Dicer supplemented with purified, recombinant TRBP or PACT (10 nM) [17]. Asterisks indicate the ^{32}P radiolabels

analysis, which requires testing various substrate concentrations, the enzyme concentration may need to be decreased and/or earlier time points may need to be analyzed.

10. The in vitro RNA processing assay can be performed in the presence of Dicer partner proteins (*Drosophila* Loquacious-PA [Loqs-PA] and Loqs-PB, and human TRBP and PACT) to test how they affect the processing. Loqs-PB and TRBP, but not Loqs-PA or PACT, were shown to alter the length of a subset of miRNAs produced by Dicers (Fig. 4) [12]. The bacterial expression constructs of these Dicer partner proteins were published [12] and are available from Addgene (#41097, 41094, 41096, 41093).
11. For the higher resolution gel electrophoresis as in Fig. 4, we use Glass Plate Set Notched, 33 × 42 cm (C.B.S. Scientific. SGP33-040A) and Sequencing Spacer Set 0.4 mm Thickness × 42 cm(l) (C.B.S. Scientific. SGS-4040).

Acknowledgments

This work was supported by a grant from American Heart Association (15SDG23220028), a grant from NIH (R01GM 116841), and funds provided by Johns Hopkins School of Medicine Department of Biological Chemistry to RF.

References

1. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409(6818):363–366
2. Hutvagner G et al (2001) A cellular function for the RNA-interference enzyme dicer in the maturation of the let-7 small temporal RNA. Science 293(5531):834–838
3. Ketting RF et al (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev 15(20):2654–2659
4. Knight SW, Bass BL (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science 293(5538):2269–2271
5. Jiang F et al (2005) Dicer-1 and R3D1-L catalyze microRNA maturation in drosophila. Genes Dev 19(14):1674–1679
6. Macrae IJ et al (2006) Structural basis for double-stranded RNA processing by dicer. Science 311(5758):195–198
7. Ma E, MacRae IJ, Kirsch JF, Doudna JA (2008) Autoinhibition of human dicer by its internal helicase domain. J Mol Biol 380(1): 237–243
8. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA (2008) In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci U S A 105(2):512–517
9. Cenik ES et al (2011) Phosphate and R2D2 restrict the substrate specificity of dicer-2, an ATP-driven ribonuclease. Mol Cell 42(2):172–184
10. Park JE et al (2011) Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature 475(7355):201–205
11. Welker NC et al (2011) Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Mol Cell 41(5):589–599
12. Fukunaga R et al (2012) Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell 151(3):533–546
13. Tsutsumi A, Kawamata T, Izumi N, Seitz H, Tomari Y (2011) Recognition of the pre-miRNA structure by drosophila dicer-1. Nat Struct Mol Biol 18(10):1153–1158
14. Lee HY, Doudna JA (2012) TRBP alters human precursor microRNA processing in vitro. RNA 18(11):2012–2019
15. Fukunaga R, Colpan C, Han BW, Zamore PD (2014) Inorganic phosphate blocks binding of pre-miRNA to dicer-2 via its PAZ domain. EMBO J 33(4):371–384
16. Sinha NK, Trettin KD, Aruscavage PJ, Bass BL (2015) Drosophila dicer-2 cleavage is mediated by helicase- and dsRNA termini-dependent states that are modulated by loquacious-PD. Mol Cell 58(3):406–417
17. Lin X et al (2016) Common miR-590 variant rs6971711 present only in African Americans reduces miR-590 biogenesis. PLoS One 11(5):e0156065
18. Kandasamy SK, Fukunaga R (2016) Phosphate-binding pocket in dicer-2 PAZ domain for high-fidelity siRNA production. Proc Natl Acad Sci U S A 113(49):14031–14036
19. Kandasamy SK, Zhu L, Fukunaga R (2017) The C-terminal dsRNA-binding domain of *Drosophila* Dicer-2 is crucial for efficient and high-fidelity production of siRNA and loading of siRNA to Argonaute2. RNA. 23(7):1139–1153. doi:10.1261/rna.059915.116. Epub 2017 Apr 17. PMID:28416567
20. Lee YS et al (2004) Distinct roles for drosophila dicer-1 and dicer-2 in the siRNA/miRNA silencing pathways. Cell 117(1):69–81
21. Lau PW et al (2012) The molecular architecture of human dicer. Nat Struct Mol Biol 19(4):436–440
22. Lau PW, Potter CS, Carragher B, MacRae IJ (2009) Structure of the human dicer-TRBP complex by electron microscopy. Structure 17(10):1326–1332
23. Wang HW et al (2009) Structural insights into RNA processing by the human RISC-loading

complex. Nat Struct Mol Biol 16(11): 1148–1153

24. Taylor DW et al (2013) Substrate-specific structural rearrangements of human dicer. Nat Struct Mol Biol 20(6):662–670

25. Kao C, Zheng M, Rudisser S (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. RNA 5(9):1268–1272

Chapter 8

Nucleic Acid-Binding Assay of Argonaute Protein Using Fluorescence Polarization

Tomohiro Miyoshi

Abstract

Nucleic acid binding by the Argonaute protein is an important trigger step in the Argonaute-dependent gene silencing system. We established an in vitro method to detect the nucleic acid binding activity of the Argonaute protein by fluorescence polarization. In this chapter, we will describe the expression and purification of the prokaryotic (*Rhodobacter sphaeroides*) Argonaute protein, and the nucleic acid-binding analysis using a Fluorescence Polarization System (Beacon 2000).

Key words Fluorescence polarization assay, Beacon 2000, *Rhodobacter sphaeroides* argonaute, Guide RNA, Purification of prokaryotic argonaute

1 Introduction

Argonaute proteins are evolutionarily conserved between eukaryotes and prokaryotes [1], and play key roles in gene silencing and host defense systems, in collaboration with small guide strands that recognize complementary sequences in the target strands [2]. Although the overall secondary structures of Argonaute proteins are very similar, each Argonaute protein associates with its respective specific combinations of RNA and DNA. In general, eukaryotic Argonaute proteins recognize RNAs as both the guide and target strands, while prokaryotic Argonaute proteins recognize each appropriate combination of RNA and DNA as the guide and target strands. The prokaryotic Argonaute proteins from *R. sphaeroides* [3, 4] and *M. piezophila* [5] preferentially associate with RNAs as the guide strands and DNAs as the target strands. In contrast, the proteins from *T. thermophilus* [6] and *P. furiosus* [7] preferentially associate with DNAs as the guide and target strands.

The binding of the guide strand to Argonaute is the first functional trigger step and a crucial reaction for the silencing mechanisms. The nucleic acid binding activities of Argonaute have mainly been measured by the double-filter method [8], using nitrocellulose-filter

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_8,

binding with 5′-^{32}P labeled nucleotides [9, 10]. Furthermore, the crystal structures of Argonaute proteins bound to guide strands have been investigated, to clarify the binding mechanism. The binary structures of Argonaute bound to guide RNAs were obtained for the human, yeast and prokaryotic systems, while the binary structure of Argonaute bound to a guide DNA was only obtained from a prokaryote. These structures revealed the following common features: the 5′ end of the guide strand is anchored within a nucleotide-binding pocket in the MID domain and the 3′ end interacts with the PAZ domain. However, the mechanisms for RNA/DNA selection by Argonaute remain poorly understood at this point.

This study is primarily focused on the guide strand binding activity of prokaryotic Argonaute. In this chapter, to assess the binding activity (Kd: equilibrium dissociation constant) of the Argonaute protein to the guide strand, protocols are presented for the calculation of the binding capacity between prokaryotic (*R. sphaeroides*) Argonaute (RsAgo) and a 5′ phosphorylated guide RNA (18 nt), using a fluorescence polarization system. Using this system, accurate Kd values of the Argonaute-guide strand complex can be obtained without radioisotope-labeled nucleic acids.

2 Materials

2.1 Expression and Purification of RsAgo Protein

1. RsAgo expression plasmid: pET-28a (Km^r) plasmid containing the RsAgo gene [3].
2. BL21 (DE3) Competent cells (Novagen).
3. LB broth (Becton, Dickinson and Company).
4. LB-kanamycin broth: LB broth, 15 mg/ml kanamycin.
5. Buffer A: 50 mM HEPES-KOH (pH 7.5), 1 M NH_4Cl, 5% glycerol and 5 mM β-mercaptoethanol.
6. Buffer B: 50 mM HEPES-KOH (pH 7.5), 1 M NH_4Cl, 5% glycerol, 15 mM imidazole and 5 mM β-mercaptoethanol.
7. Buffer C: 50 mM HEPES-KOH (pH 7.5), 1 M NH_4Cl, 5% glycerol, 300 mM imidazole and 5 mM β-mercaptoethanol).
8. Buffer D: 20 mM Tris–HCl (pH 7.6), 300 mM NaCl and 5 mM β-mercaptoethanol.
9. Buffer E: 20 mM Tris–HCl (pH 7.6), 500 mM NaCl and 5 mM β-mercaptoethanol.
10. Binding buffer: 20 mM HEPES-KOH (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 0.01 mg/ml *E. coli* total tRNA and 0.01% NP-40.
11. Shaking incubator.
12. Isopropyl β-D-1-thiogalactopyranoside (IPTG).

13. Ultrasonic processor UP200S (Hielscher).
14. Centrifuge.
15. Ni-NTA agarose (Qiagen).
16. Anion exchange column, RESOURCE Q (GE Healthcare).
17. HiLoad Superdex 200 pg column (GE Healthcare).
18. Fast protein liquid chromatography: AKTA purifier 100 (GE Healthcare).
19. Centrifugal filter: Amicon Ultra Ultracel-10 k (Merck Millipore).

2.2 Fluorescence Polarization Assay

1. Fluorescence polarization system (Beacon 2000).
2. Glass tubes (Borosilicate Glass Disposable Culture Tubes, Kimble).
3. 5′ phosphorylated (p) RNA (18 nt) labeled with 6-carboxy-fluorescein (6-FAM): 5′- p-UUACAACCUACUACCUCG-6-FAM -3′ (GeneDesign, Inc.) (*see* **Note 1**).
4. Scientific software: GraphPad Prism 6 (GraphPad Software) .

3 Methods

3.1 Expression and Purification of RsAgo Protein

1. Transform *E. coli* BL21 (DE3) competent cells with the RsAgo expression plasmid.
2. Cultivate the transformed *E. coli* cells in LB-kanamycin medium to an OD_{600} of 0.6 at 37 °C, and induce protein expression with 0.5 mM IPTG. Culture the cells further at 25 °C for 18 h, and collect them by centrifugation for 20 min at 6000 × *g*. The recombinant protein has a 6× His-tag at the N-terminus.
3. Resuspend the pellet from the 1000 ml bacterial culture in 20 ml buffer A.
4. Sonicate on ice for 10 min, 2 cycles, using an ultrasonic processor.
5. Centrifuge for 15 min at 10,000 × *g* and 4 °C.
6. Centrifuge the supernatant further for 30 min at 100,000 × *g* and 4 °C.
7. Apply the crude extract to Ni-NTA agarose resin equilibrated with buffer A, by open column chromatography.
8. Wash with buffer B with three resin volumes (*see* **Note 2**).
9. Elute with buffer C.
10. Dialyze the sample against buffer D and then pass it through an anion exchange column (Resource Q) with the AKTA system, in order to remove the nucleic acid bound to the Argonaute protein. Under these buffer conditions, the protein

sample is in the flow through fraction and the nucleic acids are trapped on the column (*see* **Note 2**).

11. Concentrate the sample with an Amicon Ultra filter and subject it to size-exclusion chromatography with buffer D, using the AKTA system.
12. Dialyze the sample fractions against buffer E, concentrate the sample by using an Amicon Ultra filter, and then store the sample at −80 °C (*see* **Note 3**).

3.2 Fluorescence Polarization Assay

1. Using 18 glass tubes, serially dilute the RsAgo protein from Subheading 3.1 with binding buffer in the total volume of 100 μl (Table 1).

Table 1
Protein concentration in each tube and theoretical total amount of protein required to measure the dissociation constant (Kd) of RsAgo binding to single-stranded RNA. In this example, the protein has a molecular weight of 87,000 Da (RsAgo) and binds to the RNA with a Kd of 1 nM

Tube No.	Protein concentration (nM)	Protein/Tube (ng)
1	0.026	0.23
2	0.039	0.34
3	0.059	0.51
4	0.088	0.77
5	0.132	1.15
6	0.196	1.71
7	0.296	2.58
8	0.444	3.86
9	0.667	5.80
10	1.000	8.70
11	1.500	13.05
12	2.250	19.58
13	3.375	29.36
14	5.063	44.05
15	7.594	66.07
16	11.391	99.10
17	17.086	148.65
18	25.629	222.97
		Total 668.48

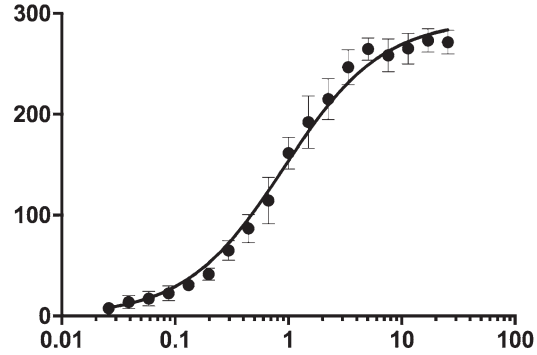


Fig. 1 Fluorescence polarization detection of the interaction between RsAgo and 18 nt single-stranded RNA. The Kd value is 0.91 ± 0.06 nM. Error bars represent s.d. values ($n = 3$)

2. Add 1 μl of 0.1 nM single-stranded RNA (18 nt) labeled with 6-FAM in the binding buffer to each tube, for a final concentration of 10 pM.
3. Incubate the tubes for 10 min at 25 °C.
4. Measure the fluorescence polarization of each tube at 25 °C, using the Beacon 2000 Fluorescence Polarization Instrument. The measurement temperature can be varied between 4 °C and 85 °C.
5. Fit the actual polarization data by nonlinear regression, using a computer program of one site-specific binding function with the GraphPad Software Prism 6. The binding curve and the dissociation constant (Kd) value are shown in Fig. 1 (*see* **Notes 4** and **5**).

4 Notes

1. The RsAgo protein interacts with 15–19 nt RNAs and prefers 5′-end uracil to adenine, guanine, and cytosine [3, 4].
2. To calculate the exact binding activity of RsAgo, the protein must be purified without nucleic acids. The nucleic acids bound to RsAgo are removed by a two-step strategy, including a high salt wash and anion exchange chromatography.
3. The RsAgo protein that is stored at −80 °C must be thawed in a high salt (500 mM NaCl) buffer to avoid precipitation of the protein.
4. The polarization value is the ratio of vertical and horizontal light intensities, and thus is a dimensionless number expressed in millipolarization (mP) units. Polarization value = (Vertical

intensity – Horizontal intensity)/(Vertical intensity + Horizontal intensity). The value is directly related to the molecular size (molecular rotation); therefore, when measuring the polarization of the guide RNA in the absence of the Argonaute protein, the labeled RNA can freely rotate and shows a low polarization value, and when the RNA is bound to the Argonaute protein, the large complex rotates more slowly and shows a high polarization value.

5. In this experiment, the 3′-end of the guide strand RNA is labeled by 6-FAM. The 3′-end of the single-stranded RNA is anchored within the nucleic acid binding pocket of the PAZ domain in the Argonaute-guide RNA binary complex [5, 11–18]. However, the deletion of the interaction between the 3′-end of the guide strand and the PAZ domain does not seem to have an effect on the binding affinity of the Argonaute-guide binary complex [19, 20].

Acknowledgments

This work was supported by the JSPS KAKENHI (Grant Numbers: 25840019 and 16K07246), the Takeda Science Foundation, the Naito Foundation, the Sasaki Environment Technology Foundation, and the Union Tool Scholarship Foundation.

References

1. Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J (2014) The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol 21:743–753
2. Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9:22–32
3. Miyoshi T, Ito K, Murakami R, Uchiumi T (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. Nat Commun 7:11846
4. Olovnikov I, Chan K, Sachidanandam R, Newman DK, Aravin AA (2013) Bacterial argonaute samples the transcriptome to identify foreign DNA. Mol Cell 51:594–605
5. Kaya E, Doxzen KW, Knoll KR, Wilson RC, Strutt SC, Kranzusch PJ, Doudna JA (2016) A bacterial Argonaute with noncanonical guide RNA specificity. Proc Natl Acad Sci U S A 113:4057–4062
6. Swarts DC, Jore MM, Westra ER, Zhu Y, Janssen JH, Snijders AP, Wang Y, Patel DJ, Berenguer J, Brouns SJ, van der Oost J (2014) DNA-guided DNA interference by a prokaryotic Argonaute. Nature 507:258–261
7. Swarts DC, Hegge JW, Hinojo I, Shiimori M, Ellis MA, Dumrongkulraksa J, Terns RM, Terns MP, van der Oost J (2015) Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. Nucleic Acids Res 43:5120–5129
8. Wong I, Lohman TM (1993) A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. Proc Natl Acad Sci U S A 90:5428–5432
9. Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ (2005) Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature 434:666–670
10. Yuan YR, Pei Y, Ma JB, Kuryavyi V, Zhadina M, Meister G, Chen HY, Dauter Z, Tuschl T, Patel DJ (2005) Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell 19:405–419

11. Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ (2008) Structure of the guide-strand-containing argonaute silencing complex. Nature 456:209–213
12. Schirle NT, MacRae IJ (2012) The crystal structure of human Argonaute2. Science 336:1037–1040
13. Elkayam E, Kuhn CD, Tocilj A, Haase AD, Greene EM, Hannon GJ, Joshua-Tor L (2012) The structure of human argonaute-2 in complex with miR-20a. Cell 150:100–110
14. Faehnle CR, Elkayam E, Haase AD, Hannon GJ, Joshua-Tor L (2013) The making of a slicer: activation of human Argonaute-1. Cell Rep 3:1901–1909
15. Nakanishi K, Ascano M, Gogakos T, Ishibe-Murakami S, Serganov AA, Briskin D, Morozov P, Tuschl T, Patel DJ (2013) Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. Cell Rep 3: 1893–1900
16. Suter SR, Sheu-Gruttadauria J, Schirle NT, Valenzuela R, Ball-Jones AA, Onizuka K, MacRae IJ, Beal PA (2016) Structure-guided control of siRNA off-target effects. J Am Chem Soc 138:8667–8669
17. Matsumoto N, Nishimasu H, Sakakibara K, Nishida KM, Hirano T, Ishitani R, Siomi H, Siomi MC, Nureki O (2016) Crystal structure of silkworm PIWI-clade argonaute siwi bound to piRNA. Cell 167:484–497
18. Schirle NT, Kinberger GA, Murray HF, Lima WF, Prakash TP, MacRae IJ (2016) Structural analysis of human argonaute-2 bound to a modified siRNA guide. J Am Chem Soc 138:8694–8697
19. Hur JK, Zinchenko MK, Djuranovic S, Green R (2013) Regulation of Argonaute slicer activity by guide RNA 3′ end interactions with the N-terminal lobe. J Biol Chem 288: 7829–7840
20. Deerberg A, Willkomm S, Restle T (2013) Minimal mechanistic model of siRNA-dependent target RNA slicing by recombinant human Argonaute 2 protein. Proc Natl Acad Sci U S A 110:17850–17855

Chapter 9

Reconstitution of RNA Interference Machinery

Shintaro Iwasaki and Yukihide Tomari

Abstract

Small RNAs, including small interfering RNAs (siRNAs) and microRNAs (miRNAs), silence protein expression from target mRNAs bearing their complementary sequences, via the formation of the effector complex called RNA-induced silencing complex (RISC). Although the mechanism of RISC assembly has been studied for nearly two decades, the detailed mechanism has still remained unclear in part due to the lack of a pure reconstitution system. Recently, we identified all the core proteins necessary for RISC assembly in flies and successfully recapitulated the assembly of catalytically active RISC with eight recombinant proteins. The reconstitution system provides a versatile framework for detailed studies of RISC assembly, including single molecule analysis as described in another chapter in this issue.

Key words siRNA, miRNA, Chaperone, Argonaute

1 Introduction

Small RNAs of 21–22 nucleotides (nt) in length, including siRNAs and miRNAs, are a major class of non-coding RNAs in cells. Small RNAs mediate cleavage, deadenylation, decapping, and/or translational repression of their complementary target mRNAs, by guiding the effector complex termed RNA-induced silencing complexes (RISCs) to the targets. Argonaute (Ago) proteins lie at the core of RISC and bind directly to small RNAs. Although many other proteins have been identified as components of RISC [1–3], here we refer to the simple small RNA:Ago complex as minimal "RISC."

Since siRNA-mediated target RNA cleavage (so-called RNA interference or RNAi) has been recapitulated in fly embryo lysate in 1999 [4], extensive biochemical efforts have been made to understand how active RISC is assembled. Experiments using lysates from fly embryos, human [5] and plant cells [6] have revealed that the assembly of RISC generally follows a complicated pathway. Typically, the pathway can be classified into two substeps: duplex loading and passenger ejection. In the duplex loading

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_9,

step, small RNA duplexes, which have been generated by RNase III protein Dicer and its partner double-stranded RNA-binding protein, are loaded into Ago proteins to form pre-RISC, with the aid of ATP [7–9] and the Hsc70/Hsp90 chaperone machinery [6, 10–12]. In the following passenger ejection step, one of the two strands (termed passenger strand) of the small RNA duplex in Ago protein is discarded from the complex. In contrast to the duplex loading step, passenger ejection does not consume ATP or require the Hsc70/Hsp90 chaperone machinery. Passenger ejection is facilitated by mismatches in the seed (guide positions (g) 2–7 from the 5′-end of the guide strand) and/or 3′ supplemental region (g12–15) [8, 9] with the help of the N domain of Ago protein [13], or the cleavage of the passenger strand by the catalytic activity of Ago itself [14–17]. The ejected passenger strand is further digested by the exonuclease C3PO [18, 19]. After passenger ejection, mature RISC is formed with Ago and a single-stranded guide small RNA, ready to recognize complementary target RNAs.

Although these crude lysate systems have been extremely useful for clarifying the molecular framework of RISC assembly, the absence of a pure reconstitution system posed an analytic hurdle to study further mechanistic details of this reaction. Recently, we identified all the necessary factors and successfully reconstituted the RISC assembly reaction [20] with eight proteins: Ago2, Dicer-2/R2D2 heterodimer, and 5 chaperone machinery proteins [heat shock protein (Hsp) 83 (a homolog of Hsp90), Heat shock cognate protein (Hsc) 70–4 (a homolog of Hsc70), Hsp70/Hsp90 organizing protein organizing protein (Hop), *Drosophila* DnaJ-like (Droj) 2 (a homolog of Hsp40), and p23]. The composition of the 5 chaperone machinery proteins is essentially the same as the ones necessary for steroid receptor maturation [21], highlighting the unexpected commonality of the two reactions; the ATP-driven chaperone machinery mediates a conformational change of client proteins to allow binding of their ligands.

Here, we describe the preparation procedures for the reconstitution system and the assay methods to evaluate the RISC assembly reaction. This system allows further detailed studies of RISC assembly as well as its application to single-molecule analysis described in another chapter in this issue.

2 Materials

1. Dounce homogenizer (WHEATON; 7 ml, “TIGHT” pestle).
2. Sonifier 250 (Branson).
3. Dynabeads Protein G for immunoprecipitation (ThermoFischer Scientific).

4. Magnetic stand [DynaMag-2 Magnet (ThermoFischer Scientific) or an equivalent one].
5. Anti-Flag M2 antibody (Sigma).
6. His A buffer: 30 mM HEPES-KOH pH 7.4, 200 mM KOAc, 2 mM $Mg(OAc)_2$, 5% glycerol, 20 mM imidazole, and 0.2 mM Tris (2-carboxyethyl) phosphine (TCEP).
7. His B buffer: 30 mM HEPES-KOH pH 7.4, 200 mM KOAc, 2 mM $Mg(OAc)_2$, 5% glycerol, 400 mM imidazole, and 0.2 mM TCEP.
8. Dilution buffer: 30 mM HEPES-KOH pH 7.4, 2 mM $Mg(OAc)_2$, 5% glycerol, and 1 mM dithiothreitol (DTT).
9. Mono Q A buffer: 30 mM HEPES-KOH pH 7.4, 20 mM KCl, 2 mM $Mg(OAc)_2$, 5% glycerol, and 1 mM DTT.
10. Mono Q B buffer: 30 mM HEPES-KOH pH 7.4, 1 M KCl, 2 mM $Mg(OAc)_2$, 5% glycerol, and 1 mM DTT.
11. Storage buffer: 30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM $Mg(OAc)_2$, 10% glycerol, and 1 mM DTT.
12. Hypotonic lysis buffer: 10 mM HEPES-KOH pH 7.4, 10 mM KCl, 1.5 mM $MgCl_2$, and 1× cOmplete EDTA-free protease inhibitor cocktail (Roche).
13. 5× Lysis buffer: 150 mM HEPES-KOH pH 7.4, 500 mM KOAc, and 10 mM $Mg(OAc)_2$.
14. Lysis buffer: 30 mM HEPES-KOH pH 7.4, 100 mM KOAc, and 2 mM $Mg(OAc)_2$.
15. Wash buffer: 30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM $Mg(OAc)_2$, 0.8 M NaCl, 1 mM DTT, and 1% Triton X-100.
16. Elution buffer 1: Lysis buffer containing 5 mM D-biotin.
17. Elution buffer 2: Lysis buffer containing 2.5 mM D-biotin, 30% glycerol, and 1 mM DTT.
18. 100 mM ATP pH ~7: dissolve ATP in RNase-free water, adjust its pH to ~7 by KOH, and store in −20 °C.
19. 500 mM Creatine monophosphate: dissolve the powder stock (Fluka) in RNase-free water right before use.
20. 2 U/μl Creatine kinase: dissolve the powder stock (Cal-Biochem) in Lysis buffer containing 50% glycerol and 1 mM DTT and store in −20 °C.
21. ATP reaction mix: mix the reagents with the following order right before use.
    (a) 5.8 μl RNase-free water.
    (b) 8 μl 5× Lysis buffer.
    (c) 4 μl 100 mM ATP pH ~7.

(d) 20 μl 500 mM Creatine monophosphate.

(e) 1.2 μl 2 U/μl Creatine kinase (Cal-Biochem).

(f) 1 μl 40 U/μl RNasin Plus (Promega).

22. Proteinase K reaction mix: 200 mM Tris–HCl, pH 7.5, 25 mM EDTA, pH 8.0, 300 mM NaCl, 2% w/v sodium dodecyl sulfate (SDS), 2 mg/ml Proteinase K, and 0.2 mg/ml glycogen.
23. Native loading dye: 15% w/v Ficoll-400, 0.01% w/v xylene cyanol, 0.01% w/v Bromophenol Blue, 0.5× Tris/Borate/EDTA (TBE), and 2 mM $MgCl_2$.
24. 12% native polyacrylamide gel: mix the reagents with the following order, load into 20 cm × 20 cm gel glasses, and wait until gel is formed.

    (a) 11.8 ml RNase-free water.

    (b) 6 ml 40(w/v)%-Acrylamide/Bis Mixed Solution (19:1).

    (c) 2 ml 5× Tris/Borate/EDTA (TBE).

    (d) 80 μl 500 mM $MgCl_2$.

    (e) 80 μl 20% Ammonium Persulfate (APS).

    (f) 40 μl Tetramethylethylenediamine (TEMED).

25. Plasmids [11, 20]

    (a) pCold-Hsp83.

    (b) pCold-Hsc70-4.

    (c) pCold-Hop.

    (d) pCold-Droj2.

    (e) pCold-p23.

    (f) pASW-Hop.

    (g) pAWH-Dicer-2.

    (h) pASW-R2D2.

    (i) pAFW-Ago2.

26. Small RNAs: Synthetic RNA oligos for an siRNA duplex with the following sequences (5′–3′):
    Guide strand (*let-7*): pUGAGGUAGUAGGUUGUAUAGU
    Passenger strand: pUAUACAACCUACUACCUCUCU
    "p" indicates the 5′ monophosphate.
27. Illustra G-25 MicroSpin columns (GE).
28. Gel dryer (Model 583, Biorad).
29. 3MM filter paper (Whatman).
30. 25 μm thin polyethylene terephthalate sheet.
31. X-tremeGENE HP (Roche).

32. PD-10 (GE Healthcare).
33. HisTrap FF crude 5 ml (GE Healthcare).
34. Mono Q 1 ml (GE Healthcare).
35. AKTA purifier (GE Healthcare) or equivalent FPLC.
36. GelCode blue stain reagent (Invitrogen).
37. Streptavidin Sepharose HP (GE Healthcare).
38. T4 polynucleotide kinase and 10× T4 Polynucleotide Kinase Buffer: 500 mM Tris–HCl pH 7.5, 100 mM $MgCl_2$, and 50 mM DTT (TAKARA).
39. FLA-7000 imaging system (Fujifilm Life Sciences).
40. [γ-^{32}P] ATP: 7000 Ci/mmol, 100 mCi/ml (MP Biomedical).

3 Methods

Fly Ago2-RISC assembly can be reconstituted with eight proteins: Ago2, Dicer2, R2D2, Hsp83, Hsc70-4, Hop, Droj2, and p23. For the purification of chaperone machinery proteins (Hsp83, Hsc70-4, Hop, Droj2, and p23), two different approaches are available: (a) co-purification of the core chaperone machinery complex (Hsp83/Hop/Hsc70-4) from *Drosophila* S2 cells (Subheading 3.3) and additional purifications of co-chaperones (Droj2 and p23) from *E. coli* (Subheading 3.1) and (b) individual purification of all the chaperone components from *E. coli* (Subheadings 3.1 and 3.2) (Fig. 1a–e). When the same molarities were used, the co-purified core chaperone machinery was more active than the combination of individually purified components (~0.15 nM vs. ~0.03 nM RISC programmed) [20]. Therefore, we used the co-purified core chaperone machinery for our single molecule analysis (*see* Chap. 10). Because R2D2 is known to be unstable in the absence of Dicer-2 [22, 23], we co-expressed streptavidin-binding peptide (SBP)-tagged R2D2 and non-tagged Dicer-2 in insect cells, and purified via the SBP tag, thus ensuring homogenous preparation of Dicer-2/R2D2 heterodimer (Subheading 3.4) (Fig. 1f). Although successful purification of eukaryotic Agos in solution has been reported [18, 19, 24–28], isolation of fly Ago2 competent for RISC assembly was quite difficult in our hands. Therefore, we take a strategy to keep Ago protein tethered on magnetic beads (or glass surface for single-molecule analysis) and directly monitor the RISC assembly reaction on beads by supplying Dcr-2/R2D2 and the chaperone components. For this, Ago2 should be immunopurified (Subheading 3.5) right before the RISC assembly reaction. The on-beads reaction allows us to wash out unloaded siRNAs and measure the amount of loaded double-stranded and single-stranded siRNAs (Subheading 3.7), which reflect the formation of pre-RISC and mature RISC, respectively (Fig. 2).

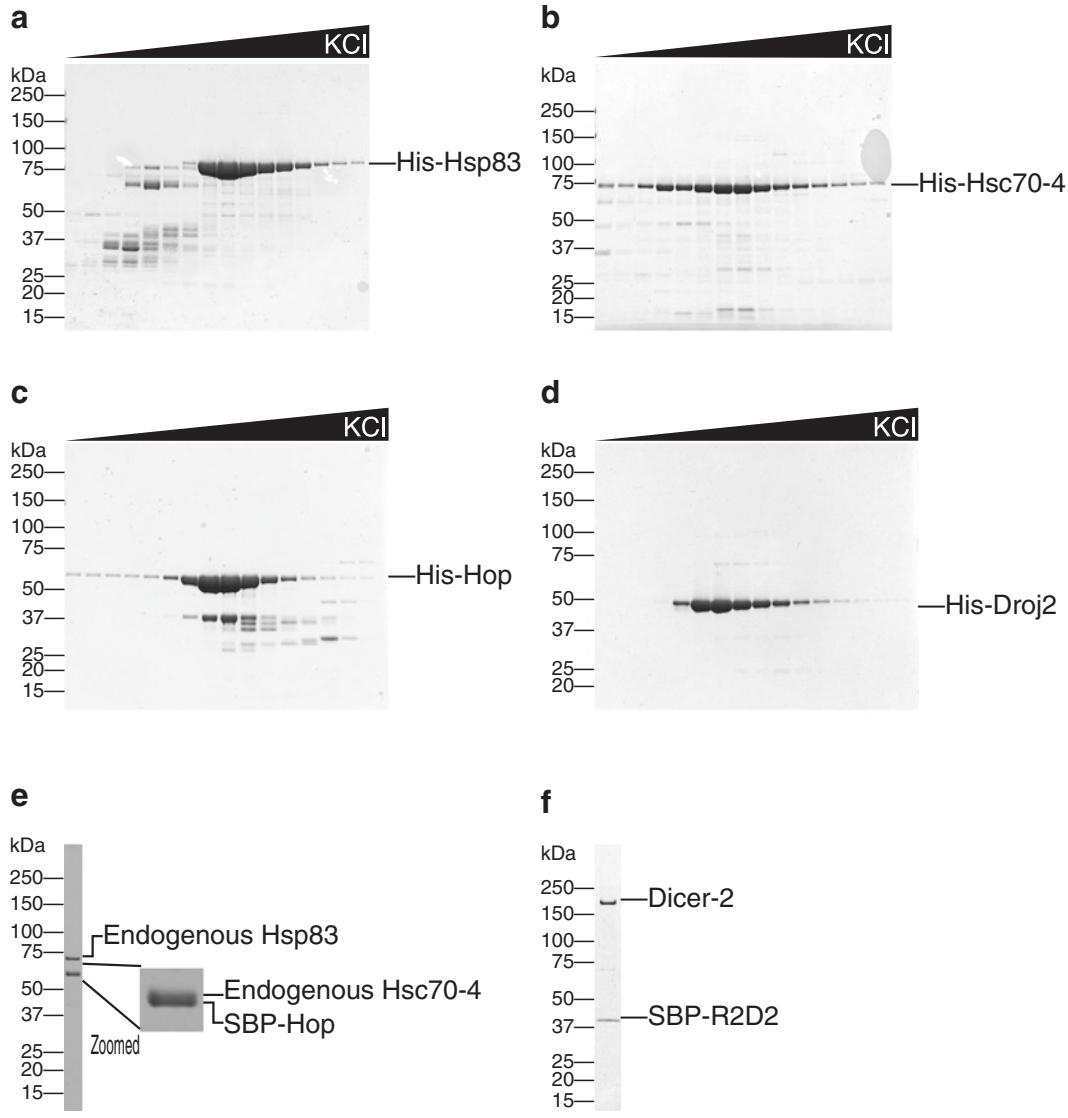


Fig. 1 Recombinant proteins used in this method. (**a**–**d**) Fractions of chaperone proteins expressed in *E. coli* by Mono Q column along salt concentration. (**e** and **f**) Core chaperone machinery and Dicer-2/SBP-R2D2 heterodimer purified from S2 cell. Proteins were stained by CBB

3.1 Recombinant His-Hsc70-4, Hop, Droj2, and p23 Protein Purifications from E. coli

1. Cultivate BL21 strain transformed with *pCold-Hsc70-4*, *Hop*, *Droj2*, and *p23* in 2 l of LB to an OD_{600} of 0.5 at 37 °C with 100 μg/ml ampicillin. Inoculate the cells with 1 mM IPTG and further cultivate them at 15 °C overnight.
2. Collect cell pellets by centrifugation at 3000 × *g* for 30 min at 4 °C.

   PAUSE POINT—Shock-freeze the cell pellet by liquid nitrogen and store at −80 °C.

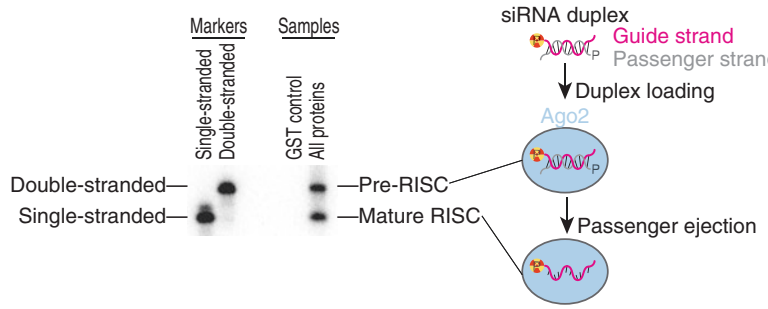


Fig. 2 Reconstitution of RISC assembly with purified factors Single-stranded siRNAs (mature RISC) and double-stranded siRNAs (pre-RISC) bound to FLAG-Ago2 after incubation of the chaperone protein and Dicer-2/SBP-R2D2 (reaction B) were run on a native acrylamide gel

3. Resuspend the cell pellets with 20 ml of His A buffer containing 1× cOmplete EDTA-free protease inhibitor cocktail, place in glass beaker, sonicate for 3 min on ice by Sonifier 250 with 60% power.
4. Centrifuge the lysate at 10,000 × *g* for 20 min at 4 °C. Hereafter, handle the samples at 4 °C during a whole procedure.
5. Load the supernatant onto HisTrap FF crude 5 ml and elute the recombinant proteins with a liner gradient from His A buffer to His B buffer using AKTA purifier (or equivalent FPLC). Monitor the abundance of the recombinant proteins by SDS-PAGE and CBB staining [GelCode blue stain reagent or an equivalent method], and then collect the peak fractions.
6. Dilute the fractions ten times with Dilution buffer and loaded onto Mono Q 1 ml. Elute the recombinant proteins with a linear gradient from Mono Q A buffer to Mono Q B buffer using AKTA purifier (or equivalent FPLC). Monitor the abundance of the recombinant proteins by SDS-PAGE and CBB staining (Fig. 1a–d), and then collect the peak fractions (*see* **Note 1**).
7. Buffer-exchange the fractions to Storage buffer by PD-10. Measure the concentration of recombinant proteins with OD_{280} and calculate their molarities based on the extinction co-efficiencies from the amino acid sequences by ExPASy ProtParam tool (http://web.expasy.org/protparam/) or equivalent ones.
8. Split the proteins into aliquots if necessary. Shock-freeze the proteins by liquid nitrogen and stored at −80 °C

3.2 Recombinant His-Hsp83 Protein Purification from E. coli

1. Follow the procedure described above (Subheading 3.1) by starting with BL21 transformed with *pCold-Hsp83*. Omit glycerol and the reducing reagent from the buffers during all the steps of purification (*see* **Note 2**).

3.3 Purification of Hsp83/SBP-Hop/Hsp70-4 Complex from S2 Cells

1. Transfect the complex of 10 μg *pASW-Hop* and 20 μl X-tremeGENE HP into 10 ml of 1.0×10^6 cells/ml *Drosophila* S2 cells placed in a 10 cm dish, following the manufacture's instruction. Prepare ten dishes of the transfected cells. Cultivate the cells for 72 h.
2. Collect the cells by centrifugation with 3000 × *g* for 5 min at 4 °C. Wash the cells pellet with PBS once. Measure the weight of the cell pellet. Hereafter, handle the samples at 4 °C during the whole procedure.
3. Resuspend the cells with Hypotonic lysis buffer of twice the volume of the cell pellet, incubate for 30 min on ice, and centrifuge with 17,000 × *g* for 20 min at 4 °C. Collect the supernatant.

   PAUSE POINT—Shock-freeze the lysate in liquid nitrogen and store at −80 °C.
4. Equilibrate 200 μl Streptavidin Sepharose HP. Spin down the beads and discard supernatant. Add 1 ml of Hypotonic lysis buffer. Repeat this wash step two times (three times wash in total).
5. Incubate 2 ml of the S2 cell lysate with the equilibrated Streptavidin beads for 1 h at 4 °C with nutation.
6. Spin down the beads and discard supernatant. Add 1 ml of Lysis buffer. Repeat this wash step four times (five times wash in total). Discard the final supernatant.
7. Add 100 μl of elution buffer 1 and incubate for 30 min at 4 °C, mixing occasionally.
8. Collect the supernatant and check the protein purity by SDS-PAGE and CBB staining (Fig. 1e) (*see* **Note 3**).
9. Measure the protein concentration with OD_{280} and calculate the molarity based on extinction co-efficiencies of the proteins predicted by ExPASy ProtParam tool, assuming the complex of Hsp83/SBP-Hop/Hsc70-4 as 2:1:1 complex.
10. Split the proteins into aliquots if necessary. Shock-freeze the proteins in liquid nitrogen and store at −80 °C.

3.4 Purification of Dicer2/SBP-R2D2 Complex from S2 Cells

1. Transfect the complex of 7.5 μg *pAWH-Dicer-2*, 7.5 μg *pASW-R2D2*, and 30 μl X-tremeGENE HP per 10 cm dish, following **step 1** of Subheading 3.3.
2. Follow the **steps 2–5** of Subheading 3.3. Use Hypotonic lysis buffer containing 1 mM DTT.
3. Spin down the beads and discard the supernatant. Add 1 ml of Lysis buffer containing 0.8 M NaCl and 1 mM DTT. Repeat this wash step five times in total. Further wash twice with lysis buffer containing 30% glycerol and 1 mM DTT. Discard the final supernatant.

4. Add the 100 µl of elution buffer 2 and incubate for 30 min at 4 °C, mixing occasionally.
5. Follow the **steps 8–10** of Subheading 3.3. Estimate the protein molarity with the assumption 1:1 Dicer-2/SBP-R2D2 complex (Fig. 1f).

3.5 Preparation of FLAG-Ago2 Expressing S2 Cell Lysate

1. Transfect the complex of 10 µg *pAFW-Ago2* and 20 µl X-tremeGENE HP per 10 cm dish, following **step 1** of Subheading 3.3.
2. Follow the **step 2** of Subheading 3.3.
3. Resuspend the cells with the same volume of Lysis buffer containing 1 mM DTT, lyse the cells with Dounce homogenizer by 20 strokes, and centrifuge at 17,000 × *g* for 20 min at 4 °C. Collect the supernatant, split into aliquots, shock-freeze in liquid nitrogen, and stored at −80 °C.

3.6 Preparation of Guide Strand-Labeled siRNA Duplexes

1. Combine the following reaction and incubate at 37 °C for 1 h (*see* **Note 4**).
   (a) 15.1 µl RNase-free water
   (b) 2 µl 10× T4 Polynucleotide Kinase Buffer
   (c) 0.4 µl 14.3 µM [γ-^{32}P] ATP (7000 Ci/mmol, 100 mCi/ml, MP Biomedical)
   (d) 1 µl 10 µM guide strand RNA
   (e) 1 µl T4 polynucleotide kinase
   (f) 0.5 µl 40 U/µl RNasin Plus
2. Load the reaction on G-25 spin column following the manufacturer's instruction and take the flow though.
3. Precipitate the RNA with 1/10 volume of 3 M NaOAc, pH 5.2, 1 µl of 20 mg/ml Glycogen, and 2.5× volume of ethanol. Centrifuge the mix at 17,000 × *g* for 15 min at 4 °C.
4. Discard the supernatant and rinse the pellet with 200 µl of 70% ethanol. Centrifuge at 17,000 × *g* for 15 min at 4 °C.
5. Discard the supernatant and dry-up the RNA pellet.
6. Resuspend the RNA with 10 µl of RNase-free water and assume the concentration as 1 µM.
7. Prepare the following reagents, incubate at 95 °C for 5 min, and then cool it at room temperature for 30 min.
   (a) 15 µl RNase-free water
   (b) 10 µl 5× Lysis buffer
   (c) 10 µl 1 µM radio-labeled guide strand
   (d) 15 µl 1 µM passenger strand RNA

8. Store the siRNA duplex at −20 °C. Use this stock as 200 nM siRNA duplex. The duplex can be used for approximately a month, as long as signals can be detected by the imager.

3.7 Immunopurification of Ago2 and the Following RISC Assembly Reaction

1. For 15 reactions, equilibrate 30 μl of Dynabeads Protein G by three time washing with Lysis buffer containing 1 mM DTT, using a magnetic stand. After the final wash, resuspend the beads with 30 μl of Lysis buffer containing 1 mM DTT. Hereafter, handle the samples at 4 °C during the whole procedure, until **step 5** of Subheading 3.7.
2. Add 1 μl of 1 mg/ml anti-Flag M2 antibody and incubate at 4 °C for 30 min with occasional mixing.
3. Wash the beads three times with lysis buffer containing 1 mM DTT, using magnetic stand. Discard the final supernatant.
4. Thaw the FLAG-Ago2 S2 cell lysate on ice, and then incubate 30 μl of the lysate with the antibody-conjugated beads at 4 °C for 1 h with occasional mixing.
5. Wash the beads five times with Wash buffer and rinse twice with Lysis buffer containing 1 mM DTT. Split the beads into 15 tubes.
6. Add the following reaction A or B to the beads right after discarding the final supernatant. Incubate the reaction at 25 °C for 90 min (*see* **Note 5**).

   Reaction A:

   (a) 1 μl ATP reaction mix
   (b) 2 μl 1× Lysis buffer
   (c) 1 μl 22 μM SBP-Hop/Hsp83/Hsc70-4 complex
   (d) 2 μl 42 μM His-Droj2
   (e) 2 μl 11 μM His-p23
   (f) 1 μl 200 nM Dicer-2/SBP-R2D2
   (g) 1 μl 200 nM guide strand labeled siRNA duplex

   Reaction B:

   (a) 1 μl ATP reaction mix
   (b) 0.5 μl 88 μM His-Hsp83
   (c) 3 μl 7.4 μM His-Hsc70-4
   (d) 0.5 μl 44 μM His-Hop
   (e) 2 μl 42 μM His-Droj2
   (f) 1 μl 22 μM His-p23.
   (g) 1 μl 200 nM Dicer-2/SBP-R2D2
   (h) 1 μl 200 nM guide strand labeled siRNA duplex

7. Wash the beads four times with Wash buffer, using a magnetic stand (*see* **Note 6**).
8. Add 100 μl of Proteinase K reaction mix and incubate at 25 °C for 90 min.
9. Add 250 μl of ethanol and centrifuge at 17,000 × *g* for 15 min at 4 °C.
10. Discard the supernatant and dry up the pellet.
11. Resuspend the RNA with 5 μl of native loading dye.
12. Prepare the markers of double-stranded and single-stranded siRNAs by dilution of the labeled siRNA duplex and the guide strand by native loading dye, respectively (2 nM marker would be comparable with samples).
13. Load 2.5 μl of the sample onto 12% native polyacrylamide gel and run the gel in 0.5× TBE containing 2 mM $MgCl_2$ for 1 h at 4 °C.
14. Place the gel on filter paper and cover it with the polyethylene terephthalate sheet on the gel.
15. Dry the gel with gel dryer at 80 °C for 30 min.
16. Acquire the image by a FLA-7000 imaging system (Fig. 2) (*see* **Notes 7** and **8**).

4 Notes

1. Hsc70-4 could distribute widely along fractions during Mono Q purification step, so that final concentration may be lower than other chaperone proteins.
2. The addition of reducing reagent during protein purification may inactivate Hsp83 for the RISC assembly reaction. The importance in cysteine residues in Hsp90 for its function [29] and an inhibitor targeting the cysteines [30] suggested the possibility that reduction of cysteine may abrogate its function in RISC assembly.
3. The migrations of endogenous Hsc70-4 and SBP-Hop are quite similar on SDS-PAGE (Fig. 1e).
4. The guide strand is labeled by phosphate exchanging reaction between γ-position phosphate in ATP and 5′ monophosphate of guide strand by T4 polynucleotide kinase.
5. Use equimolar concentration of Glutathione S-transferase (GST) in the reconstitution system as a negative control.
6. Do not keep the RISC assembly reaction and purified RNAs higher than the room temperature (25 °C). Double-stranded RNA may be unwound during the handling and may cause the

overestimation of the actual efficiency of passenger strand ejection during RISC assembly.

7. The upper band and lower band represent pre-RISC containing double-stranded siRNA and mature RISC containing single-stranded siRNA, respectively (Fig. 2).
8. If the passenger strand instead of the guide strand is radiolabeled in siRNA duplex, only pre-RISC containing double-stranded siRNA would be observed.

Acknowledgments

We are grateful to members of Tomari laboratory and Iwasaki laboratory for critical reading of this manuscript. This work is supported by in part by Grants-in-Aid for Scientific Research on Innovative Areas ('Functional machinery for non-coding RNAs' 21115002 and 'Non-coding RNA neo-taxonomy' 26113007) (to Y.T.), and a Grant-in-Aid for Scientific Research on Innovative Areas ('nascent chain biology' JP17H05679) and Grant-in-Aid for Young Scientists ('Start-up' 23870004 and 'A' JP17H04998) (to S.I.) from The Ministry of Education, Culture, Sports, Science and Technology in Japan.

References

1. Hock J, Weinmann L, Ender C et al (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. EMBO Rep 8:1052–1060
2. Landthaler M, Gaidatzis D, Rothballer A et al (2008) Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA 14:2580–2596
3. Meister G, Landthaler M, Peters L et al (2005) Identification of novel Argonaute-associated proteins. Curr Biol 15:2149–2155
4. Tuschl T, Zamore PD, Lehmann R, Bartel DP, Sharp PA (1999) Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev 13:3191–3197
5. Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110:563–574
6. Iki T, Yoshikawa M, Nishikiori M et al (2010) In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. Mol Cell 39:282–291
7. Nykanen A, Haley B, Zamore PD (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell 107:309–321
8. Kawamata T, Seitz H, Tomari Y (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol 16:953–960
9. Yoda M, Kawamata T, Paroo Z et al (2010) ATP-dependent human RISC assembly pathways. Nat Struct Mol Biol 17:17–23
10. Miyoshi T, Takeuchi A, Siomi H, Siomi MC (2010) A direct role for Hsp90 in pre-RISC formation in Drosophila. Nat Struct Mol Biol 17:1024–1026
11. Iwasaki S, Kobayashi M, Yoda M et al (2010) Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. Mol Cell 39:292–299
12. Johnston M, Geoffroy MC, Sobala A, Hay R, Hutvagner G (2010) HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. Mol Biol Cell 21:1462–1469

13. Kwak PB, Tomari Y (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. Nat Struct Mol Biol 19:145–151
14. Rand TA, Petersen S, Du F, Wang X (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell 123: 621–629
15. Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell 123:607–620
16. Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC (2005) Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. Genes Dev 19:2837–2848
17. Leuschner PJ, Ameres SL, Kueng S, Martinez J (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Rep 7:314–320
18. Liu Y, Ye X, Jiang F et al (2009) C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. Science 325:750–753
19. Ye X, Huang N, Liu Y et al (2011) Structure of C3PO and mechanism of human RISC activation. Nat Struct Mol Biol 18:650–657
20. Iwasaki S, Sasaki HM, Sakaguchi Y, Suzuki T, Tadakuma H, Tomari Y (2015) Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature 521:533–536
21. Smith DF, Toft DO (2008) Minireview: the intersection of steroid receptors with molecular chaperones: observations and questions. Mol Endocrinol 22:2229–2240
22. Liu Q, Rand TA, Kalidas S et al (2003) R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. Science 301:1921–1925
23. Liu X, Jiang F, Kalidas S, Smith D, Liu Q (2006) Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. RNA 12:1514–1520
24. Schirle NT, MacRae IJ (2012) The crystal structure of human Argonaute2. Science 336:1037–1040
25. Elkayam E, Kuhn CD, Tocilj A et al (2012) The structure of human Argonaute-2 in complex with miR-20a. Cell 150:100–110
26. Nakanishi K, Weinberg DE, Bartel DP, Patel DJ (2012) Structure of yeast Argonaute with guide RNA. Nature 486:368–374
27. Nakanishi K, Ascano M, Gogakos T et al (2013) Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. Cell Rep 3:1893–1900
28. Faehnle CR, Elkayam E, Haase AD, Hannon GJ, Joshua-Tor L (2013) The making of a slicer: activation of human Argonaute-1. Cell Rep 3:1901–1909
29. Mollapour M, Neckers L (2012) Post-translational modifications of Hsp90 and their contributions to chaperone regulation. Biochim Biophys Acta 1823:648–655
30. Zhang Y, Dayalan Naidu S, Samarasinghe K et al (2014) Sulphoxythiocarbamates modify cysteine residues in HSP90 causing degradation of client proteins and inhibition of cancer cell proliferation. Br J Cancer 110:71–82

Chapter 10

Single-Molecule Analysis for RISC Assembly and Target Cleavage

Hiroshi M. Sasaki, Hisashi Tadakuma, and Yukihide Tomari

Abstract

RNA-induced silencing complex (RISC) is a small RNA–protein complex that mediates silencing of complementary target RNAs. Biochemistry has been successfully used to characterize the molecular mechanism of RISC assembly and function for nearly two decades. However, further dissection of intermediate states during the reactions has been warranted to fill in the gaps in our understanding of RNA silencing mechanisms. Single-molecule analysis with total internal reflection fluorescence (TIRF) microscopy is a powerful imaging-based approach to interrogate complex formation and dynamics at the individual molecule level with high sensitivity. Combining this technique with our recently established in vitro reconstitution system of fly Ago2-RISC, we have developed a single-molecule observation system for RISC assembly. In this chapter, we summarize the detailed protocol for single-molecule analysis of chaperone-assisted assembly of fly Ago2-RISC as well as its target cleavage reaction.

Key words Single-molecule imaging, RNA interference, Argonaute, RISC, Small interfering RNA, Total internal reflection fluorescence microscopy

1 Introduction

Small RNAs such as small interfering RNAs (siRNAs) and microRNAs (miRNAs) induce post-transcriptional gene regulation or RNA silencing, by which gene expression is repressed through mRNA degradation and/or translation repression. At the heart of the molecular basis of RNA silencing, RNA-induced silencing complexes (RISCs) work as the effector complex that carries out catalytic cleavage of target mRNA, inhibition of translation initiation, and/or recruitment of downstream silencing factors like the deadenylation complexes [1, 2]. RISCs are composed of two components: (1) single-stranded small RNAs, typically at the length of ~22 nucleotides, which work as the guide to target mRNAs using their sequence complementarity and (2) Argonaute family proteins, which share a characteristic four-domain organization.

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_10,

A decade of research, particularly biochemistry of RNA interference (RNAi) in *Drosophila*, has largely revealed the molecular mechanism of RISC assembly [3–6]. In particular, RISC assembly can be divided into two steps: duplex loading and passenger ejection. For example, in the *Drosophila* RNAi pathway, a small RNA duplex, which is processed from a long double-stranded RNA precursor by a heterodimer of RNase III-family endonuclease Dicer-2 and its partner protein R2D2, is handed over from a Dicer-2/R2D2 heterodimer to Argonaute 2 (Ago2) protein, to produce precursor RISC (pre-Ago2-RISC). Next, at the passenger ejection step, Ago2 cleaves one of the strands called the passenger strand through its catalytic activity and releases the cleaved products to form a mature Ago2-RISC, which contains only a single-stranded RNA called the guide RNA. It is also known that, while duplex loading is promoted by ATP, passenger ejection occurs spontaneously and does not require ATP hydrolysis [4, 6–10]. Biochemical studies in flies, mice, and plants have indicated that the Hsp90/Hsp70 chaperone machinery is required to load small RNA duplexes into Ago proteins in an ATP-dependent manner, but not for the passenger ejection step [7, 8, 11].

Biochemical approaches such as phosphorimaging with native agarose gel electrophoresis of RNA–protein complexes or native polyacrylamide gel electrophoresis of Ago-bound RNAs combined with lysate-based in vitro RISC assembly system have been used to investigate the basic process of RISC biogenesis [9, 10, 12–14]. As for *Drosophila* Ago1 pathway or human Ago2 pathway, for instance, small RNA duplexes radiolabeled with ^{32}P at the 5′ end can readily form pre-RISCs and mature RISCs in cell lysate at 25 °C, whereas incubation at 15 °C allows only duplex loading and prevents the conversion from pre-RISCs into mature RISCs [9, 14]. Native agarose gel electrophoresis can separate pre-RISCs and mature RISCs by mobility, therefore a pulse-chase experiment of the RISC assembly reaction at 15 or 25 °C with native agarose gel electrophoresis can monitor the kinetics of complex formation. However, as these assays are performed on ensemble of molecules and their temporal resolution is at the order of minutes, they provide only limited information about individual RISCs and cannot further dissect the intermediate states of RISC assembly. Moreover, in *Drosophila* Ago2 pathway, it remains unclear how Dicer-2/R2D2 heterodimer, which is essential for Ago2-RISC assembly, hands over a small RNA duplex to Ago2. Thus, many gaps remain in our understanding of RISC biogenesis.

Single-molecule analysis with total internal reflection fluorescence (TIRF) microscopy is a powerful imaging-based approach to interrogate complex formation of individual molecules with high sensitivity. Moreover, monitoring biochemical process in real time, single-molecule analysis can detect short-lived events/intermediates and reveal the dynamics of the complex formation. Combining

this technique with our recently established in vitro reconstitution system of fly Ago2-RISC assembly using seven recombinant proteins, Dicer-2, R2D2, Hsp90, Hsc70, Hop, p23, and Droj2, we have developed a single-molecule observation system for RISC assembly [15]. This system can investigate the fundamental steps in duplex loading as well as passenger ejection with a temporal resolution of 1 s and reveal the population structure of the intermediate states in RISC assembly. Similar single molecule technique is also applicable to the target cleavage reaction by RISC [16–18]. Our recent single-molecule analysis of *Drosophila* Ago2-RISC [16] unveiled how RISC accurately recognize and cleave complementary target RNAs and release the cleaved fragments. In this chapter, we summarize the detailed protocol for single-molecule analysis of chaperone-assisted assembly of fly Ago2-RISC as well as the target cleavage reaction.

2 Materials

2.1 General Buffers, Reagents, and Cells

1. 5× lysis buffer: 150 mM HEPES-KOH (pH 7.5), 500 mM KOAc, 10 mM $Mg(OAc)_2$.
2. 1× lysis buffer: 30 mM HEPES-KOH (pH 7.5), 100 mM KOAc, 2 mM $Mg(OAc)_2$.
3. 1 M dithiothreitol (DTT).

2.2 TIRF Microscope Setup

We constructed a prism-type custom TIRF microscope using an inverted microscope, lasers, a fluorescence image splitting system, and a back-illuminated electron-multiplying charge-coupled device camera (Fig. 1) [19]. Here is the list of the major components of our custom microscope.

1. Inverted type microscope IX71 (Olympus).
2. Oil immersion objective lens, UAPON 150× OTIRFM, NA 1.45 (Olympus). (Other objectives such as 60× or 100× are also usable. The choice of objective magnification depends on the camera sensor size, pixel size and the density of the signal and is important for precise 2D or 3D trajectory analysis [20, 21]).
3. Ultra-stable manual XY stage for Olympus inverted type microscope (Chukousha).
4. Anti-vibration table, 1500 × 1000 mm (Heltz).
5. Ar laser, 514 nm (Melles Griot): 514–532 nm diode-pumped solid-state (DPSS) laser is also usable.
6. He-Ne laser, 633 nm (SOC): 635–640 nm laser is also usable.
7. Back-illuminated electron multiplying charge coupled device (EMCCD) camera, iXon3 DU-897E-CSO-#BV, 512 × 512

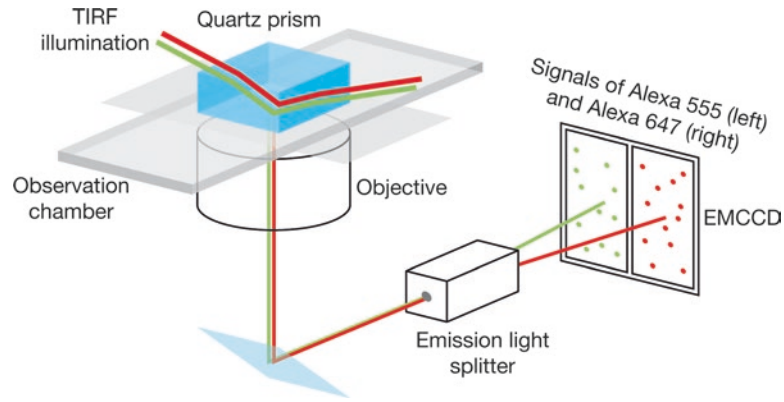


Fig. 1 Schematic of the single-molecule observation setup. A prism-type custom TIRF microscope is used to selectively illuminate fluorescently labeled molecules on the surface. An observation chamber is set onto the objective lens with the coverslip-side down. Halo-Ago2 molecules are tethered on the slide glass surface (upper side of the chamber). The gap between the quartz prism and slide glass is filled with glycerol, allowing the incident light to enter into the slide glass at the designed angle. The sample is excited by both 514 nm and 633 nm lasers simultaneously at various power densities depending on the experiments (*see* the main text). For detection, an emission light splitter separates the emission lights of the *green* (e.g., Alexa 555) and *red* (e.g., Alexa 647) dyes from each other and projects them to the single EMCCD chip side by side

pixels (Andor Technology): A camera with another format such as 1024 × 1024 is also usable. With enough photons, scientific complementary metal-oxide-semiconductor (sCMOS) cameras could be used instead of EMCCD cameras. For a camera with a larger image sensor, the dimension of image splitter's field of view should be checked.

8. EMCCD acquisition software, SOLIS version 4.23.30003.0 (Andor Technology).
9. Fluorescence image splitter, DualView2 (Photometrics) with a 635 nm filter cube (Dichroic mirror, T635LPXR (Chroma)): In the detection pathway, the DualView2 module separates spatially identical but spectrally distinct emission lights and projects onto the camera chip side by side.
10. Emission filter, HQ582/50 m (Chroma).
11. Emission filter, HQ700/100 m (Chroma).
12. Aluminum mirror, TFA-30C05-4 (OptoSigma).
13. Dichroic mirror, FF560-Di01-25 × 36 (Semrock): To combine green and red laser.
14. Synthetic quartz prism, 20 × 20 × 6 mm, OPSQ-20S06-4P-3 (OptoSigma).

15. Quarter waveplate, 514 nm, WPQ-5145-04 M (OptoSigma).
16. Quarter waveplate, 633 nm, WPQ-6328-04 M (OptoSigma).
17. Variable ND filter.
18. Custom aluminum holders and bases to build up the optical system.

2.3 Ago2 Protein Labeling with HaloTag Ligands

1. 1 μg/μl pAHaloFW-Ago2 wild type, D965A, or Y897E: Prepared by Qiagen midi-prep kit.
2. 1 μg/μl pAHaloFW-TEV-Ago2 wild type, D965A.
3. 1 μg/μl pAHaloFW-3C-Ago2 wild type, D965A.
4. X-treamGENE HP DNA transfection reagent (Roche).
5. *Drosophila* S2 cells.
6. *Drosophila* Schneider's medium (Thermo Fisher).
7. Fetal bovine serum (FBS) (Thermo Fisher).
8. Dounce homogenizer, 7 ml, "TIGHT" pestle (Wheaton).
9. 1× PBS.
10. Protease inhibitor cocktail (Roche).
11. HaloTag TMR ligand (Promega): For quantification of Halo-Ago2 proteins in lysate. Dissolve in DMSO to make 100 μM of the stock solution and store at −30 °C.
12. HaloTag PEG-biotin ligand (Promega): For surface tethering of Halo-Ago2 on the single-molecule observation chamber. Dissolve in DMSO to make 100 μM of the stock solution and store at −30 °C.
13. NAP-5 column (GE Healthcare).

2.4 Preparation of Fluorescently Labeled siRNA Duplexes

2.4.1 RNA Oligos

1. *let-7* siRNA guide strand, 5′-phosphorylated, 3′-Alexa 647-labeled (Gene design).
2. *let-7* siRNA guide strand, 5′-hydroxy, 3′-Alexa 647-labeled (Gene design).
3. *let-7* siRNA guide strand, 5′-phosphorylated, 3′-Alexa 555-labeled (Gene design).
4. *let-7* siRNA passenger strand, 5′-phosphorylated, 3′-Alexa 555-labeled (Gene design).
5. *let-7* siRNA passenger strand, 5′-phosphorylated, 3′-Alexa 647-labeled (Gene design).

2.4.2 PAGE Purification

1. 40% acrylamide gel mix.
2. 10× TBE.
3. 200 mM Ammonium persulfate (APS).
4. Tetramethylethylenediamine (TEMED).

5. Power unit.
6. Electrophoresis gel chamber (Biocraft).
7. Ice-filled plastic bar: ϕ20 × 400 mm plastic tube half-filled with water. Frozen at −20 °C lying on its long side.
8. 2× Native-PAGE loading dye: 25% glycerol, 2 mM $MgCl_2$, 0.01% bromo-phenol blue, 0.01% xylene cyanol, 0.02% tartrazine, and 0.5× TBE.
9. LAS-3000 (Fujifilm).
10. 2× PK buffer: 100 mM Tris–HCl (pH 7.5), 200 mM NaCl, 2 mM EDTA, 1% SDS.
11. Ethanol.
12. Glycogen.

2.5 Preparation of the Observation Chamber [22]

2.5.1 Materials and Reagents

1. Synthetic quartz glass slides, 26 × 56 × 1 mm (Matsunami Glass).
2. Coverslips, 24 × 36 mm (Matsunami Glass).
3. Methanol.
4. Ethanol.
5. Acetic acid.
6. 0.1 M KOH.
7. Aminosilane LS-4265, 100 g (Shin-Etsu Silicone): Prepare and store the stocks as follows:
   (a) Aliquot aminosilane 6.4 ml per bottle, put a silicone rubber stopper on the bottle, and crimp an aluminum cap.
   (b) Insert a 27G needle through the rubber cap for a vent. Next, insert another 27G needle and inject argon or nitrogen gas into the bottle at 0.1 MPa for 20 s.
   (c) Remove the vent needle first and inject more argon or nitrogen gas for 10 s.
   (d) Remove the gas injection needle.
   (e) The stocks can be kept at 4 °C.
8. PEG-biotin, Biotin-CONH-PEG-O-C3H6-CONHS, 500 mg (Rapp Polymere).
9. SUNBRIGHT PEG5000, 1 g (NOF): Aliquot 43 mg of the powder per tube and then cover the top with Parafilm. Store the stocks at −30 °C.
10. 50 mM MOPS-KOH (pH 7.5): Prepare and store the stocks as follows:
    (a) Mix 5.23 g of MOPS and 43 ml of MilliQ water and then add 3 ml of 5 M KOH. You may store this 0.5 M MOPS-KOH stock at −30 °C protected from light. When the

solution color change to light yellow over time, make a fresh stock.

(b) Dilute the 0.5 M stock tenfold with MilliQ water to make 50 mM stock.

(c) Make 300 μl aliquots and store them at −30 °C.

11. Silica gel desiccants.

2.5.2 Equipment

1. Low-alkali glass bottle with rubber cap, Amber, 10 ml (NEG): For aminosilane storage.
2. Aluminum seal crimp cap (NEG): For aminosilane storage.
3. Hand crimper for aluminum seals, 20 mm (TGK): For aminosilane storage.
4. Argon (or nitrogen) gas: For aminosilane storage.
5. Slide rack for 20 slides (Matsunami Glass).
6. Stainless steel deep vat, 84 × 140 × 67 mm (Sansho): For washing quartz slides and coverslips.
7. Stainless steel deep vat, 96 × 160 × 73 mm (Sansho): For aminosilane coupling reaction.
8. Glass staining dish for 15 slides (AS ONE): For washing quartz slides.
9. Glass jar, ϕ60 × 90 mm (Pyrex): For washing coverslips.
10. Coverslip rack for ten slips (Ikemoto Scientific Technology).
11. Magnetic stirrer.
12. Teflon star head magnetic stir bar, diameter 22 mm, height 15 mm (AS ONE).
13. Sonicator water bath.
14. Parafilm: cut into 16 × 25 mm pieces before use.
15. Vacuum sealer: Compatible with both plastic bags and vacuum canisters.
16. Vacuum seal canister.

2.6 Single-Molecule Imaging

1. 3MM Filter paper, 10 × 100 mm (Whatmann).
2. 2.5 mg/ml NeutrAvidin solution (Invitrogen): Dissolve 5 mg of NeutrAvidin in 2 ml of MilliQ water. Make 100 μl of aliquots and flash-freeze with liquid nitrogen. Store at −30 °C.
3. Double-sided tape: 24 × 36 mm, window size 16 × 24 mm, thickness 25 μm (3 M).
4. PCR cooler (Eppendorf).
5. Objective oil.
6. Glycerol.

2.7 In Vitro Reconstructed System of RISC Assembly for Single-Molecule Imaging

1. 20 mM ATP (Sigma).
2. Creatine monophosphate (Sigma): Prepare a fresh stock of 500 mM creatine monophosphate in MilliQ water on the day of use.
3. Creatine phosphokinase, 10 U/μl (Calbiochem).
4. RNasin Plus (Promega).
5. ATP regenerating system: Mix 5.8 μl of MilliQ water, 8 μl of 5× lysis buffer, 4 μl of 20 mM ATP, 20 μl of 500 mM creatine monophosphate, 1.2 μl of 10 U/μl creatine kinase, and 1 μl of RNasin Plus.
6. 200 nM siRNA duplex stock.
7. 5% Biolipidure-203 (NOF).
8. 1 mg/ml Yeast tRNA mix (Sigma).
9. Glucose oxidase, 50,000 U (Sigma): Dissolve in MilliQ water to make 5000 U/ml of the stock solution. Make 10 μl of aliquots and flash-freeze with liquid nitrogen. Store at −80 °C. Dilute 1 μl of the stock solution into 9 μl of MilliQ water before use.
10. Catalase: Dissolve 100 mg (Sigma) of catalase in MilliQ water to make 5000 U/ml of the stock solution. Make 10 μl of aliquots and flash-freeze with liquid nitrogen. Store at −80 °C.
11. Glucose: Dissolve 4.5 g of glucose in 10 ml of MilliQ water to make 450 mg/ml of stock solution. Make 10 μl of aliquots and flash-freeze with liquid nitrogen. Store at −80 °C.
12. Oxygen scavenging system: Mix 3.4 μl of MilliQ water, 1.6 μl of 5× lysis buffer, 1 μl of glucose, 1 μl of glucose oxidase, and 1 μl of catalase at each time of use.
13. (±)-6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox), (Sigma): Dissolve 10 mg of Trolox in 200 μl of ethanol to make 200 mM of the stock solution. Store at 4 °C protected from light for up to 2 weeks.

2.8 Software

1. Image J.
2. OpenSIFkai plugin for Image J: Modified to open the SIF files taken by the latest version of Andor SOLIS. https://imagej.nih.gov/ij/plugins/open-sif.html.
3. Template Matching and Slice Alignment plugin for Image J [23].
4. A homemade program to obtain time trajectories of the fluorescence intensities of Alexa 555 and 647, as described previously [24].
5. Excel (Microsoft).
6. KaleidaGraph (Synergy).
7. Igor Pro (Wavemetrics).

2.9 Target Cleavage Reaction

Basically, the materials required for the target cleavage reaction are similar to those for the RISC assembly analysis, as briefly summarized below.

1. General buffers and reagents are the same as in Subheading 2.1.
2. TIRF microscope setup is similar as in Subheading 2.2 except that the Ar laser is switched to DPSS laser, 515 nm (cobolt).
3. For lysate preparation, wild-type *Drosophila* S2 cells are used.
4. Guide/passenger/target sequences are as follows:
   (a) *let-7* siRNA guide strand, 5′-phosphorylated, 3′-Dye-labeled, Dye = Cy3 or Cy5 (IDT).
   (b) *let-7* siRNA passenger strand, 5′-phosphorylated, non-labeled (IDT or Japan Bio Services).
   (c) Target strand, 5′-biotin-labeled, 3′-Dye-labeled, Dye = Alexa 555 or ATTO 647 N (Japan Bio Services).
   (d) Target strand, 5′-Dye-labeled, 3′-biotin-labeled, Dye = Alexa 555 or ATTO 647 N (Japan Bio Services).
   (e) 2′-*O*-Methylated target strand, 5′-biotin-labeled, 3′-Dye-labeled, Dye = Alexa 555 or ATTO 647 N, for fluorophore lifetime measurement (Japan Bio Services).
   (f) 2′-*O*-Methylated target strand, 5′-Dye-labeled, 3′-biotin-labeled, Dye = Alexa 555 or ATTO 647 N, for fluorophore lifetime measurement (Japan Bio Services).
5. For chamber preparation, same materials are used.
6. For single-molecule imaging, similar materials are used, but proteases are omitted.
7. For cleavage reaction observation, similar materials are used, but Biolipidure-203 and yeast tRNA mix are omitted.
8. For software, similar materials are used.

3 Methods

3.1 Ago2 Protein Labeling with Halo-Tag Ligands

3.1.1 Halo-Ago2 Expression

1. Centrifuge approximately 2×10^8 S2 cells in 50-ml tube at $1500 \times g$ for 5 min at room temperature.
2. Aspirate the supernatant, resuspend the cell pellet in 200 ml of antibiotics-free medium at 1×10^6 cells/ml, and transferred into 100-mm dishes (10 ml per each).
3. Mix 200 μl of 1 μg/μl Ago2 plasmid vector and 20 ml of serum-free antibiotics-free medium. Then mix 400 μl of X-treamGENE HP and incubate for 15 min at room temperature.
4. Add 530 μl of the mixture to single dishes of S2 cells and gently swirl to make sure the solution is well mixed.
5. Incubate the cells for 72 h at 25 °C.

3.1.2 Lysate Preparation

1. Centrifuge the cells in the four 50-ml tubes at 1500 × *g* for 5 min at room temperature. After centrifugation, wash the pellet in 20 ml of 1× PBS and centrifuge it again.
2. Resuspend the cells in the same volume of ice-cold 1× lysis buffer containing 1 mM DTT and 1× protease inhibitor cocktail.
3. Transfer the cells into a pre-chilled Dounce homogenizer.
4. Homogenize the cells by 20 strokes on ice.
5. Transfer the homogenized sample to 1.5-ml tubes and centrifuge them at 17,000 × *g* for 10 min at 4 °C.
6. Transfer the supernatant to new 1.5-ml tubes and centrifuge them at 17,000 × *g* for 10 min at 4 °C to clear the lysate.
7. Collect the supernatant.
8. Mix 4.5 μl of the supernatant and 0.5 μl of 10 μM HaloTag-TMR ligand. Incubate for 30 min at 25 °C and then perform SDS-PAGE analysis to check the concentration of the Ago2 protein.

3.1.3 Halo-Ago2 Labeling

1. Mix 400 μl of the lysate and 4 μl of 100 μM HaloTag-PEG-biotin ligand.
2. Incubate the reaction mixture for 30 min at 25 °C.
3. To remove the free ligands, apply the reaction mixture to a NAP-5 column equilibrated with 1× lysis buffer containing 1 mM DTT.
4. Apply 370 μl of 1× lysis buffer containing 1 mM DTT and discard the flow through.
5. Apply 480 μl of 1× lysis buffer containing 1 mM DTT and collect the flow through.
6. Make 10 μl aliquots of the collected flow through and flash-freeze with liquid nitrogen.
7. Store at −80 °C. The lysate containing biotin-labeled Ago2 may be stored for up to ~6 months.

3.2 Preparation of Fluorescently Labeled siRNA Duplexes

3.2.1 Annealing

1. Dissolve oligo pellets in MilliQ water to make 100 μM stocks and store at −80 °C.
2. Mix 10 μl of 100 μM single-stranded guide strand RNA, 10 μl of 100 μM single-stranded passenger strand RNA, 20 μl of 5× lysis buffer, and 60 μl MilliQ water.
3. Incubate the mixture in a heating block at 95 °C for 3 min.
4. Transfer the mixture from the heating block to the bench-top at room temperature and anneal it for 30 min. The annealed sample may be stored at −80 °C until PAGE purification.

*3.2.2 PAGE Purification (see **Note 1**)*

1. The electrophoresis should be performed in a cold room.
2. Chill 500 ml of 0.5× TBE at 4 °C before use.

3. Make a 12% native-polyacrylamide gel. Mix 6.6 ml of 40% acrylamide gel mix, 13.2 ml of MilliQ water, 4.4 ml of 5× TBE, 88 μl of 200 mM APS. Then mix 44 μl of TEMED and slowly pour the gel mixture into the glass plates, and immediately insert a 1-mm thick 5-well comb between the glass plates.
4. When the gel has solidified, carefully remove the comb and silicon-rubber gasket, and wash the wells with MilliQ water. The gel may be stored at 4 °C for a few days.
5. Set the gel in the electrophoresis apparatus and fill the upper and lower reservoirs of the electrophoresis tank with pre-chilled 0.5× TBE. Then remove any air bubbles trapped beneath the bottom of the gel.
6. Place the four ice-filled plastic bars in the running buffer to keep it at low temperature.
7. Pre-run the gel at 500 V for 30 min at 4 °C.
8. Mix 50 μl of the annealed RNA duplex sample and 50 μl of the native-PAGE loading dye.
9. Replace the ice-filled plastic bars with new frozen ones.
10. Directly load 20 μl of the sample into a well.
11. Run the gel at 500 V for ~60 min. Every 30 min, replace the ice-filled plastic bars with new frozen ones.
12. After electrophoresis, slowly open the glass plates and transfer the gel onto a thin clear plastic sheet. The separated duplex can be seen as colored bands.
13. Using LAS-3000, take fluorescent images with both red and green excitation lights to analyze the minor band caused by the excess of the guide or passenger strand.
14. Excise the band corresponding to the RNA duplexes with a clean razor blade and place the gel slices in a 1.5-ml tube.
15. Add 400 μl of 2× PK buffer and cover it with aluminum foil.
16. Gently shake the sample at room temperature for overnight.
17. The next day, centrifuge the tubes containing the gel slices for 5 min at maximum speed, room temperature.
18. Collect the supernatant from each tube, leaving the gel slice behind, and transfer into a fresh 1.5-ml tube.
19. Set up a DNA precipitation as follows:
    (a) Add 4 μl of glycogen to each tube.
    (b) Vortex vigorously for 30 s to mix.
    (c) Add 1 ml of ice-cold 100% ethanol to each tube.
    (d) Vortex vigorously for 15 s to mix.
20. Centrifuge the precipitation for 30 min at 21,500 × *g*, 4 °C.
21. Carefully remove the supernatant.

22. Add 1 ml of ice-cold 70% ethanol.
23. Centrifuge the precipitation for 5 min at 21,500 × *g*, 4 °C.
24. Carefully remove the supernatant and air dry the pellets by placing the open tubes on a bench for ~10 min.
25. Add 50 μl of 1× lysis buffer per tube and vortex vigorously to resuspend the pellets.
26. Mix 1 μl of the unpurified annealed duplex (10 μM) and 9 μl of 1× lysis buffer to make 1 μM electrophoresis standard duplex.
27. Mix 2 μl of the samples or the standard duplex with 2 μl of the native-PAGE loading dye.
28. Perform electrophoresis as described above.
29. Analyze the gel with LAS-3000 and calculate the concentration of the sample based on the band density of the standard.
30. Dilute the sample with 1× lysis buffer into 1 μM. The 1 μM stock can be stored at −80 °C.
31. Dilute the 1 μM stock with 1× lysis buffer into 400 or 200 nM and make 2 μl aliquots, and flash-freeze in liquid nitrogen.
32. Store aliquots at −80 °C.

3.3 Preparation of the PEG-Coated Quartz Slides and Coverslips

3.3.1 Pre-Cleaning of the Quartz Slides and Coverslips

1. Set 12 quartz slides in a slide rack and place them in the staining dish filled with 0.1 M KOH.
2. Set 12 coverslips in a coverslip rack and place them in the jar filled with 0.1 M KOH.
3. Sonicate the slides and coverslips for 10 min at room temperature.
4. Wash them in a small vat filled with fresh MilliQ water.
5. Transfer the slides and coverslips to the dish and jar filled with ethanol, respectively.
6. Sonicate the slides and coverslips for 10 min.
7. Wash them in the small vat filled with fresh MilliQ water.
8. Transfer the slides and coverslips to the dish and jar filled with fresh MilliQ water, respectively.
9. Sonicate the slides and coverslips for 10 min.
10. Wash them in the small vat filled with fresh MilliQ water.

3.3.2 Aminosilane Coupling

1. Mix 340 ml of methanol, 2.58 ml of acetic acid, 6.4 ml of aminosilane, and 12.75 ml of MilliQ water in a large stainless deep vat. Place a magnetic stirrer into it.
2. Transfer the slides and coverslips to the stainless vat with the reaction mixture. This procedure should be performed in a clean bench.
3. Incubate them for 20 min with stirring at room temperature.

4. Transfer them to the small stainless vat with fresh MilliQ water in a clean bench, and wash them for 5 min with continuous MilliQ water current.
5. Place the slides and coverslips on lab wipes and dry them in a clean bench for 30 min. A compressed air duster may also be used to blow the water away.

3.3.3 PEGylation of the Slides and Coverslips

1. Dissolve 1 mg of biotin-PEG in 20 μl of 50 mM MOPS-KOH (pH 7.5) to make a 50 mg/ml biotin-PEG solution.
2. Dissolve 43 mg of PEG in 215 μl of 50 mM MOPS-KOH (pH 7.5) to make a 200 mg/ml PEG solution.
3. Mix 2.4 μl of the biotin-PEG solution and 117.6 μl of the PEG solution to make a 0.5% biotin-PEG/PEG solution. When oversudsing, spin down the sample.
4. Place 9 μl of the 0.5% biotin-PEG/PEG solution on each quartz slide and then place a piece of Parafilm to make the center area of the slide covered with the solution. This procedure should be performed in a clean bench.
5. Place 8 μl of the PEG solution on each coverslip and then place a piece of Parafilm to make the center area of the slip covered with the solution. This procedure should be performed in a clean bench.
6. Incubate at room temperature for 3 h in a clean bench.
7. Before removing the piece of Parafilm, mark the area covered with the solution to make sure where is PEG-coated. Also, write down some marks to distinguish the front and back sides.
8. Remove the piece of Parafilm in fresh MilliQ water and rinse the slide or slip thoroughly with 20 strokes (for ~30 s). Dry using compressed air.
9. Place the slides and coverslips on lab wipes and further dry in a clean bench for 15 min.
10. Place the coated slides in a vacuum plastic bag and then seal the bag with a vacuum sealer.
11. Place the coverslips in a 25 ml plastic tube, place it inside a vacuum canister with silica gel desiccants to avoid hydrolysis of PEG, and then vacuum the canister.
12. Store them at −30 °C. The slides and coverslips may be kept for 1 month.

3.4 Single-Molecule Imaging

3.4.1 Preparation of Single-Molecule Observation Chamber (Fig. 2)

1. Make sure which side of the slide and coverslip is PEG-coated.
2. Place the double-sided tape on the flat bottom (back side) of a PCR ice cooler and cool it for 15 s. This step allows the protective plastic foil to be removed easily.
3. Apply the double-sided tape on the PEG-coated side of the coverslip and remove the other protective foil.

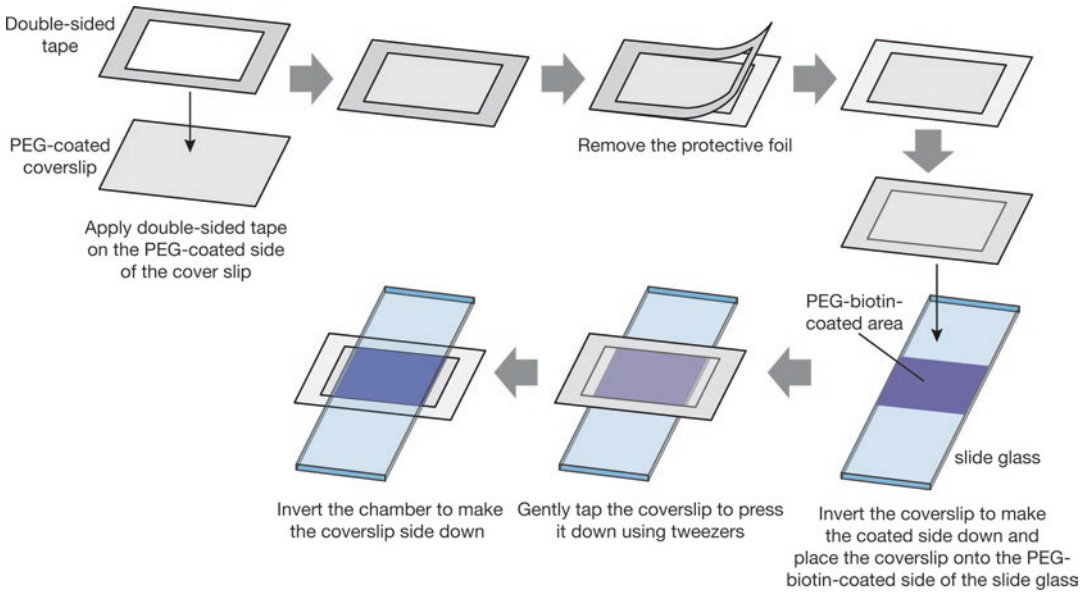


Fig. 2 Preparation of the single-molecule observation chamber. The observation chamber was made by combining a PEG-biotin-coated slide glass and a PEG-coated coverslip with double-sided tape in between. The sample chamber volume was ~15 μl. *See* Subheading 3.4 for details

4. Invert the coverslip to make the PEG-coated side down and place it onto the PEG-coated side of the slide. Use tweezers to gently tap the coverslip to press it down. Then invert the chamber to make the coverslip side down.
5. The sample chamber volume should be ~15 μl. Fluid can be flown through the chamber by placing the pipette tip in one hole and gently expelling the liquid from the pipette through the chamber. Use stripes of filter paper to suck the excess of liquid out from the other hole of the chamber.
6. Infuse 15 μl of 2.5 mg/ml NeutrAvidin solution into the chamber and incubate for 2 min. Wash the chamber with 50 μl of 1× lysis buffer.
7. Flow 20 μl of a 1:5 or 1:10 dilution (for snapshot or continuous observation, respectively) of S2 lysate containing biotin-labeled Halo-Ago2 into the chamber. Incubate for 2 min.
8. Wash the chamber fourth with 25 μl of 1× lysis buffer containing 800 mM NaCl.
9. Wash the chamber fourth with 25 μl of 1× lysis buffer.
10. Flow 20 μl of 1× lysis buffer containing 1% Biolipidure-203 and 50 μg/ml tRNA mix into the chamber. Incubate for 5 min to block the surface.

3.4.2 Sample Preparation

1. Prepare the ATP regenerating system on ice.
2. Prepare the fresh solution of the oxygen scavenging system on ice.
3. Prepare the RISC assembly system for single-molecule observation on ice:

   2 μl of 5× lysis buffer.

   4 μl of MilliQ water.

   4 μl of 5% Biolipidure-203.

   2 μl of ATP regenerating system.

   1 μl of 400 nM the Dicer-2/R2D2 binary complex.

   1 μl of 24 μM the Hop/Hsp90/Hsc70 ternary complex.

   1 μl of 15 μM DroJ2.

   1 μl of 54 μM p23.

   2 μl of the oxygen scavenging system.

   1 μl of the siRNA duplex (400 nM for snap observation or 200 nM for continuous monitoring).

   1 μl of 1 mg/ml tRNA mix.

   0.2 μl of 200 mM Trolox.

3.4.3 Snapshot Observation of RISC Assembly

1. Before the experiments, the excitation lights are aligned to generate the evanescent field. Make sure that the incident lights are totally internally reflected at the quartz–water interface and are well overlapped. The critical angle for the interface between water and quartz is ~65.6 degrees for visible light. In our experiments, the incident angle of excitation light at the quartz-water surface was set to be 69 degrees to make the evanescent field thin. For the snapshot observation, the sample is excited simultaneously with both 514 nm and 633 nm lasers at a power density of 2.5 W/mm^2 and 360 mW/mm^2, respectively. The frame rate is set to ten frames per second (fps). The temperature of EMCCD is set to −85 °C.
2. Flow 20 μl of the RISC assembly system into the chamber and incubate for 60 min at room temperature.
3. Wash the chamber twice with 25 μl of 1× lysis buffer.
4. Wash the chamber fourth with 25 μl of 1× lysis buffer containing 800 mM NaCl.
5. Wash the chamber fourth with 25 μl of 1× lysis buffer.
6. Flow 20 μl of 1× lysis buffer containing oxygen scavenging system and 2 mM Trolox into the chamber.
7. Place immersion oil on the objective lens and set the chamber onto the microscope stage. The coverslip should face down.

8. Place ~100 μl of glycerol on the upper side of the quartz slide. Then place the prism on the quartz slide. Glycerol fills the gap between the quartz prism and slide glass, allowing the incident light to enter into the slide glass at the designed angle.
9. Excite the sample simultaneously with both 514 nm and 633 nm lasers at power densities of 2.5 W/mm^2 and 360 mW/mm^2, respectively. Make sure that the incident lights are totally internally reflected at the quartz–water interface and are well overlapped.
10. Select and image three random and non-overlapping places from one chamber for 1 s at a frame rate of 10 fps. Save the images as Andor SIF format. FITS (Flexible Image Transport System) format is also usable, in which the file header part is more clearly described than SIF. FITS files can be opened by ImageJ without additional plugins.

*3.4.4 Continuous Monitoring of RISC Assembly (see **Note 2**)*

1. For the continuous monitoring, excite the sample simultaneously with both 514 nm and 633 nm lasers at power densities of 250 mW/mm^2 and 36 mW/mm^2, respectively. Because the power densities of excitation lights are set to tenfold lower than those for snapshot observation to minimize the photobleaching effect and therefore the photon budget is limited, the electron multiplying (EM) gain is set to the maximum (1000×). The frame rate is set to 1 fps.
2. Prepare the single-molecule observation chamber.
3. Place immersion oil on the objective lens and set the chamber onto the microscope stage. The coverslip should face down.
4. Place ~100 μl of glycerol on the upper side of the quartz slide. Then place the prism on the quartz slide.
5. Infuse 20 μl of the RISC assembly system into the chamber for continuous monitoring.
6. Observe the sample immediately for 20 min with simultaneous excitation of both 514 nm and 633 nm lasers.

3.4.5 Data Processing and Image Analysis for Snapshot Observation Data

1. Open the SIF file as a stack on ImageJ using the Open_SIFkai plugin. A slight modification of the original source code of Open_SIFkai plugin might be required to match it with the version of SOLIS software. No modification is required for FITS files.
2. Average the image over ten frames and then deconvolute. Essentially the same results are obtained regardless of deconvolution; however, deconvolution makes the following steps easier.
3. Select all the spots in every three images taken from one chamber using 6-pixel diameter circular regions of interest (ROIs) (*see* **Note 3**). To measure the background intensity, also select 30 ROIs of the background randomly.

4. Measure the fluorescent intensity of the spots using a built-in function of Image J. Then measure and calculate the average and the standard deviation (σ) of the background intensity in the same way.
5. Analyze only spots with fluorescence intensities greater than 3 σ of the background intensity in order to exclude shot noise.

3.4.6 Data Processing and Image Analysis for Continuous Monitoring Data

1. Open the SIF file as a stack on ImageJ using the Open_SIFkai plugin.
2. Perform frame averaging on each frame using a window size of five frames to improve the signal-noise ratio. That is, the new frame i is produced from the average projection of the original frames $i - 2$, $i - 1$, i, $i + 1$, and $i + 2$.
3. Correct stage drift using a slice alignment plugin.
4. To pick all of the Alexa 555 and 647 colocalized fluorescent spots appeared during the monitoring, apply Image J built-in smooth filter on each frame, and then create a maximum intensity projection over all frames in stack.
5. Select all the spots in every three images taken from one chamber using 6-pixel diameter circular regions of interest (ROIs). To measure the background intensity, also select 30 ROIs of the background randomly.
6. Generate integrated intensity traces of all selected ROIs using a built-in function of Image J. Then measure and calculate the average and the standard deviation (σ) of the background intensity.
7. Subtract the background intensity +3 σ from intensity traces.
8. Identify concurrent landing events of Alexa 555 and 647. Reject events that do not represent concurrent landing from the following analysis.
9. Find the appearance time of the simultaneous Alexa 555 and 647 signals (t_1) and subtract it from the disappearance time of either or both signals (t_2).
10. Measure duration times ($t_2 - t_1$).

3.5 Observation of the Target Cleavage Reaction by RISC

Basically, the observation method for the target cleavage reaction is similar to that for the RISC assembly analysis, as briefly summarized below.

1. Prepare lysate as described above by Dounce homogenization of S2 cells using 1× lysis buffer containing 1 mM DTT and 1× complete EDTA-free protease inhibitor cocktail.
2. Prepare the duplexes as in Subheading 3.2. but can be used without PAGE purification, as excess passenger stand will be readily degraded in *Drosophila* S2 cell lysate.

3. Prepare PEG-coated quartz slides and coverslips as in Subheading 3.3.
4. Single-molecule imaging is conducted as follows:
   (a) Program fluorescent Ago2-RISC with ~50 nM siRNA duplex fluorescently labeled on the guide strand in S2 lysate.
   (b) Prepare the chamber as in Subheading 3.4.
   (c) Infuse NeutrAvidin into the chamber and wait a few minutes, then wash by 50 μl of 1× lysis buffer.
   (d) Flow target RNAs (2.5–4 nM) double-labeled by a biotin and a fluorescent dye into the chamber. After a few minutes, wash it by 50 μl of 1× lysis buffer. Target density on the glass surface affects the overall multiple-turnover reaction speed.
   (e) Flow 50 μl of naive (non-programmed) S2 lysate containing oxygen scavenging system and ATP regenerating system into the chamber.
   (f) Flow 50 μl of the lysate containing fluorescent Ago2-RISC, oxygen scavenging system and ATP regenerating system into the chamber. To control the reaction speed, the lysate with fluorescent Ago2-RISC can be diluted by naive (non-programmed) S2 lysate or by decreasing the siRNA concentration at the RISC programming step.
   (g) Monitor cleavage reaction using the EMCCD camera with 0.3–1 fps. The condition of laser power and camera frame rate should be suitably chosen depending on the combination of dye-labeled samples. For example, Cy3-guide with 5′-ATTO 647 N–labeled target and Cy5-guide with 5′-Alexa 555-labeled target pairs show strong fluorescence resonance energy transfer (FRET, ~100%), thus dramatically shortening the fluorophore lifetime. Therefore, the laser intensity should be decreased to obtain long enough fluorophore lifetime compared to the duration of the cleavage reaction.
   (h) Data processing and image analysis are essentially the same as those for RISC assembly.

4 Notes

1. The excess of either Alexa555 or 647 labeled RNA caused significant non-specific fluorescent spots on the glass. Therefore, native PAGE purification of fluorescently labeled RNA duplexes is essential for this experiment.

2. In the microscopic observation, mechanical drift in the XY plane as well as focus drift in the Z axis is a severe problem for long time monitoring. To overcome this problem, two approaches—hardware and software—can be used. For the hardware improvement, an ultrastable stage, a motorized Z-drift compensator and/or a rigid microscope body will be helpful. Also minimizing the room temperature change, which affects the metal heat extension (length) and thus causes the stage, microscope body and/or mirror drift, will improve the stability. For the post-production improvement, drift markers are useful. We used a non-specific aggregate, which was often observed in our continuous monitoring condition, as a drift marker for XY drift correction. Biotinylated fluorescent beads, Qdots, or gold nanoparticles could also be used as a fiducial drift marker.
3. Recent progress in single-molecule localization microscopy provides us several pieces of Gaussian fitting-based spot finding software [25]. Using one of these software applications to find single-molecule spots or events automatically improves the throughput of single-molecule analysis.

Acknowledgements

We thank S. Iwasaki for constructing vectors for protein expression and establish the in vitro reconstitution system of fly Ago2-RISC assembly. This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas ('Functional machinery for non-coding RNAs' 21115002 and 'Non-coding RNA neo-taxonomy' 26113007) (to H.T. and Y.T.), and Grants-in-Aid for challenging Exploratory Research (24657115 and 26650047) (to H.M.S.) from The Ministry of Education, Culture, Sports, Science and Technology in Japan.

References

1. Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136(4):642–655
2. Kawamata T, Tomari Y (2010) Making Risc. Trends Biochem Sci 35(7):368–376
3. Liu Q, Rand TA, Kalidas S, Du F, Kim HE, Smith DP, Wang X (2003) R2d2, a bridge between the initiation and effector steps of the *Drosophila* Rnai pathway. Science 301(5641):1921–1925
4. Pham JW, Pellino JL, Lee YS, Carthew RW, Sontheimer EJ (2004) A dicer-2-dependent 80s complex cleaves targeted Mrnas during Rnai in *Drosophila*. Cell 117(1):83–94
5. Tomari Y, Matranga C, Haley B, Martinez N, Zamore PD (2004) A protein sensor for siRNA asymmetry. Science 306(5700):1377–1380
6. Tomari Y, Du T, Haley B, Schwarz DS, Bennett R, Cook HA, Koppetsch BS, Theurkauf WE, Zamore PD (2004) Risc assembly defects in the drosophila Rnai mutant Armitage. Cell 116(6):831–841
7. Iki T, Yoshikawa M, Nishikiori M, Jaudal MC, Matsumoto-Yokoyama E, Mitsuhara I, Meshi

T, Ishikawa M (2010) In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone Hsp90. Mol Cell 39(2):282–291

8. Iwasaki S, Kobayashi M, Yoda M, Sakaguchi Y, Katsuma S, Suzuki T, Tomari Y (2010) Hsc70/hsp90 chaperone machinery mediates Atp-dependent Risc loading of small RNA duplexes. Mol Cell 39(2):292–299
9. Kawamata T, Seitz H, Tomari Y (2009) Structural determinants of Mirnas for Risc loading and slicer-independent unwinding. Nat Struct Mol Biol 16(9):953–960
10. Yoda M, Kawamata T, Paroo Z, Ye X, Iwasaki S, Liu Q, Tomari Y (2010) Atp-dependent human Risc assembly pathways. Nat Struct Mol Biol 17(1):17–23
11. Miyoshi T, Takeuchi A, Siomi H, Siomi MC (2010) A direct role for Hsp90 in pre-Risc formation in drosophila. Nat Struct Mol Biol 17(8):1024–1026
12. Betancur JG, Tomari Y (2012) Dicer is dispensable for asymmetric Risc loading in mammals. RNA 18(1):24–30
13. Kawamata T, Yoda M, Tomari Y (2011) Multilayer checkpoints for Microrna authenticity during Risc assembly. EMBO Rep 12(9): 944–949
14. Kawamata T, Tomari Y (2011) Native gel analysis for Risc assembly. Methods Mol Biol 725: 91–105
15. Iwasaki S, Sasaki HM, Sakaguchi Y, Suzuki T, Tadakuma H, Tomari Y (2015) Defining fundamental steps in the assembly of the drosophila Rnai enzyme complex. Nature 521(7553): 533–536
16. Yao C, Sasaki HM, Ueda T, Tomari Y, Tadakuma H (2015) Single-molecule analysis of the target cleavage reaction by the drosophila Rnai enzyme complex. Mol Cell 59(1):125–132
17. Jo MH, Shin S, Jung SR, Kim E, Song JJ, Hohng S (2015) Human Argonaute 2 has diverse reaction pathways on target Rnas. Mol Cell 59(1):117–124
18. Salomon WE, Jolly SM, Moore MJ, Zamore PD, Serebrov V (2015) Single-molecule imaging reveals that Argonaute reshapes the binding properties of its nucleic acid guides. Cell 162(1):84–95
19. Miyazono Y, Hayashi M, Karagiannis P, Harada Y, Tadakuma H (2010) Strain through the neck linker ensures Processive runs: a DNA-Kinesin hybrid Nanomachine study. EMBO J 29(1):93–106
20. Yildiz A, Forkey JN, Mckinney SA, Ha T, Goldman YE, Selvin PR (2003) Myosin V walks hand-over-hand: single Fluorophore imaging with 1.5-nm localization. Science 300(5628):2061–2065
21. Thompson RE, Larson DR, Webb WW (2002) Precise nanometer localization analysis for individual fluorescent probes. Biophys J 82(5):2775–2783
22. Zhou ZP, Shimizu Y, Tadakuma H, Taguchi H, Ito K, Ueda T (2011) Single molecule imaging of the trans-translation entry process via anchoring of the tagged ribosome. J Biochem 149(5):609–618
23. Tseng Q, Duchemin-Pelletier E, Deshiere A, Balland M, Guillou H, Filhol O, Thery M (2012) Spatial organization of the extracellular matrix regulates cell-cell junction positioning. Proc Natl Acad Sci U S A 109(5):1506–1511
24. Ueno T, Taguchi H, Tadakuma H, Yoshida M, Funatsu T (2004) GroEL mediates protein folding with a two successive timer mechanism. Mol Cell 14:423–434
25. Small A, Stahlheber S (2014) Fluorophore localization algorithms for super-resolution microscopy. Nat Methods 11(3):267–279

Chapter 11

Profiling Open Chromatin Structure in the Ovarian Somatic Cells Using ATAC-seq

Kensaku Murano, Yuka W. Iwasaki, and Haruhiko Siomi

Abstract

The *a*ssay for *t*ransposase-*a*ccessible *c*hromatin using *seq*uencing (ATAC-seq) was recently established as a method to profile open chromatin, which overcomes the sample size limitations of the alternative methods DNase/MNase-seq. To investigate the role of Piwi in heterochromatin formation around transposable element loci, we have used ATAC-seq to examine chromatin accessibility at target transposable elements in a *Drosophila* cultured cell line, ovarian somatic cells (OSCs). In this chapter, we describe our method to profile open chromatin structure in OSCs using ATAC-seq.

Key words ATAC-seq, Chromatin accessibility, Heterochromatin, Small RNA, Transposons, *Drosophila*, OSC

1 Introduction

The gene expression pattern of a cell is characterized by a defined set of *cis*-regulatory elements that are accessible to *trans*-acting factors in the chromatin regulatory network. The chromatin state is dynamically regulated in a cell type-specific manner. The fundamental subunit of chromatin is the nucleosome core particle. Nucleosome packaging restricts protein binding and interferes with cellular reactions that use a DNA template, including transcription, replication, and repair [1]. In eukaryotes, all DNA-templated reactions occur in the context of chromatin. Local modulation of DNA accessibility provides an opportunity to influence these DNA-templated reactions. An accessible/open chromatin structure marks active regulatory elements, such as promoters, transcription start sites, enhancers, and insulators [2]. Deoxyribonuclease (DNase) or micrococcal nuclease (MNase) digestion coupled with high-throughput DNA sequencing (DNase-seq and MNase-seq) has been widely used to identify active regulatory regions and profile the epigenome [1, 3].

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_11,

The *a*ssay for *t*ransposase-*a*ccessible *c*hromatin using sequencing (ATAC-seq) provides an alternative to DNase-seq and MNase-seq as a way to measure open chromatin. ATAC-seq uses the prokaryotic Tn5 transposase to tag regulatory regions by inserting sequencing adapters into accessible regions of the genome [4] (Fig. 1). Therefore, genomic regions with open chromatin regions, such as transcription start sites, generate an accumulation of ATAC-seq reads, whereas chromatin-dense regions, such as heterochromatinized transposable element (TE) loci, result in decreased ATAC-seq signals [5]. ATAC-seq captures open chromatin sites using a simple two-step protocol, because tagging with Tn5 transposase results in a concomitant fragmentation of the genome (Fig. 1b). This makes the method faster and more sensitive than the established methods of DNase-seq and MNase-seq [4].

Mobilization of TEs can result in deleterious mutations in the host genome. Emerging evidence indicates that PIWI-clade proteins of Argonautes are involved in TE silencing in *Drosophila* gonadal somatic and germline cells [6, 7]. PIWI-interacting RNAs (piRNAs) form piRNA-induced silencing complexes (piRISCs) with PIWI proteins to repress TEs and maintain the integrity of the germline genome. The recent establishment of the Ovarian Somatic Cell (OSC) line allows researchers to study the molecular functions of factors

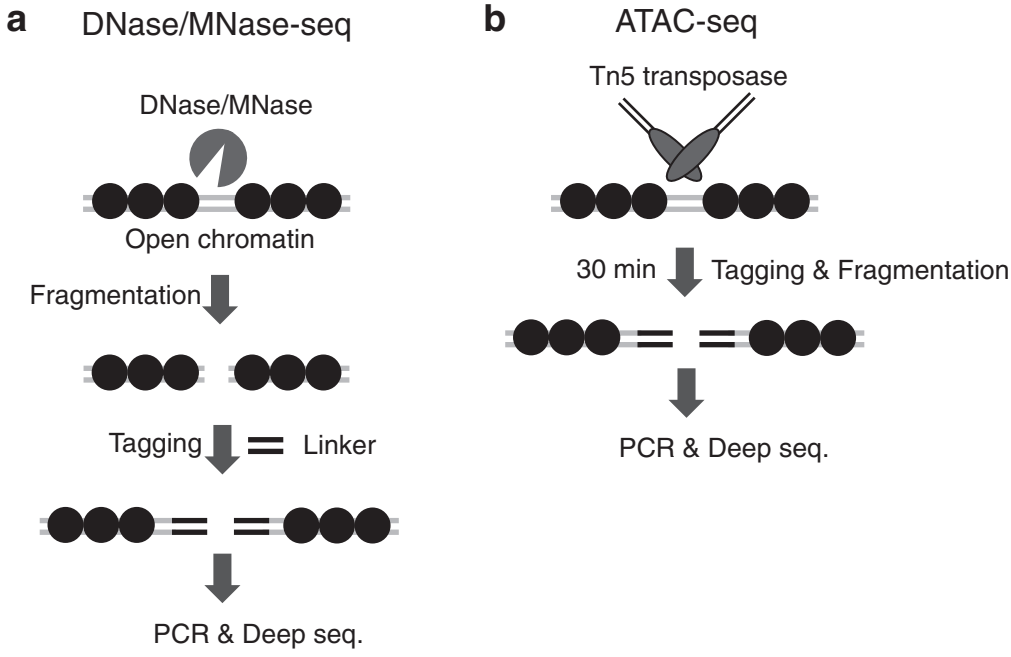


Fig. 1 ATAC-seq probes open chromatin. (**a**) DNase/MNase-seq consists of three steps: fragmentation, tagging with linker DNA, and PCR and deep sequencing. (**b**) ATAC-seq uses prokaryotic Tn5 transposase to simultaneously cut and ligate linker DNA for deep sequencing at regions of increased accessibility

involved in the primary piRNA pathway in *Drosophila* [8, 9]. Earlier studies showed that OSCs reproduce the Piwi-piRNA pathway observed in somatic cells of *Drosophila* ovaries, and therefore represent a very powerful system to elucidate how Piwi-piRISCs repress TEs. It was shown that Piwi-piRNA complexes repress TE transcription by forming a heterochromatin structure marked with H3K9 trimethylation and the linker histone H1 at TE loci [5, 10, 11]. To explore the role of Piwi in heterochromatin formation, we took advantage of ATAC-seq to evaluate the chromatin accessibility of TE loci in OSCs. The ATAC-seq data strongly support the notion that Piwi-piRNA modulates chromatin accessibility to repress TEs [5].

To profile the open chromatin structure in *Drosophila* OSCs using ATAC-seq, we had to partially modify the original ATAC-seq protocol described by Buenrostro et al. [4]. Briefly, we developed an optimized protocol for use on OSC to prepare DNA library for ATAC-seq because the genome size of *Drosophila* is approximately a twentieth part of that of humans. A concentration of primers was also optimized to amplify DNA library in PCR, because a higher concentration of primers causes the production of DNA derived from the primers, which probably interferes with the deep sequencing.

In this chapter, we describe the details of (1) DNA library preparation for ATAC-seq optimized for OSCs; and (2) deep sequencing and bioinformatic analysis of ATAC-seq data.

2 Materials

1. *Drosophila* Ovarian Somatic Cells (OSCs) (*see* **Note 1**).
2. 0.5 ml Siliconized tube (Sarstedt).
3. Dulbecco's Phosphate-buffered saline (D-PBS, 8 g/l NaCl, 0.2 g/l KCl, 0.2 g/l KH_2PO_4, 1.15 g/l Na_2HPO_4 (anhydrous)).
4. Lysis buffer: 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL CA-630 (Sigma-Aldrich).
5. Tagment DNA buffer (Nextera DNA sample Preparation Kit, Illumina).
6. Tagment DNA Enzyme 1 (Nextera DNA Sample Preparation Kit, Illumina).
7. MinElute PCR purification kit (Qiagen).
8. NEBNext High-Fidelity 2× PCR Master Mix (New England Biolabs).
9. SYBR Green I Nucleic Acid Gel Stain (10,000× concentrate, Thermo Fisher).

10. High Sensitivity DNA Reagents and High Sensitivity DNA Chips for use with the Agilent 2100 Bioanalyzer System (Agilent Technologies).
11. MiSeq Reagent Kit v3 150 cycles (Illumina).
12. HT1 included in MiSeq Reagent Kit v3 150 cycles (Illumina).
13. 0.1 N NaOH (prepared freshly).
14. Bioanalyzer (Agilent Technologies).
15. Primers (a complete list is available in Buenrostro et al. [4]). Primer 1.

Ad1_noMX: AATGATACGGCGACCACCGAGATCTACACTC GTCGGCAGCGTCAGATGTG.

Barcoded primer 2 (*see* **Note 2**).

Ad2.1_TAAGGCGA: CAAGCAGAAGACGGCATACGAGATTC GCCTTAGTCTCGTGGGCTCGGAGATGT.

Ad2.2_CGTACTAG: CAAGCAGAAGACGGCATACGAGATCT AGTACGGTCTCGTGGGCTCGGAGATGT.

Ad2.3_AGGCAGAA: CAAGCAGAAGACGGCATACGAGATTT CTGCCTGTCTCGTGGGCTCGGAGATGT.

Ad2.4_TCCTGAGC: CAAGCAGAAGACGGCATACGAGATGC TCAGGAGTCTCGTGGGCTCGGAGATGT.

Ad2.5_GGACTCCT: CAAGCAGAAGACGGCATACGAGATAG GAGTCCGTCTCGTGGGCTCGGAGATGT.

3 Methods

3.1 Preparation of DNA Library

1. Harvest and count OSCs.
2. Transfer 1×10^6 cells to a 0.5-ml siliconized sampling tube (*see* **Note 3**).
3. Centrifuge cells at $500 \times g$, 4 °C for 2 min.
4. Remove and discard supernatant.
5. Wash cells in 50 μl of PBS. Centrifuge cells at $500 \times g$, 4 °C for 2 min.
6. Remove and discard supernatant.
7. Resuspend cells in 50 μl of ice-cold lysis buffer by gentle pipetting.
8. Centrifuge cells immediately at $500 \times g$, 4 °C for 2 min, twice. After the first centrifugation, change the direction of the tube, and spin the samples again (*see* **Note 4**).
9. Remove and discard supernatant.

10. Resuspend nuclei pellet in 50 μl of the transposition reaction mix by gentle pipetting.

Transposition reaction mix	
Milli-Q water	22.5 μl
Tagment DNA buffer	25 μl
Tagment DNA Enzyme 1	2.5 μl

11. Incubate the transposition reaction at 37 °C for 30 min in a water bath. Mix the reaction by gentle tapping every 10 min during the incubation.
12. Stop the reaction by adding 250 μl of buffer PB included in the Qiagen MinElute PCR purification kit.
13. Purify the fragmented and tagged DNA fragment (transposed DNA) using a Qiagen MinElute PCR purification kit according to the manufacturer's instructions.
14. Elute the transposed DNA in 12 μl of elution buffer included in the kit.
15. Store the purified transposed DNA at −20 °C if necessary.
16. To amplify the transposed DNA by PCR, combine the following in a 0.2-ml PCR tube (*see* **Note 2**).

Transposed DNA	10 μl
Milli-Q water	13.7 μl
PCR primer 1 (10 μM)	0.63 μl
Barcoded PCR primer 2 (10 μM)	0.63 μl
NEBNext High-Fidelity (2× PCR mix)	25 μl

17. Thermal cycle as follows (*see* **Note 5**):

72 °C	5 min	
98 °C	30 s	
98 °C	10 s	⎡
63 °C	30 s	⎢ 5 cycles
72 °C	1 min	⎣

18. Perform quantitative PCR using a real-time PCR machine. A quantitative PCR side reaction allows determination of the appropriate number of PCR cycles. Combine the following in a 0.2-ml PCR tube.

PCR-amplified DNA (From step 3.1.17)	5 µl
Milli-Q water	2.85 µl
PCR primer 1 (1 µM)	0.625 µl
Barcoded PCR primer 2 (1 µM)	0.625 µl
10× SYBR Green I	0.9 µl
NEBNext High-Fidelity (2× PCR mix)	5 µl

98 °C	30 s	
98 °C	10 s	20 cycles
63 °C	30 s	
72 °C	1 min	

19. Determine the additional number of cycles needed (X cycles) from the quantitative PCR plot prior to saturation (maximum fluorescent intensity) (*see* **Note 6**, Fig. 2).
20. Run the remaining 45 µl PCR reaction.

98 °C	30 s	
98 °C	10 s	X cycles
63 °C	30 s	
72 °C	1 min	

21. Purify the amplified library with a Qiagen MinElute PCR Purification Kit. Elute 20 µl with EB.
22. Assess the quality and concentration of the library using a Bioanalyzer (*see* **Note 7**, Fig. 3).

3.2 Denaturation of the DNA Library for Deep Sequencing by MiSeq

1. Prepare 4 nM of library (mix samples and diluted with Milli-Q water) (*see* **Note 8**).

2. Mix the following (Denatured Library). Incubate at room temperature for 5 min (*see* **Note 9**).

4 nM Library	10 μl
0.1 N NaOH	10 μl

3. Mix the following (final 20 pM in 0.5 mM NaOH). Keep on ice.

Denatured Library (2 nM)	10 μl
Iced HT1	990 μl

4. Mix the following (final 15 pM). Keep on ice.

Denatured Library (20 pM)	750 μl
Iced HT1	250 μl

5. Load 600 μl sample from step 3.2.4 (15 pM) to the correct position on the cartridge tray for MiSeq.

3.3 Deep Sequencing and Bioinformatic Analysis of ATAC-seq Data

This section describes an example bioinformatic analysis of sequenced reads from an ATAC-seq library, to annotate the chromatin accessibility of OSCs in a genome-wide manner.

1. Perform deep sequencing using an Illumina sequencing system as per the manufacturer's instructions (*see* **Note 10**).
2. Check the quality of the reads using software such as FastQC [12], and remove adapters prior to mapping of the ATAC-seq reads. Reads that are too short should be filtered out at this point (*see* **Note 11**).
3. Map the reads using alignment software such as bowtie2 [13]. Analyze the size distribution of the mapped reads using Picard [14] to confirm the quality of the library using the pattern of nucleosomes (Fig. 4a) (*see* **Notes 12** and **13**).
4. Detect peaks using peak-calling software such as MACS2 [15]. Analyze and annotate the detected peaks (Fig. 4b) (*see* **Notes 14** and **15**).
5. Visualize the mapped reads using a genome browser or viewer, such as the University of California Santa Cruz (UCSC) genome browser [16] (Fig. 4c) (*see* **Note 16**).

4 Notes

1. The method for culture of OSCs is described precisely by Saito [8].
2. The different barcoded-primers should be used for each sample. Primer concentration is critical to the quality of the DNA library for sequencing. A higher concentration of primers causes the production of DNA derived from the primers (Fig. 3a), which probably interferes with the deep sequencing. The level of primer-derived DNA also depends on the amount of transposed DNA. The concentration of primers should be optimized for each experimental condition.
3. 1×10^6 OSCs are required to prevent loss of the sample during library preparation. A siliconized sample tube is strongly recommended to avoid loss of permeabilized cell nuclei, because they are sticky and tend to attach to tube walls.
4. Centrifugation should be performed twice, changing the direction of the tube, to avoid loss of samples when using a fixed-angle rotor.
5. The first 5-min extension at 72 °C is required to allow extension of both ends of the primer after transposition.
6. According to the result of the quantitative PCR side reaction, determine the additional number of cycles that corresponds to, or exceeds, one-third of the maximum fluorescent intensity, to reduce GC and size bias in the PCR. In the case of Fig. 2, the additional number of cycles is five; therefore, the total PCR cycle number is ten.
7. Determine the concentration of the DNA library using an Agilent 2100 Bioanalyzer system (Fig. 3c).
8. We obtained acceptable ATAC-seq data from a mixed-DNA library prepared from five samples. Data were collected as 50-bp paired-end reads from MiSeq. Other Illumina deep sequencing systems such as HiSeq are also suitable for collecting ATAC-seq reads.
9. Although 0.2 N NaOH is recommended for denaturing the DNA library in the manufacturer's instructions, we use 0.1 N NaOH, based on our experience.
10. We recommend performing ATAC-seq with at least two replicates, in order to confirm the correlation between two samples. The reads should be analyzed as paired-ends, because the length of ATAC-seq reads varies widely from dozens to thousands of base pairs. Analyzing the distribution of read length is important to confirm the quality of the ATAC-seq library (Fig. 4a).
11. FastQC provides a modular set of sequence quality check analyses and gives warnings if the sequenced data has problems. The

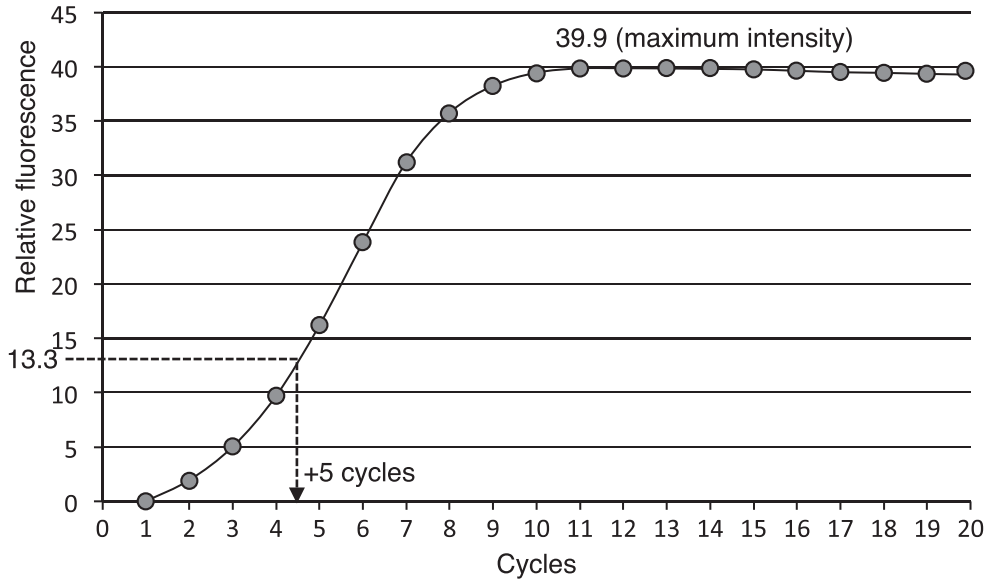


Fig. 2 Representative amplification plots of quantitative PCR side reactions. To avoid GC and size bias in the PCR, the appropriate number of PCR cycles should be determined using a quantitative PCR side reaction, which corresponds to or exceeds one-third (*blue dashed line*) of the maximum fluorescent intensity. In this experiment, the additional number of cycles is five. We determine the additional numbers of cycles required for each sample according to its amplification plot. A negative control experiment should be performed in parallel, without the transposed DNA, because primer-derived DNA is amplified in the absence of transposed-DNA derived from the OSC nuclei

results can be used to examine if adapter trimming is necessary for the data (by looking at information such as overrepresented sequences). Depending on the length of the ATAC-seq insert size and the sequencing read length, part of the adapter can be present within the read, so try trimming the reads prior to the mapping step if necessary. Also, omit reads that are too short for further analysis (we filtered out sequences shorter than 50 nt). The size threshold can be optimized depending on the analysis, but avoid too short reads which can be mapped to unspecific genomic regions. This can be achieved using an in-house script, or software such as Cutadapt [17].

12. When mapping reads using Bowtie2 [13], it will be important to specify the –X option, which allows the mapping of fragments up to an indicated length. We used 4000 to map reads up to 4000 bp, in order to accept long fragments in the ATAC-seq library. Default settings are used for the other options, but it should be optimized depending on the analysis. Refer to the project website for further information on options. Also, pre-built indexes may be useful for the mapping step. There are sets of genome sequence and pre-built indexes for various alignment software programs, including Bowtie2, at Illumina iGenomes [18].

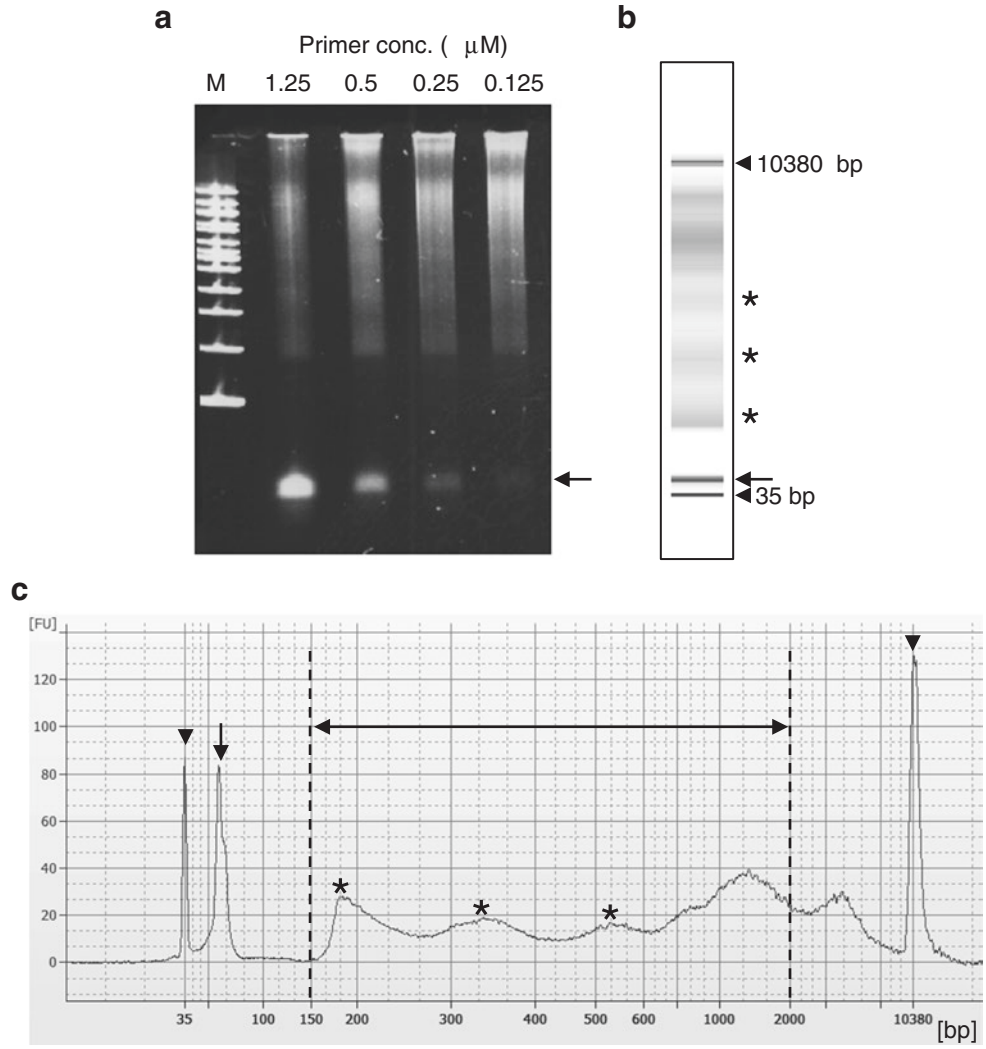


Fig. 3 Fragment sizes in an amplified ATAC-seq library. (**a**) DNA libraries were separated by 6% native polyacrylamide gel electrophoresis. A high concentration of primers caused the production of DNA derived from the primers (*arrow*). M indicates the 100-bp DNA ladder. (**b**, **c**) Electrophoresis patterns from a Bioanalyzer. The ATAC-seq library shows a nucleosome ladder with repeating unit size (*asterisks*). This electrophoresis pattern is a hallmark of the DNA library for ATAC-seq. *Arrow heads* indicate the DNA size marker. *Arrows* indicate DNA derived from the primers. The concentration of the DNA library should be calculated from the area between the *two dashed lines* (150 and 2000 bp) (*see* **Note 7**). Primer-derived DNA probably interferes with deep sequencing. The primer concentration should be optimized to avoid the production of primer-derived DNA (*see* **Note 2**)

13. ATAC-seq reads may contain a higher proportion of mitochondrial DNA than expected. The proportion of mitochondrial DNA within ATAC-seq reads depends on the cell/tissue types used. We recommend omitting the reads mapped to mitochondrial sequences before proceeding with further analysis.

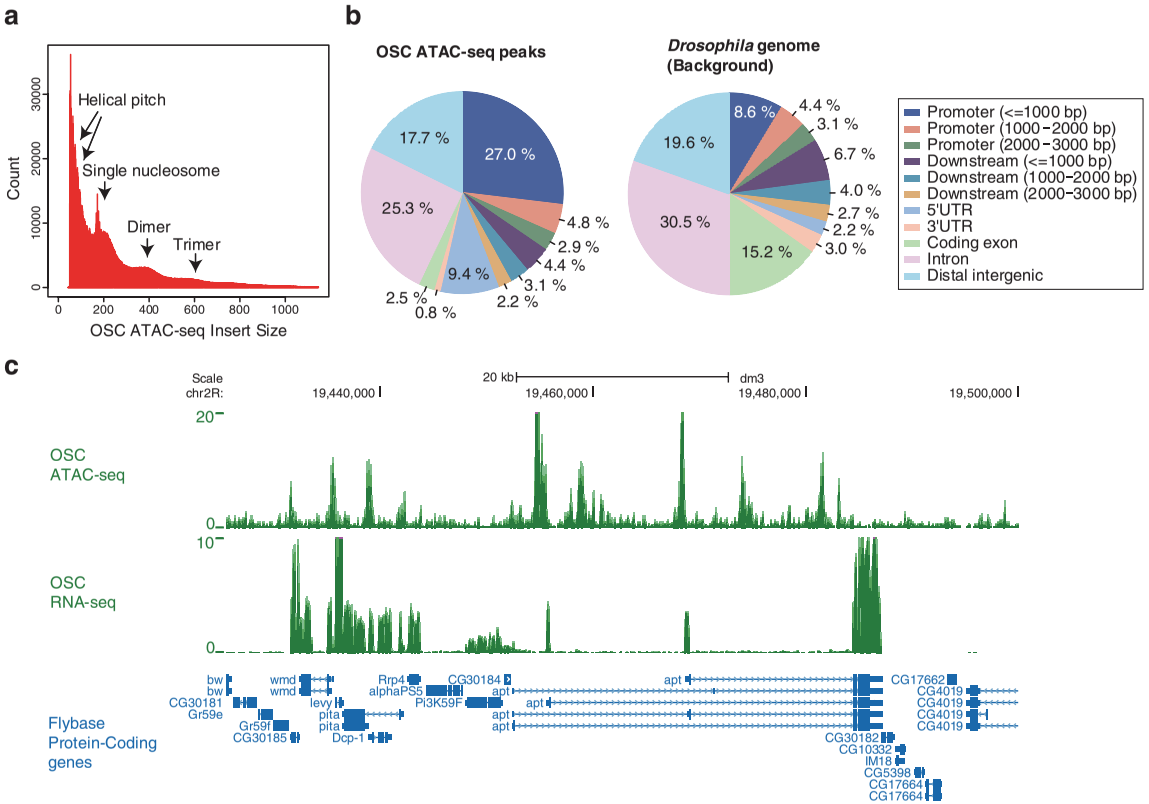


Fig. 4 Analysis of ATAC-seq data from OSCs. Previously published ATAC-seq data from OSCs treated with control small interfering RNAs [5] are used. (**a**) Insert size histogram for ATAC-seq reads calculated using Picard [14]. The distribution shows single, dimer, and trimer nucleosome patterns, and also the helical pitch of the DNA (with a periodicity of ~10.5 bp). (**b**) Annotation of ATAC-seq peaks calculated using CEAS [19]. Enrichment of ATAC-seq peaks at the promoter (≤1000 bp) and 5′ untranslated regions can be detected in the OSC ATAC-seq peaks (*left*) compared with the background *Drosophila* genome (*right*). This suggests higher accessibility of the chromatin structure near the transcription start site. (**c**) Example of ATAC-seq read distribution at Chromosome 2R visualized using a custom track on the UCSC genome browser [16]. OSC ATAC-seq reads are visualized along with OSC RNA-seq reads [5]. ATAC-seq and RNA-seq reads are shown in reads per million mapped reads. The positions of Flybase protein-coding genes are overlaid. ATAC-seq peaks, indicating higher chromatin accessibility, can be detected at the promoter regions of some protein-coding genes. The RNA-seq distribution shows that the genes with higher accessibility in their promoter regions are expressed in OSCs

14. When analyzing OSC ATAC-seq reads using MACS2 [15], we use the "–g dm –nomodel –nolambda" option. This specifies the genome size for *Drosophila* (–g option), not building the shifting model (–nomodel option), and not considering the local bias at peak candidate regions (–nolambda option). Refer to the project website for further information on options. Downstream analysis, such as comparison of the peaks from two different ATAC-seq data, can also be performed using MACS2.
15. Annotation of peaks can be performed using in-house scripts, or software such as CEAS [19] (Fig. 4b). As in the case of CEAS, there are tools developed for ChIP-seq analysis that can also be useful for ATAC-seq analysis. For extended annotation of the genes neighboring the detected peaks, web resources such as DAVID [20] may be used to identify Gene Ontology [21] terms enriched within the list of genes.
16. Using Custom Tracks from the UCSC genome browser [16], the distribution of two different sequencing analyses can be overlaid for comparison purposes (in the case of Fig. 4c, ATAC-seq, and RNA-seq data). Other information, such as protein-coding genes, repeats, and conservation between species, can also be overlaid to provide a comparison.

Acknowledgments

This work was supported by grants from the Kato Memorial Bioscience Foundation to K.M., and Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) to Y.W.I. and H.S.

References

1. Bell O, Tiwari VK, Thoma NH, Schubeler D (2011) Determinants and dynamics of genome accessibility. Nat Rev Genet 12(8):554–564. doi:10.1038/nrg3017
2. Gross DS, Garrard WT (1988) Nuclease hypersensitive sites in chromatin. Annu Rev Biochem 57:159–197. doi:10.1146/annurev.bi.57.070188.001111
3. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutyavin T, Lajoie B, Lee BK, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, Navas PA, Stamatoyannopoulos G, Iyer VR, Lieb JD, Sunyaev SR, Akey JM, Sabo PJ, Kaul R, Furey TS, Dekker J, Crawford GE, Stamatoyannopoulos JA (2012) The accessible chromatin landscape of the human genome. Nature 489(7414):75–82. doi:10.1038/nature11232
4. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10(12):1213–1218. doi:10.1038/nmeth.2688
5. Iwasaki YW, Murano K, Ishizu H, Shibuya A, Iyoda Y, Siomi MC, Siomi H, Saito K (2016)

Piwi modulates chromatin accessibility by regulating multiple factors including histone H1 to repress transposons. Mol Cell 63(3):408–419. doi:10.1016/j.molcel.2016.06.008

6. Iwasaki YW, Siomi MC, Siomi H (2015) PIWI-interacting RNA: its biogenesis and functions. Annu Rev Biochem 84:405–433. doi:10.1146/annurev-biochem-060614-034258
7. Siomi MC, Sato K, Pezic D, Aravin AA (2011) PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol 12(4):246–258. doi:10.1038/nrm3089
8. Saito K (2014) RNAi and overexpression of genes in ovarian somatic cells. Methods Mol Biol 1093:25–33. doi:10.1007/978-1-62703-694-8_3
9. Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC (2009) A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature 461(7268):1296–1299. doi:10.1038/nature08501
10. Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K (2013) DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. Genes Dev 27(15):1656–1661. doi:10.1101/gad.221515.113
11. Sienski G, Donertas D, Brennecke J (2012) Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151(5):964–980. doi:10.1016/j.cell.2012.10.040
12. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
13. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. doi:10.1038/nmeth.1923
14. Picard Tools. http://broadinstitute.github.io/picard/
15. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). Genome Biol 9(9):R137. doi:10.1186/gb-2008-9-9-r137
16. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. Genome Res 12(6):996–1006. doi:10.1101/gr.229102. Article published online before print in May 2002
17. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17(1):10–12. doi:http://dx.doi.org/10.14806/ej.17.1.200
18. iGenomes. http://support.illumina.com/sequencing/sequencing_software/igenome.html
19. Shin H, Liu T, Manrai AK, Liu XS (2009) CEAS: cis-regulatory element annotation system. Bioinformatics 25(19):2605–2606. doi:10.1093/bioinformatics/btp479
20. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4(1):44–57. doi:10.1038/nprot.2008.211
21. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25(1):25–29. doi:10.1038/75556

Chapter 12

Assessing miR-451 Activity and Its Role in Erythropoiesis

Dmitry A. Kretov, Andrew M. Shafik, and Daniel Cifuentes

Abstract

The ability to microinject small RNAs and mRNAs into zebrafish embryos, of different genetic backgrounds, allows for the precise dissection of microRNA processing pathways at the molecular level, while simultaneously provides insight into their physiologic role. Here, we apply such an approach to determine the impact of Argonaute 2 in the processing of miR-451, a vertebrate-specific microRNA required for terminal erythrocyte differentiation. This was achieved using fluorescent microRNA reporter sensor assays and phenotype rescue experiments.

Key words Zebrafish, MicroRNA, Ago2, Microinjection, miR-451, Erythrocyte

1 Introduction

Recent advances in genome editing approaches (e.g., using Zinc Finger Nucleases, TALENs, and CRISPRs) have allowed for the routine mutagenesis of specific gene targets in model organisms. However, there is a growing need to determine the specificity of these genome-editing tools. To this end, several assays/strategies have been developed to quantify on-target versus off-target effects [1]. In addition, the specificity of the various genome-editing methods can be assessed by comparing different mutant alleles generated by independent targeting sequences, or by complementation experiments that reintroduce the wild-type copy of the mutated gene.

Zebrafish is particularly suited to test the latter, as the rescue of the mutant phenotype is possible by microinjecting the mRNA of interest into the early embryo. This process enables researchers to discard off-target effects resulting from the mutagenesis process and confirm the causal relationship between the mutated genes and the affected molecular mechanisms.

Furthermore, the impact of such molecular mechanisms on the whole organism can be quickly determined as zebrafish embryos develop rapidly [2]. That is, one-day-old zebrafish already show

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_12,

a brain, beating heart, twitching muscles and flowing blood, permitting a comprehensive analysis of embryogenesis.

These features are particularly useful to study microRNA biogenesis pathways in zebrafish. MicroRNAs, and some of their pathway components, were first described using genetic means [3, 4]. However, precise molecular mechanisms regarding microRNA action were determined soon after through in vitro experiments [5–8]. Zebrafish has the advantage of directly integrating both approaches in a vertebrate model organism [9–11]. miR-451 is a vertebrate-specific microRNA that participates in terminal erythrocyte differentiation [12–14]. Interestingly, miR-451 processing by-passes Dicer cleavage and requires the slicer activity of Ago2 [9, 15, 16]

Here we describe a method to detect microRNA activity using fluorescent reporters in zebrafish, and precisely outline the role of Ago2 in miR-451 processing and erythropoiesis by performing a rescue experiment with small RNAs.

2 Materials

2.1 MicroRNA Reporter Assay

1. Forward and reverse oligos to clone miR-451 sensor (8-mer miR-451 seed sites are in bold): 5′-CAAGTAGCTC GAGCCTAGATAACTACCTGCACTGTG**AACGGT TA**GATAAGACTACCTGCACTGTG**AACGGTTA**GAT AACTATAGTGAGTCGTATTACG-3′, 5′-CGTAATA C G A C T C A C T A T A G T T A T C T A A C C G T TCACAGTGCAGGTAGTCTTATCTAACCGTTCACA GTGCAGGTAGTTATCTAGGCTCGAGCTACTTG-3′. Prepare a 100 μM stock solution for each oligo.
2. pre-miR-451 hairpin: 5′-rArArArCrCrGrUrUrArCr-CrArUrUrArCrUrGrArGrUrUrUrArGrUrA rArUrGrGrUr-ArArGrGrGrUrUrCrUrG-3′: Dilute with RNase-free water to make a 1 mM stock solution.
3. miR-451 duplex: 5′-rArArArCrCrGrUrUrArCrCrArUrUrAr-CrUrGrArGTT-3′; 5′-rCrUrCrArGrUrArArUrGrGrUrArAr-CrGrGrUrArUTT-3′. Dilute with RNase-free water to make a 1 mM stock solution.
4. pCS2-TagRFPT.zf1, Addgene plasmid #61390 [17]. This plasmid encodes a photostabilized version of zebrafish codon optimized red fluorescent protein (RFP) from *Entacmaea quadricolor*.
5. 1× TE buffer: 10 mM Tris–HCl pH 8.0, 1 mM EDTA.
6. Gibson assembly reaction master mix.
7. In vitro transcription kit with SP6 RNA polymerase.
8. NotI restriction enzyme (New England Biolabs).

9. System water: water from the fish facility supply. Specific water parameters are: pH: 7.2 to 7.6; conductivity: 500 μS; temperature: 28.5 °C and no traces of ammonia, nitrite, or nitrate.
10. Methylene Blue stock solution: dissolve to 0.1% (w/v) in system water.
11. Blue water: dilute 2 mL of methylene blue stock solution per liter of system water. Final concentration is 0.0002% (w/v) (*see* **Note 1**).
12. Petri dishes, of 100 mm and 35 mm diameter.
13. Agarose coating: pour melted 2% agarose in blue water on pippetor dishes (*see* **Note 2**).
14. 10 mL Pippetor.
15. 5¾ in. disposable controlled drop glass Pasteur pipettes.
16. Protease from *Streptomyces griseus* (Sigma-Aldrich), commonly called pronase. Dissolve 5 mg/mL in system water. Aliquot 2 mL in a 15 mL tube and store at −20 °C. Thaw and dilute to 1 mg/mL prior to use.
17. Injection plate cast (Adaptive Science Tools).
18. Thin wall glass filaments for injection needles, 4 in. long, 1 mm of outer diameter/0.75 mm inner diameter.
19. Fine point tweezers.
20. 20 μL microloader tips.
21. Micrometer slide.
22. Mineral oil (Sigma-Aldrich).
23. 3-in. diameter mini tea strainer stainless steel (OXO GoodGrips, Adaptive Science Tools).

2.2 Rescue of Erythropoiesis Defects in Ago2 Mutant

1. N-Phenylthiourea (1-phenyl-2-thiourea, PTU) (Sigma-Aldrich): dissolve to 0.3% (w/v) in system water for a 100× stock solution (*see* **Note 3**).
2. O-dianisidine stock solution: Dissolve 1.4 mg of O-dianisidine (Sigma-Aldrich) in 2 mL of 96% ethanol (*see* **Note 4**).
3. 100 mM sodium acetate pH 4.5.
4. 30% hydrogen peroxide solution.
5. Phosphate buffered saline (PBS; 10×): 100 mM H_3PO_4, 1.37 M NaCl, and 27 mM KCl, pH 7.4.
6. 4% Paraformaldehyde (PFA): Dissolve 2 g of PFA in 30 mL of deionized water with 100 μL of 100 mM NaOH. Once PFA is dissolved, add 5 mL of 10× PBS and adjust to pH 7.
7. 12-well plastic cell culture plate.

3 Methods

3.1 MicroRNA Reporter Assay

3.1.1 MicroRNA Reporter Sensor Cloning

1. To construct the miR-451 sensor, anneal the forward and reverse oligonucleotides containing the miR-451 target sites by mixing 10 μL Forward Primer + 10 μL Reverse Primer + 10 μL TE buffer. Incubate in a beaker with boiling water and let it cool down until it is not hot to the touch.
2. Clone annealed oligos with two 8-mer seed sites for miR-451 in the 3'untranslated region (UTR) of enhanced yellow fluorescent protein (EYFP) using the Gibson assembly method [18] following the manufacturer's instructions in pCS2+ vector to create the miR-451 sensor (*see* **Note 5**).
3. Linearize 5 μg of the miR-451 sensor sites with NotI, a restriction enzyme that cuts after the SV40 polyA signal (*see* **Note 6**).
4. In parallel, linearize pCS2-TagRFPT.zf1 using NotI (*see* **Note 7**).
5. Purify the DNA using a spin column kit and transcribe the DNA using the in vitro transcription kit with SP6 RNA polymerase. Incubate the reaction for 3 h at 37 °C.
6. Prepare separate tubes with the small RNAs at 10 μM and the third tube with a mix of mRNAs of the EYFP reporter and the TagRFP control at 0.1 μg/μL each (*see* **Note 8**).

3.1.2 Embryo Collection

1. The night before the injection, set individual pairs in 1.5 L breeding tanks. These tanks have a perforated insert that will allow the eggs go through and away from the parents that otherwise would eat them. Separate each male and female by inserting a transparent divider. Being in the same tank with the male but without contact opportunity will prime the female to lay eggs.
2. Next morning, remove the dividers from the breeding tanks to reunite the couple. 10 min after the parents start to lay eggs, transfer the parents with the perforated insert into a new breeding tank (*see* **Note 9**).
3. Pour the water of the first breeding tank through a tea strainer into a new tank to collect the fertilized eggs and flush the eggs with system water (*see* **Note 10**).
4. Invert the strainer over a 100 mm Petri dish (without the lid) and flush system water on the strainer to release the eggs into the dish and take them to the injection station.
5. Transfer the embryos with the help of the pipettor and a glass Pasteur pipette into a 35 mm Petri dish. Remove all the water and cover the eggs with pronase 1 mg/mL (*see* **Note 11**).
6. Incubate with pronase until at least two embryos start to pop-out of the chorion (embryo outer shell) (*see* **Note 12**).
7. Immediately immerse the Petri dish in a 250 mL beaker with 100 mL of system water (*see* **Note 13**).

8. Decant the water in the beaker until only a small amount of water covers the embryos in a corner of the tilted beaker (*see* **Note 14**).
9. In a quick motion, pour approximately 200 mL of water into the beaker, aiming directly over the embryos (*see* **Note 15**).
10. Gently transfer the embryos with the Pasteur pipette into a new agarose-coated Petri dish, remove the embryos that are leaking yolk and transfer the rest to the injection plate seeding the groves (*see* **Note 2**).

3.1.3 Embryo Microinjection

1. Make injection needles by pulling glass capillaries in a needle puller (*see* **Note 16**).
2. Load the needle with the small RNA or mRNA of interest with the help of a plastic loading tip and place it into the needle holder (*see* **Note 17**).
3. Break the tip of the needle with a Dumont forcep and position the tip of the needle inside a drop of mineral oil over the micrometer slide. Press the injection pedal and evaluate the diameter of the drop with the help of the engraved ruler on the micrometer slide (Fig. 1). Either increase the injection time or break further the needle until the desired drop diameter is reached (*see* **Note 18**).
4. Inject 1 nL of the EYFP miR-451 reporter and control TagRFP mixture in every embryo and then reinject them with either pre-miR-451, miR-451 duplex or nothing (Fig. 2a) (*see* **Note 19**).
5. After injection, transfer the embryos into an agarose-coated 100 mm Petri dish and incubate the embryos at 28.5 °C (*see* **Note 20**).
6. Coat a 60 mm Petri dish with agarose and pierce the agarose to form holes in order to hold the embryos in place (*see* **Note 21**).
7. Eight hours post fertilization, collect the embryos and arrange them in the holes of the imaging plate with water to the desired distribution (*see* **Note 22**).
8. Image the fluorescence, from the EYFP reporter (excitation peak 515 nm, emission peak 528 nm) [19]. Then, without moving the plate or the position of the embryos, image the fluorescence from the injection control TagRFP (excitation peak 555 nm, emission peak 584 nm) [20] (Fig. 2b) (*see* **Note 23**).

3.2 Rescue of Erythropoiesis Defects in Ago2 Mutant

1. In vitro transcribed capped mRNA of mouse Ago2 from the pCS2+/*Mm*Ago2 vector [9] using an SP6 RNA polymerase as described in the **steps 3–5** of Subheading 3.1.1.
2. Inject either pre-miR-451, miR-451 duplex or Ago2 mRNA as described before in the Subheading 3.1.3. into homozygous mutant Ago2 embryos. In parallel, collect wild-type embryos as a control. Incubate the embryos at 28.5 °C.

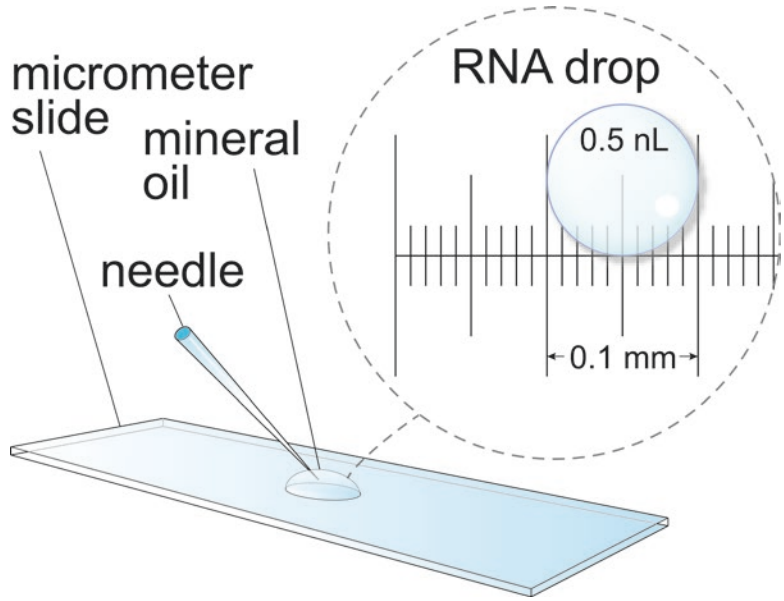


Fig. 1 Calibration of injection needle. A drop of oil covers the ruler in the micrometer slide. When the tip of the injection needle loaded with the mRNA solution is immersed in the oil, press the injection pedal once and measure the diameter of the ejected drop of mRNA against the backdrop of the micrometer ruler. A drop with a diameter of 0.1 mm will supply 500 pL, hence press the pedal twice when injecting each embryo to achieve the desired injection volume of 1 nL

3. At 9 h post-fertilization, add N-phenylthiourea (PTU) to the plates to a final concentration of 0.003% (w/v) to induce the oxidative stress in the embryo (*see* **Note 3**).
4. At 48 h post-fertilization, collect the embryos and proceed to stain hemoglobinized erythrocytes with *O*-dianisidine reagent [21, 22] (*see* **Note 24**).
5. Prepare the hemoglobin staining solution: Mix 2 mL of deionized water, 2 mL of O-dianisidine stock solution, 0.5 mL of 100 mM sodium acetate and 100 μL of 30% hydrogen peroxide (*see* **Note 4**).
6. Transfer phenyl-thiourea treated embryos to a 12-well plate, remove as much water as possible and gently add 2 mL of the staining solution. Incubate embryos with the staining solution in the dark for 15 min, at room temperature.
7. Remove the *O*-dianisidine staining solution and then wash three times with 2 mL of 1× PBS. Be very gentle at this step, since the embryos will become very brittle and are prone to bursting (*see* **Note 25**).

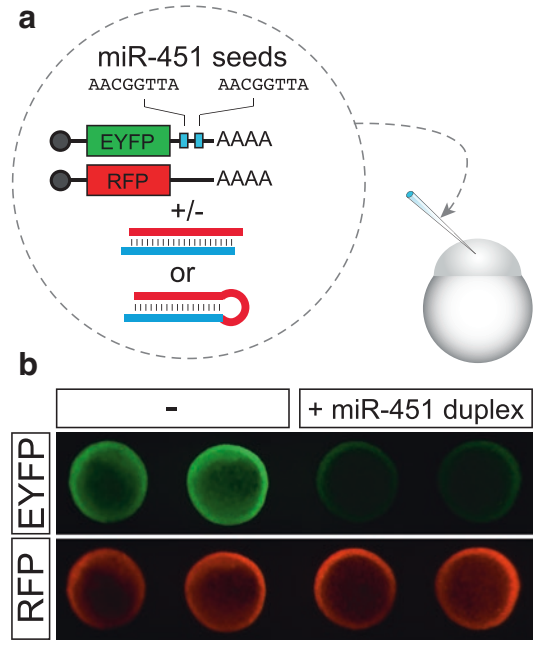


Fig. 2 In vivo microRNA sensor assay. (**a**) mRNA encoding the microRNA sensor (EYFP with two miR-451 target sites in its 3'UTR) is mixed with TagRFP mRNA that will serve as injection control to compare fluorescence between embryos. In addition, the embryos will be injected with different small RNAs to test their silencing efficiency. (**b**) Typical fluorescence result at 8 h after injection, were the duplex-injected wild type embryos have lost the fluorescence from the EYFP microRNA sensor but still show the same level of red fluorescence than the control embryos. The image shows an apical view from the animal pole of the zebrafish embryos

8. The embryos are ready for immediate imaging in a dissection scope. Alternatively, fix the embryos with 4% PFA and incubate at 4 °C for 24 h, then wash again three times with 2 mL of 1× PBS and store at 4 °C.
9. Take pictures of embryos lying flat on their backs (supine position) (Fig. 3). To achieve this position, prepare a block of 2% agarose 5 mm high and put in a 100 mm Petri dish with blue water. Make a vertical cut in the block with a blade and insert the tail of the embryo into the cut with forceps (*see* **Note 26**).
10. The morphology of the embryo makes it difficult to take the entire field in focus in a single picture. Therefore, take several pictures at different focal planes and then merge the parts in focus using a photo-editing software.

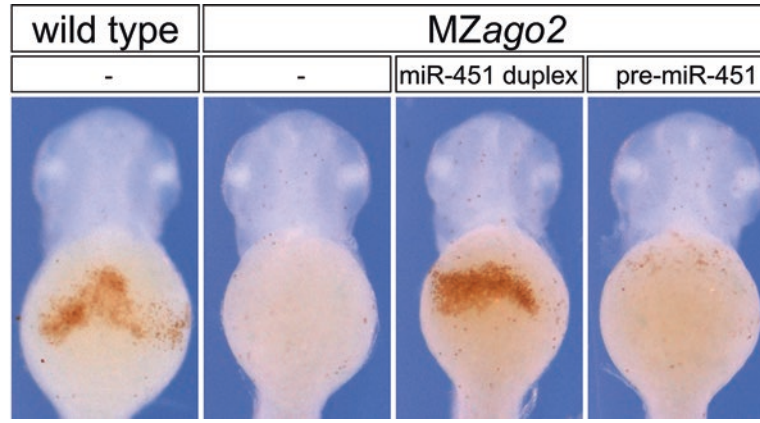


Fig. 3 *O*-dianisidine staining of erythrocytes in zebrafish embryos. The *orange-brown* staining reveals the presence of blood in wild-type but not in Ago2 mutant embryos (*see* **Note 27**). Only injection of miR-451 duplex but not pre-miR-451 hairpin (which requires Ago2 to mature) rescues erythropoiesis in Ago2 mutant embryos treated with PTU

4 Notes

1. Methylene blue inhibits fungi growth that would prey on the embryos, therefore use blue water for all embryo manipulations.
2. A thin coat will suffice, so pour the excess of agarose onto the next Petri dish or back to the original container. The agarose layer will protect the embryos without chorion that otherwise will explode in contact with the plastic ware.
3. It inhibits melanin production [23] but also creates oxidative stress that will synergize with the loss of mature miR-451 to induce anemia [24].
4. Protect the solution from light, handle with gloves.
5. Plasmid pCS2+ contains an SP6 promoter necessary for in vitro transcription and a potent SV40 polyadenylation signal necessary to instruct polyadenylation of the mRNA. EYFP is brighter and matures faster than enhanced green fluorescent protein (EGFP) [19], facilitating the downstream analysis.
6. Make sure that the linearizing enzyme does not cut EYFP or the microRNA sites. If possible, NotI is the preferred enzyme as it is an 8-base cutter. The goal is to obtain at least 1 μg of DNA in 6 μL for the in vitro transcription. Usually linearizing 5 μg in a 50 μL reaction will provide the desired concentration after the purification step. Maintain RNase-free conditions during the whole procedure. Longer incubation times may exhaust the cap nucleotide and result in non-capped, non-translatable mRNAs.

7. This red fluorescent protein will be the injection control used to normalize the fluorescence of the miR-451 sensor.
8. Prepare all the solutions in nuclease-free water.
9. Fertilized eggs will take around 40 min to conduct the first cell cycle. While it is important to maximize egg production, it is critical to perform the injections at one-cell stage. Therefore, 10 min of egg laying is enough for fish in prime condition to produce the eggs necessary for a regular experiment.
10. The strainer will not retain fish excrements and bad eggs after flushing, facilitating egg collection.
11. To avoid aspirating embryos, place the tip of the glass Pasteur pipette against the bottom corner of the 35 mm Petri dish. This results in the aspiration of the water, but not the embryos.
12. Digestion time may vary. At the point when the first embryos start to emerge from the chorion, this indicates that the chorion of the other embryos have softened and will be easy to remove.
13. Do not decant the embryos into the beaker but immerse the plate because without the protection of the chorion the embryos will explode when in contact with air. From this point, embryos must always be covered with water. The choice of a 250 mL beaker facilitates the immersion of the 35 mm Petri dish.
14. After decanting most of the water, do not return the beaker to a vertical position as the embryos will be exposed to the air and explode.
15. It is the instant pouring of water that creates the pressure necessary to squeeze the embryos out of their softened chorions.
16. The capillary used to make the injection needle must have a filament in order to lead the solution down to the tip. Needle preparation and calibration should be made prior to remove the dividers from the breeding tanks and avoid unnecessary delays during the injection process.
17. Although less than 1 μL of RNA is usually required per experiment, it is necessary to set up the micropipette to a larger volume in order to generate enough suction force to aspirate the RNA solution with these type of capillar-like disposable plastic tips.
18. Surface tension will create a dome-shaped oil drop on the slide. It is important that the drop does not flatten out, as the precise measurement of the drop diameter has to occur while the drop is falling in the oil and before it splashes against the slide surface.
19. A major source of experimental variability is the calibration of the needle. Therefore, it is preferred to inject the components independently rather than preparing a master mix for each

condition. In this way, we can compare, for example, the fluorescence levels across all embryos.

20. Manipulation and injection can be very traumatic to the embryo and a fraction of them may burst, polluting the water. It is very important to remove dead or dying embryos regularly to maintain the integrity of the other embryos.
21. The diameter of a standard 10 μL plastic tip is about the same diameter of the embryo's yolk. Connect the tip to a tube of the vacuum line in order to aspirate the portion of perforated agar at the moment of punching.
22. Manipulate the embryos into position by carefully aspirating and ejecting them with the help of a glass Pasteur pipette. Alternatively, an eyebrow hair glued to a 200 μL plastic tip will also move the embryos gently. Do not use forceps or other rigid instrument, as these will easily perforate the yolk and kill the embryo.
23. To be able to compare fluorescence levels, take a picture with all the embryos in the same field. Alternatively, when the number of embryos is large or high resolution is needed, pictures with overlapping frames using the same microscope settings and stitch them together with a photo-editing software.
24. The staining can be performed as early as 30 h post fertilization, but then slight developmental delays can be confused with differences in erythrocyte content. Therefore, it is better to stain embryos after 48 h post-fertilization.
25. A positive staining will develop as orange-brown color (Fig. 3). Add the PBS very slowly; mixing the ethanol (in the staining solution) with the water (in the buffer) will create currents that may damage the embryos.
26. The supine position expose the ducts of Cuvier, enlarged vessels on the ventral side that carry the blood into the heart and is where most of the blood of the embryo will accumulate during the staining.
27. Ago2 mutant embryos are designated as MZ*ago2*. MZ stands for *Maternal and Zygotic* and indicated that both these embryos and their mothers are homozygous mutants for the gene Ago2.

Acknowledgment

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development-NIH grant R00HD071968.

References

1. Pattanayak, V., Guilinger, J. P. & Liu, D. R. Determining the specificities of TALENs, Cas9, and other genome-editing enzymes. Methods Enzymol 546, 47–78, doi:10.1016/B978-0-12-801185-0.00003-9 (2014).
2. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. Dev Dyn 203, 253–310, doi:10.1002/aja.1002030302 (1995).
3. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell 106(1):23–34
4. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854
5. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409(6818):363–366
6. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425(6956):415–419
7. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ (2004) Processing of primary microRNAs by the microprocessor complex. Nature 432(7014):231–235. doi:10.1038/nature03049
8. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R (2004) The microprocessor complex mediates the genesis of microRNAs. Nature 432(7014):235–240. doi:10.1038/nature03120
9. Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, Wolfe SA, Giraldez AJ (2010) A novel miRNA processing pathway independent of dicer requires Argonaute2 catalytic activity. Science 328(5986):1694–1698
10. Yoda M, Cifuentes D, Izumi N, Sakaguchi Y, Suzuki T, Giraldez AJ, Tomari Y (2013) Poly(A)-specific ribonuclease mediates 3′-end trimming of Argonaute2-cleaved precursor microRNAs. Cell Rep 5(3):715–726. doi:10.1016/j.celrep.2013.09.029
11. Giraldez AJ, Cinalli RM, Glasner ME, Enright AJ, Thomson JM, Baskerville S, Hammond SM, Bartel DP, Schier AF (2005) MicroRNAs regulate brain morphogenesis in zebrafish. Science 308 (5723):833–838
12. Rasmussen KD, Simmini S, Abreu-Goodger C, Bartonicek N, Di Giacomo M, Bilbao-Cortes D, Horos R, Von Lindern M, Enright AJ, O'Carroll D The miR-144/451 locus is required for erythroid homeostasis. The Journal of experimental medicine 207 (7):1351–1358
13. Pase L, Layton JE, Kloosterman WP, Carradice D, Waterhouse PM, Lieschke GJ (2009) miR-451 regulates zebrafish erythroid maturation in vivo via its target gata2. Blood 113 (8):1794–1804
14. Dore LC, Amigo JD, Dos Santos CO, Zhang Z, Gai X, Tobias JW, Yu D, Klein AM, Dorman C, Wu W, Hardison RC, Paw BH, Weiss MJ (2008) A GATA-1-regulated microRNA locus essential for erythropoiesis. Proc Natl Acad Sci U S A 105 (9):3333–3338
15. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature 465 (7298):584–589
16. Yang JS, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. Proc Natl Acad Sci USA 107 (34):15163-15168
17. Horstick EJ, Jordan DC, Bergeron SA, Tabor KM, Serpe M, Feldman B, Burgess HA (2015) Increased functional protein expression using nucleotide sequence features enriched in highly expressed genes in zebrafish. Nucleic Acids Res 43(7):e48. doi:10.1093/nar/gkv035
18. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6(5):343–345. doi:10.1038/nmeth.1318
19. Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol 20(1):87–90. doi:10.1038/nbt0102-87
20. Shaner NC, Lin MZ, McKeown MR, Steinbach PA, Hazelwood KL, Davidson MW, Tsien RY (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins. Nat Methods 5(6):545–551. doi:10.1038/nmeth.1209
21. Iuchi I, Yamamoto M (1983) Erythropoiesis in the developing rainbow trout, Salmo gairdneri irideus: histochemical and immunochemical detection of erythropoietic organs. J Exp Zool 226(3):409–417. doi:10.1002/jez.1402260311

22. Detrich HW 3rd, Kieran MW, Chan FY, Barone LM, Yee K, Rundstadler JA, Pratt S, Ransom D, Zon LI (1995) Intraembryonic hematopoietic cell migration during vertebrate development. Proc Natl Acad Sci USA 92(23):10713–10717
23. Karlsson J, von Hofsten J, Olsson PE (2001) Generating transparent zebrafish: a refined method to improve detection of gene expression during embryonic development. Mar Biotechnol (NY) 3(6):522–527. doi:10.1007/s1012601-0053-4
24. Yu D, dos Santos CO, Zhao G, Jiang J, Amigo JD, Khandros E, Dore LC, Yao Y, D'Souza J, Zhang Z, Ghaffari S, Choi J, Friend S, Tong W, Orange JS, Paw BH, Weiss MJ (2010) miR-451 protects against erythroid oxidant stress by repressing 14-3-3zeta. Genes Dev 24(15): 1620–1633

Chapter 13

Functional Analysis of MicroRNAs in Neurogenesis During Mouse Cortical Development

Wei Zhang, Xiaoxia Zeng, and Li Zeng

Abstract

The advantage of using in utero electroporation is that it can study the gene function during neurodevelopment in vivo. Using functional analysis of a microRNA (miRNA) gene as an example, this protocol describes a set of techniques that are crucial for the success of neurogenesis studies, including mice time mating, plasmid preparation, *utero* electroporation following miRNA injection into mice embryonic brain ventricle, labeling of proliferating cells with EDU (ethynyldeoxyuridine), cryosectioning, immunofluorescence staining, and confocal microscopic analysis. This chapter also provides detailed technical tips regarding experimental planning, mouse surgery, multi-embryo injection with different plasmids, electroporation, and maintenance of pregnant mother with post-electroporated embryo.

Key words In utero electroporation, MicroRNA, Cerebral cortex, Immunohistochemistry, Neural proliferation, Neurogenesis, Cortical development

1 Introduction

It has been more than a decade since in utero electroporation was first developed to introduce DNA into mouse embryos [1, 2]. This application has opened a new window to look at neural development in mammals. Briefly, DNA is injected into the open ventricles of the developing brain and the application of an electrical current causes the formation of transient pores in cell membranes, allowing for the uptake of DNA into the cells. Not only it is capable of labeling the developing cells by expressing a fluorescent protein (e.g., GFP), but also it can manipulate genes of interest by overexpressing or even loss-of-function of a gene by expressing the shRNAs [3]. Due to its ability to temporally and spatially target a specific population of neurons in the brain, in utero electroporation has been most widely used for the analysis of brain development of neocortex, embryonic retina and thalamus, hypothalamus, cerebellar Purkinje cells, auditory brainstem, and hindbrain [4]. The convenience of this technology led to a large body of studies

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_13,

on the gene functions and early neural development such as neurogenesis and neuronal migration [5, 6]. More recently, it has also been used to address late-phase developmental questions such as dendritic patterning and axonal path-finding [7, 8]. In addition, due to the feasibility of animal models via in utero electroporation, it caught the attention of physiologists and behaviorists, such as long-term expression of a light gated protein ChR2 channels by in utero electroporation enabled the analysis of the network connectivity and neurons involved in learning in adult animals [9, 10]. Manent and others [11] reported that expression of DCX ameliorated the susceptibility to epileptic seizures in DCX knockdown animal. These results suggest in utero electroporation can have an impact not only on a molecular and cellular level, but also at circuit and behavioral levels even after developmental periods.

MicroRNAs (miRNAs) are small noncoding RNA molecules that function in the transcriptional and post-transcriptional regulation of gene expression in a variety of organisms. miRNAs exert their function by targeting mRNA molecules, typically triggering their degradation or translational repression by the RNA-induced silencing complex (RISC) [12]. The functional role of miRNAs in neuronal development has recently received increasing attention; an increasing number of studies on dysregulation of miRNAs in neurodegenerative disease also suggested that alterations in the miRNA regulatory pathway could contribute to the disease pathogenesis [13]. However, the specific miRNAs that regulate the proliferation and differentiation of neural progenitor cells (NPCs) during early cortical development are not well established; molecular mechanisms underlying the pathologic implications of dysregulated miRNA expression and regulation involved in neurodegenerative diseases remain largely unknown as well.

In this protocol, we introduced miRNAs into mouse utero via in utero electroporation and analyzed the function of miRNAs during mouse early cortical development. Specifically, we investigated miRNA's function in regulating proliferation and differentiation of NPCs.

2 Materials

2.1 In Utero Electroporation

1. Pregnant mice: Time-mate C57BL/6 mice to obtain the pregnant mice. Put one male and one female mouse into the same cage. The next morning if a vaginal plug is found on the female, then it is designated as the embryonic day 0.5 (E0.5). All experiments are to be performed in accordance with the approved protocols set by the institutional animal care and use committee.
2. 2.0–5.0 μg/μl plasmid: prepared with Qiagen EndoFree Maxi prep Kit. Fast-green (final concentration at 0.1%) should be added prior to injection to label the injection position.

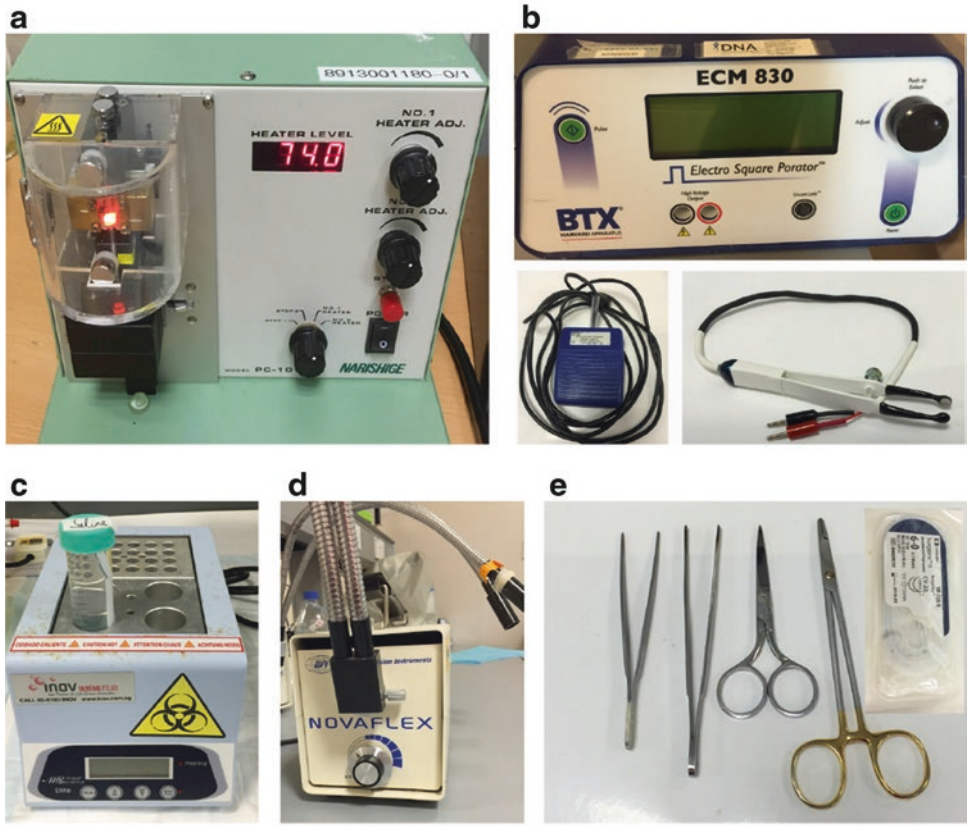


Fig. 1 Machines and tools for surgery. (**a**) PC-10 Puller. (**b**) Electroporator, foot-controlled pedal and electrode. (**c**) Heater. (**d**) Novaflex model FO-150 lighter. (**e**) Surgical tools and sutures

3. Micropipettes for plasmid injection (Fig. 1a, *see* **Notes 1** and **2**).
4. Electroporator: ECM 830 (BTX), Electrode (7 mm), and a foot-controlled pedal (Fig. 1b).
5. Heater (Fig. 1c).
6. 10% Povidone-iodine.
7. 70% ethanol.
8. Lighter (World Precision Instruments, Novaflex model FO-150) (Fig. 1d).
9. Two blunt forceps (Fig. 1e).
10. Scissors (Fig. 1e).
11. Needle holder (Fig. 1e).
12. Sterilized gauze.
13. Saline.
14. Sutures.
15. Analgesic drugs: Carprofen, 5 mg/kg.

2.2 miRNA Functional Analysis in Neural Proliferation and Differentiation

We show how to study the function of miRNA (miR-128) on neural progenitor cells (NPCs)'s proliferation and differentiation during cortical development by in utero electroporation [14, 15].

1. Pregnant mice for miRNA functional analysis: C57BL/6 mice are time-mated as above (Subheading 2.1, **item 1**), and choose the E13.5 pregnant mice for in utero electroporation.
2. miRNA/cGFP dual-expression construct with CMV promoter driving expression of miR-128 and with Ef1a driving expression of cGFP (Fig. 2a). Prepare the plasmids as per above (Subheading 2.1, **item 2**).
3. Stereomicroscope (Zeiss, Stemi 2000, Fig. 2b).
4. Cryosection machine (Leica, CM 1950, Fig. 2c).
5. Liquid blocker pen.
6. Confocal Laser Scanning Biological Microscope (Olympus FV1000, Fig. 2d).
7. Click-iT®-Plus EDU Alexa Fluor®647 Kit (Invitrogen).
8. Mouse anti-NEUN (#MAB 377, Millipore).
9. Anti-Mouse-Alexa Fluor 555 (Invitrogen).
10. Dapi (Sigma-Aldrich).
11. Mounting medium (Vector).
12. 4% paraformaldehyde (PFA).
13. PBS.
14. 30% Sucrose in PBS.

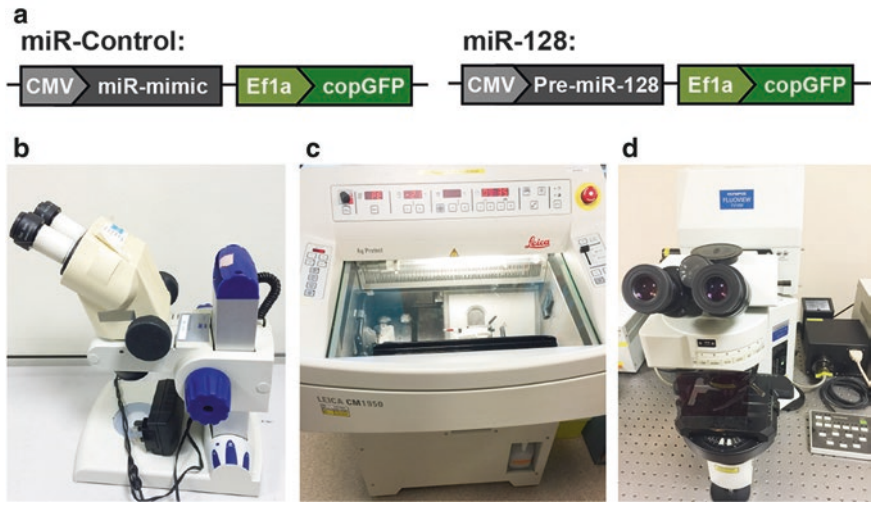


Fig. 2 Tools. (**a**) The expression construct of miR-Control and miR-128. (**b**) Stereomicroscope. (**c**) Cryosection machine. (**d**) Confocal Fluorescence microscope

3 Methods

3.1 In Utero Electroporation

1. Anesthesia of mouse: Administer a cocktail of ketamine (100 mg/kg) and xylazine (10 mg/kg) by intraperitoneal injection (i.p.). The anesthesia time will last for about 2 h.
2. Surgical skin preparation: After making sure that the animal is fully anesthetized by toe pinch reflex disappearance, disinfect the abdominal skin.
3. Exposure of the embryos: Cut a 2–2.5 cm length incision along the midline of the abdominal skin with scissors (Fig. 3a, *see* **Note 3**). Cover the skin with sterilized gauze with an open window (3–4 cm) in the middle. Cut the muscle wall about 2–2.5 cm length along the midline underneath the skin (Fig. 3b). Drop some saline on the gauze around the open place. Pull out one side of the embryos gently and place on the gauze (Fig. 3c, *see* **Note 4**), drop some warm saline on embryos from time to time to keep them moist (Fig. 3d, *see* **Note 5**).
4. Prepare glass pipette and filling of plasmid solution: Pinch the glass pipette tip with forceps, fill the glass pipette with about 2–3 μl plasmid (mixed with fast-green) using a 5 ml syringe connected with glass pipette through a plastic hose (Fig. 4a, b).
5. Microinjection of plasmids into embryo brain ventricles: Find an embryo whose head is visible, fix the embryo head position with a blunt forceps and push one side of the head against the uterus (Fig. 5a). Make a penetration through the uterus and skull between the eye and lambdoid suture (Fig. 5b), into the lateral ventricle on one side of the brain. Inject plasmid together with fast-green into the ventricle (Fig. 5c).
6. Electroporation: Drop some saline on electrode surfaces and the uterus wall; clasp the positive electrode on the injected uterus side, the negative electrode against opposite side (Fig. 6a). Keep

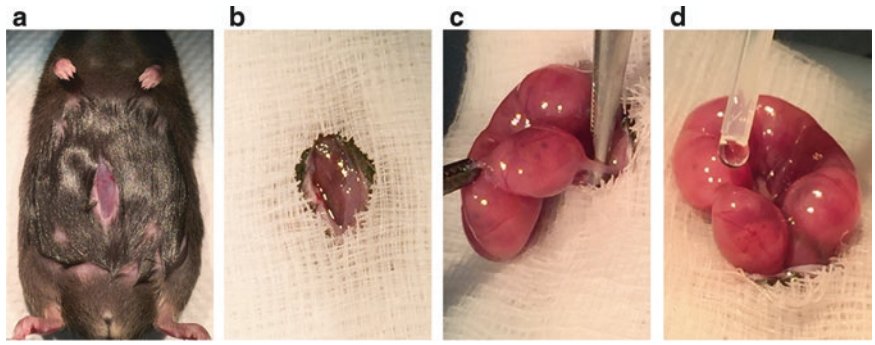


Fig. 3 Surgery and exposure of embryos. (**a**) Cut skin. (**b**) Cut the muscle. (**c**) Pull out one side embryos. (**d**) Drop saline to keep embryo moisture

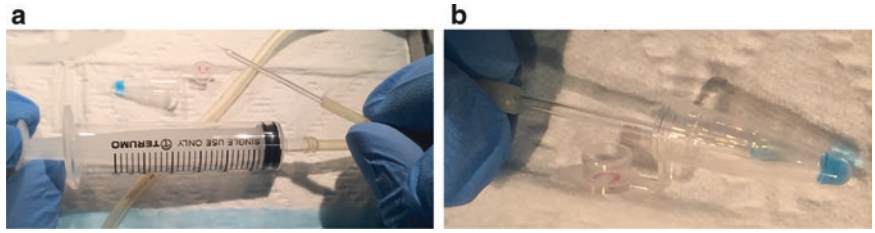


Fig. 4 Uptake plasmid. (**a**) Connect glass pipette to the syringe with a plastic hose. (**b**) Fill the glass pipette with plasmid

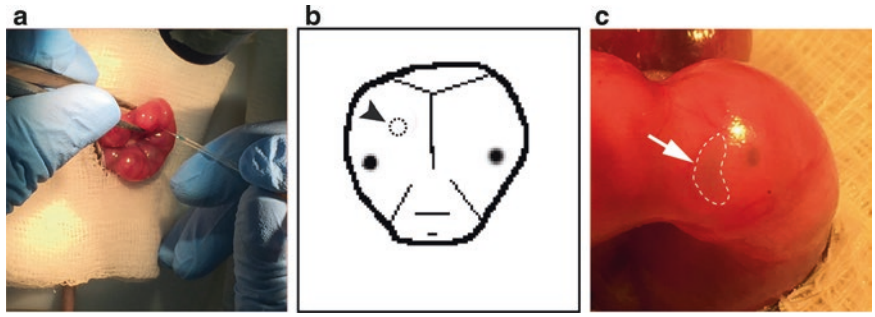


Fig. 5 Microinjection. (**a**) Hold the embryo. (**b**) Microinjection place (*dotted circle*). (**c**) Arrow shows the plasmid in the lateral ventricle

the embryonic head between two electrodes. Keep dropping some saline on the electrodes and uterus; make sure the contact surface between electrodes and uterus is filled with saline (Fig. 6b). Deliver the electric pulses with a foot-controlled pedal. During electroporation, stretching responses from the embryo can be observed. Drop saline on embryos to keep them moist after electroporation (Fig. 6c).

7. Return embryos back into the abdomens: Drop some saline on the abdominal cavity and all embryos before putting them back. Using one blunt forceps to lift the abdominal muscle and skin, and concurrently, use another blunt forceps to insert embryo body one by one back into abdominal cavity gently (Fig. 7a). Drop some saline into the abdominal cavity and then pull another side of embryos out gently. Repeat above microinjection and electroporation steps. After all the embryos have been put back into the abdominal cavity, drop some saline into the cavity to let the embryos be positioned more naturally.
8. Suturing of the abdominal wall: Suture the muscle wall first (Fig. 7b) then suture the skin (Fig. 7c, *see* **Note 6**).
9. Recovery from anesthesia and post-surgery: Put the mouse back into the cage and cover with one gauze to keep the mouse

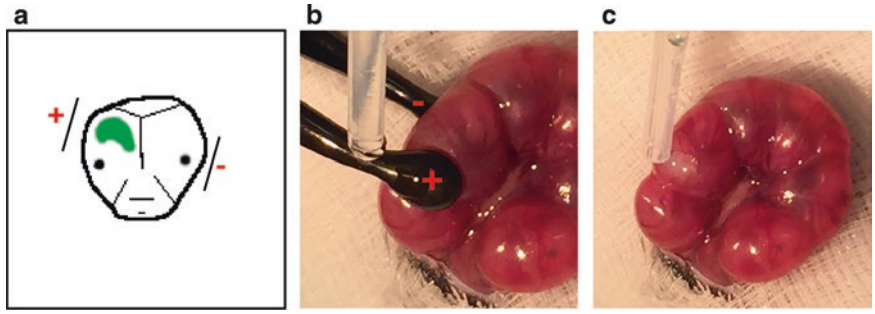


Fig. 6 Electroporation. (**a**) Electrodes direction. (**b**) Electroporation. (**c**) Drop saline on the embryo post-electroporation

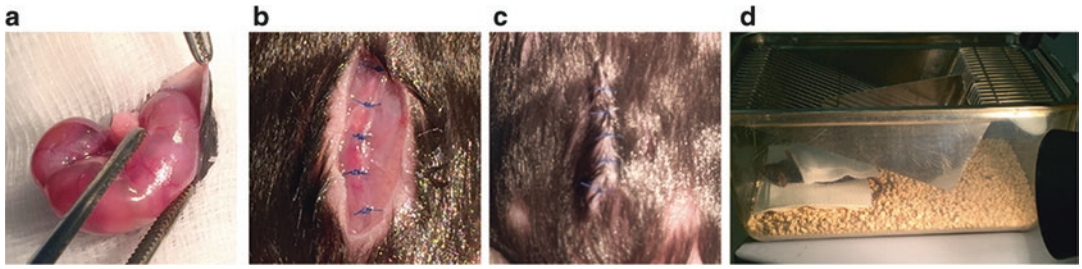


Fig. 7 Post-surgery. (**a**) Put the embryos back into abdominal cavity. (**b**) Suture the muscle wall. (**c**) Suture the skin. (**d**) Keep the mouse warm until it wakes up

warm with warm lamp until it wakes up (Fig. 7d, *see* **Note 7**), then put back into the animal room. Monitor the mouse closely during first 2 days after surgery. Provide analgesic drugs daily to the mouse during first 3 days after surgery.

3.2 miRNA Functional Analysis in Neural Proliferation

3.2.1 EDU Labeling

1. Pick the E13.5 pregnant mice to conduct in utero electroporation as above Subheading 3.1. Twenty three hours after in utero electroporation, inject 100 μl EDU (10 mM) into the pregnant mouse by i.p., Embryo to be collected 1 h later (*see* **Notes 8** and **9**).
2. Embryo samples collection. At E14.5, terminate the pregnant mouse with CO_2. Cut the abdominal skin and muscle, pull out the embryos and cut off from the pregnant mouse. Separate miR-Control and miR-128 injected embryos and place them on different dishes with cold PBS. Leave them on ice for brain tissue collection. Put one embryo in a 6 cm dish with cold PBS, cut off the head and separate out brain from the skull under stereomicroscope, and then put in one 1.5 ml tube to fix with 4% PFA.
3. Embryo samples fixation. The embryonic brain samples should be fixed for at least 6–8 h in 4% PFA at 4 °C. 24–48 h after fixation then change into 30% sucrose in PBS. Store brain sample at 4 °C until it sinks down at the bottom of the tube.

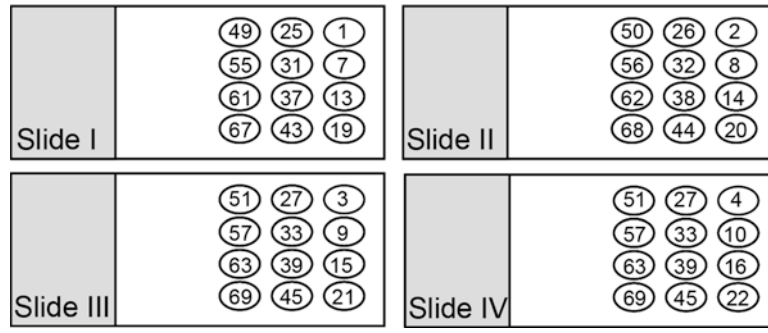


Fig. 8 Attach the cryosectioned brain slices onto the glass slide

4. Cryosections. After brain samples have sunk down in 30% sucrose, cut the coronal brain section with cryosection machine (Fig. 2c). The thickness is 16 μm. Every six sections of brain slice should be attached onto the same glass slide (Fig. 8). After brain slices were dried up at the room temperature, store the slices at 4 °C for short-term storage and −20 °C for long-term storage.

3.2.2 Immunohistochemistry Staining of EDU

Immunohistochemistry staining of EDU following the kit staining protocol, as summarized below:

1. Take out one glass slide from 4 °C or −20 °C. Dry it at room temperature for at least 30 min.
2. Circle the brain tissue with the liquid blocker pen.
3. Wash twice with PBS.
4. Wash twice with PBS containing 3% BSA.
5. Add 200–300 μl of 0.5% Triton®X-100 in PBS and incubate for 20 min at room temperature for permeabilization.
6. Remove permeabilization buffer and wash twice with PBS containing 3% BSA.
7. Add 200–300 μl of Click-iT®Plus reaction cocktail on the tissue.
8. Incubate for 30 min at room temperature (avoid the light).
9. Remove the reaction cocktail and wash once with PBS containing 3% BSA.
10. Wash with PBS twice and then mount glass coverslip with mounting medium.

3.2.3 Cell Counting and Results Analysis

1. Observe the EDU stained brain slices under Olympus FV1000 confocal microscope. Take pictures of all brain slices which contain GFP positive cells under the high magnification (40×).

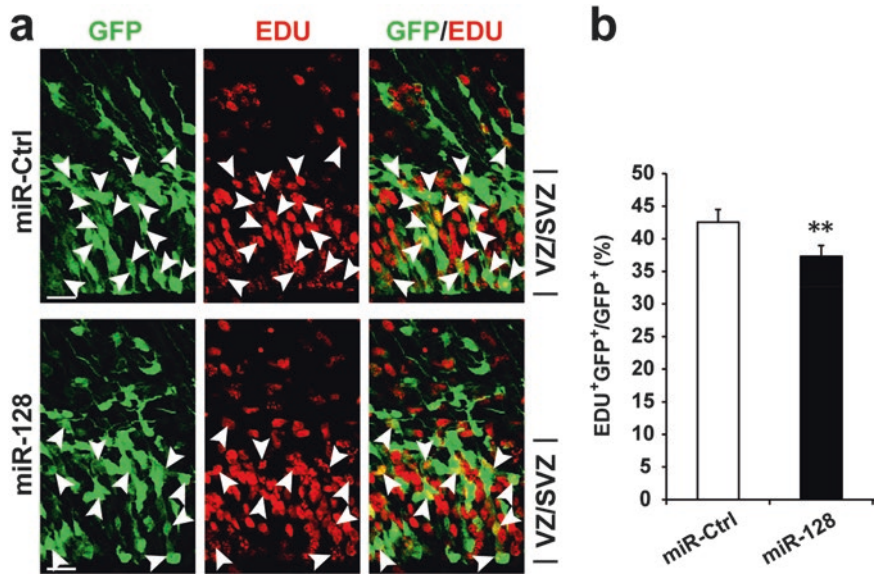


Fig. 9 miR-128 regulates NPCs proliferation in mouse cortical development by in utero electroporation. miR-128 and miR-control were introduced into mouse uterus at E13.5 via in utero electroporation. At E14.5, mouse was labeled with EDU by i.p., for 1 h before brain tissue was collected. (**a**) EDU immunohistochemistry staining was performed, arrow heads (*white*) indicate GFP and EDU double-positive cells. Scale bar = 20 μm. (**b**) Analysis of counting results by Student's *t* test. $^{**}p < 0.01$. All data shown are means with standard deviation (mean ± SD)

2. Count all GFP positive cells in one brain slice first, followed by counting GFP and EDU double-positive cells in the same brain slice, and then calculate the EDU positive ratio as below:

   EDU positive ratio (%) = (Number of GFP and EDU double-positive cells/Number of GFP positive cells) × 100.

   Normally in one embryo, there are about 8–10 brain slices containing GFP-positive cells. So that in each experiment condition (miR-Control and miR-128), there are about 8–10 groups of EDU positive ratio. The average of the ratios derived from the 8–10 groups of GFP positive cells are considered as the EDU positive ratio representing the embryo.
3. Perform Student's *t* test using values of at least three embryos from each experiment condition (Fig. 9).

3.3 miRNA Functional Analysis in Neural Differentiation

3.3.1 Embryo Samples Collection, Fixation, and Cryosection

1. We chose E13.5 pregnant mice to conduct in utero electroporation as above Subheading 3.1.
2. Collect embryo brain samples at E18.5 (*see* **Note 10**). Fixation and cryosection of brain tissue will follow the same procedure as above (Subheadings 3.2.1–3.2.3).

3.3.2 Immunohistochemistry Staining of NEUN

The method is summarized as below:

1. Take out one glass slide from 4 °C or −20 °C. Dry in room temperature for at least 30 min.
2. Circle the tissue with liquid blocker pen.
3. Wash twice in PBS.
4. Block the samples with 1% BSA + 0.1% Triton®X-100 in PBS at room temperature for 30 min.
5. Add 200–300 μl of primary antibody (anti-NEUN, mouse source, 1:200 dilutions in PBS containing 1% BSA and 0.1% Triton®X-100) at 4 °C overnight.
6. Remove primary antibody and wash three times with PBS, 5 min each time.
7. Add 200–300 μl of secondary antibody (anti-Mouse-Alexa Fluor 555, 1:200 dilution in PBS) for 45 min at room temperature (avoid the light).
8. Remove secondary antibody and wash three times with PBS, 5 min each time.
9. Mount glass coverslip with mounting medium.

3.3.3 Cell Counting and Results Analysis

1. Observe above NEUN stained brain slices under Olympus FV1000 confocal microscope. Take pictures of all brain slices which have GFP positive cells in the CP Zone under the high magnification (40×).
2. Count all GFP positive cells in one brain slice first, followed by counting GFP and NEUN double-positive cells in the same brain slice, and then calculate the NEUN positive ratio as below:
   NEUN positive ratio (%) = GFP and EDU numbers/GFP numbers × 100%
3. Follow the above Subheading 3.2.3 for counting results analysis, and the results are shown in Fig. 10.

4 Notes

In utero electroporation require meticulous experimental planning, like timed mating to achieve pregnant mice, surgery and taking out designated timed embryo, multi-embryo injection of miRNA and control plasmids into embryo brain ventricle, electroporation of embryo head and maintenance of pregnant mother with post-electroporated embryo. Delicate hands are needed for experimenter to dissect out embryo, handling of embryo during ventricle injection and electroporation, and putting back embryo into mother abdomen; preparation of fine glass micropipette tip are also very critical for the injection of miRNA and survival of embryo after utero electroporation in mother uterus. Below are specific notes:

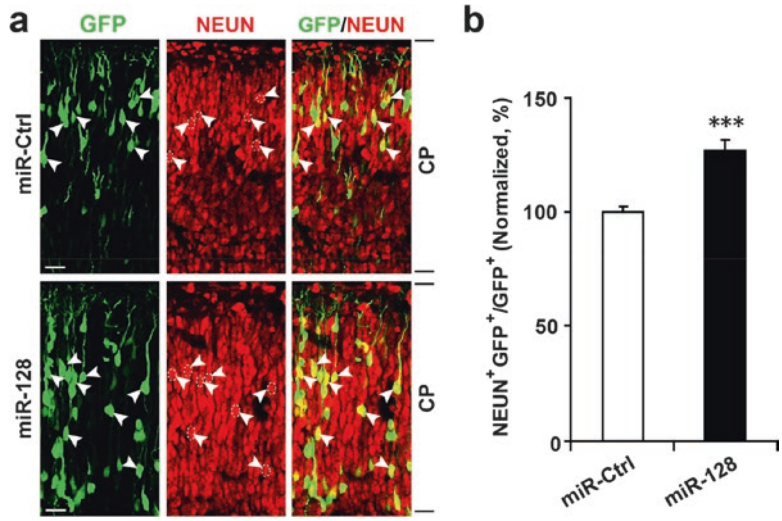


Fig. 10 miR-128 regulates NPCs differentiation in mouse cortical development by in utero electroporation. miR-128 and miR-control were introduced into mouse uterus at E13.5 via in utero electroporation and collected brain tissues at E18.5. (**a**) NEUN immunohistochemistry staining was performed; arrow heads (*white*) indicate GFP and NEUN double-positive cells. Scale bar = 20 μm. (**b**) Analysis of counting results (normalized to miR-Ctrl) by Student's *t* test, ***$p < 0.001$. All data shown are means with standard deviation (mean ± SD)

1. Capillary glass tubing with flame polished ends is used to make glass capillaries. OD = 1.5 mm, ID = 0.86 mm, Length = 7.5 cm. NARISHIGE (PC-10 model) machine (Fig. 1a), set as: NO. 2 heater, heater level 74.0 °C.
2. The capillary glass tip should be pinched before using; the diameter of the top is about 8–10 μm. Too big is not good for embryo survival, too small will be difficult to fill with plasmids. If the glass pipette is broken or felt hard to penetrate, it requires changing a new one. The glass pipette should be under UV for at least 15 min before using.
3. The abdominal incision should not be too big or too small. Too big of an incision is not good for mouse recovery, but too small may squeeze the embryos and affect their survival.
4. After pulling out one side of embryos, the last one from the uterine vagina region has a lace connected with the abdominal wall. If it is required to inject different plasmids, it can start to record from this one, for example, R-1 (right side NO.1 embryo).
5. During the operation, embryos should be kept moist by dropping warm saline on them, especially before microinjection. After sending all embryos back into the abdominal cavity, continue dropping enough warm saline. It also helps embryos to be moved back to their natural position.

6. The muscle and skin must suture separately.
7. The total surgery time should be as short as possible, it's better to be less than 1 h.
8. Although both EDU and BRDU (Bromodeoxyuridine) can label the proliferating cells, it's better to use EDU if the plasmids are tagged with GFP or other fluorescent tags. This is due to that BRDU staining requires HCl treatment and this treatment will be quenching the GFP fluorescence signal intensity.
9. For proliferation study by using EDU labeling, the labeling time is about 1–3 h. If too long, the labeled cells will migrate out of VZ/SVZ zone.
10. When conducting the neural differentiation study, collect embryos on E17.5 or E18.5 before the pregnant mother gives birth.

Acknowledgments

We thank Dr. Paul Kim and Ms. Neo Yan Zhuang for their technical support. This research was supported by the Singapore Ministry of Health's National Medical Research Council grant to W. Z and Translational Clinical Research grant to L.Z.

Reference

1. Saito T, Nakatsuji N (2001) Efficient gene transfer into the embryonic mouse brain using in vivo electroporation. Dev Biol 240(1): 237–246. doi:10.1006/dbio.2001.0439. S0012-1606(01)90439-7 [pii]
2. Tabata H, Nakajima K (2001) Efficient in utero gene transfer system to the developing mouse brain using electroporation: visualization of neuronal migration in the developing cortex. Neuroscience 103(4):865–872. doi:S0306452201000161 [pii]
3. Matsuda T, Cepko CL (2007) Controlled expression of transgenes introduced by in vivo electroporation. Proc Natl Acad Sci U S A 104(3):1027–1032. doi:10.1073/pnas.0610155104. 0610155104 [pii]
4. David LS, Aitoubah J, Lesperance LS, Wang LY (2014) Gene delivery in mouse auditory brainstem and hindbrain using in utero electroporation. Mol Brain 7:51. doi:10.1186/s13041-014-0051-4. s13041-014-0051-4 [pii]
5. Chen G, Sima J, Jin M, Wang KY, Xue XJ, Zheng W, Ding YQ, Yuan XB (2008) Semaphorin-3A guides radial migration of cortical neurons during development. Nat Neurosci 11(1):36–44. doi:10.1038/nn2018. nn2018 [pii]
6. Sanada K, Tsai LH (2005) G protein beta-gamma subunits and AGS3 control spindle orientation and asymmetric cell fate of cerebral cortical progenitors. Cell 122(1):119–131. doi:10.1016/j.cell.2005.05.009. S0092-8674(05)00454-X [pii]
7. Cancedda L, Fiumelli H, Chen K, Poo MM (2007) Excitatory GABA action is essential for morphological maturation of cortical neurons in vivo. J Neurosci 27(19):5224–5235. doi:10.1523/JNEUROSCI.5169-06.2007. 27/19/5224 [pii]
8. Wang CL, Zhang L, Zhou Y, Zhou J, Yang XJ, Duan SM, Xiong ZQ, Ding YQ (2007) Activity-dependent development of callosal projections in the somatosensory cortex. J Neurosci 27(42):11334–11342. doi:10.1523/JNEUROSCI.3380-07.2007. 27/42/11334 [pii]
9. Huber D, Petreanu L, Ghitani N, Ranade S, Hromadka T, Mainen Z, Svoboda K (2008) Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice.

Nature 451(7174):61–64. doi:10.1038/nature06445. nature06445 [pii]

10. Petreanu L, Huber D, Sobczyk A, Svoboda K (2007) Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. Nat Neurosci 10(5):663–668. doi:10.1038/nn1891. nn1891 [pii]
11. Manent JB, Wang Y, Chang Y, Paramasivam M, LoTurco JJ (2009) Dcx reexpression reduces subcortical band heterotopia and seizure threshold in an animal model of neuronal migration disorder. Nat Med 15(1):84–90. doi:10.1038/nm.1897. nm.1897 [pii]
12. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297. doi:S0092867404000455 [pii]
13. Kawahara H, Imai T, Okano H (2012) MicroRNAs in neural stem cells and neurogenesis. Front Neurosci 6:30. doi:10.3389/fnins.2012.00030
14. Zhang W, Kim PJ, Chen Z, Lokman H, Qiu L, Zhang K, Rozen SG, Tan EK, Je HS, Zeng L (2016) MiRNA-128 regulates the proliferation and neurogenesis of neural precursors by targeting PCM1 in the developing cortex. Elife 5. doi:10.7554/eLife.11324
15. Zhang W, Thevapriya S, Kim PJ, Yu WP, Shawn Je H, King Tan E, Zeng L (2014) Amyloid precursor protein regulates neurogenesis by antagonizing miR-574-5p in the developing cerebral cortex. Nat Commun 5:3330. doi:10.1038/ncomms4330. ncomms4330 [pii]

Chapter 14

Cellular Approaches in Investigating Argonaute2-Dependent RNA Silencing

Cai Zhang, Joonbae Seo, and Takahisa Nakamura

Abstract

In mammals, there are four Argonaute (Ago) family proteins that play crucial roles in RNA silencing, a process wherein microRNA (miRNA) mediates inhibition of target mRNA translation. Among the Ago proteins, Argonaute2 (Ago2) uniquely possesses an endoribonuclease (slicer) activity that is critical for the biogenesis of specific miRNAs and mRNA cleavage. This Ago2 slicer activity is required for postnatal development. Despite its important roles, there are still gaps in our understanding of the mechanistic basis of Ago2's unique functions in vivo due to a limited availability of experimental tools. In order to investigate Ago2's functions, we generated a new cellular model of Ago2-deficiency in 3T3 mouse embryonic fibroblasts (MEFs). This cell line can be used for investigating general Ago2 functions, but also for further understanding of Ago2's unique characteristics including the slicer activity, specific amino acid residues, and domains in Ago2 by reconstitution of Ago2 mutants. Here, we describe the methods for establishing Ago2-deficient MEFs and for reconstituting the MEFs with an Ago2 mutant lacking its slicer activity by means of a retrovirus-mediated gene transfer.

Key words Ago2, 3T3 protocol, Mouse embryonic fibroblast

1 Introduction

Argonaute (Ago) proteins are essential components of RNA-induced silencing complex (RISC) and play a central role in RNA silencing processes [1]. Among four Ago proteins in mammals, Ago2 is regarded as a special member since only this family member retains endoribonuclease (slicer) activity that generates specific microRNAs (miRNAs) and cleaves target mRNAs [2–6]. Ago2-deficient mice are embryonic lethal and mice lacking Ago2's slicer activity die shortly after birth with anemia, indicating essential roles of Ago2- and its slicer activity-dependent RNA silencing [2, 6]. In addition to these developmental regulations, it has become clear that Ago2 is involved in the regulation of a variety of cellular and physiological events including tumorigenesis, hypoxia, bone marrow maturation, and pancreatic β cell expansion [7–10].

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_14,

Previous reports demonstrated that there are several amino acid residues critical for Ago2 functions; slicer activity (D669), RNA recognition (e.g., K533, Q545, K570), epidermal growth factor signaling activity (Y393), and Akt-mediated RNA silencing regulation (S388) [8, 11, 12]. Furthermore, there are several domains critical for Ago2 function including PAZ, an RNA binding module that recognizes the 3′ end of miRNA, or PIWI, a protein domain found in piwi proteins and a large number of related nucleic acid-binding proteins [13]. To investigate the roles of amino acid residues and domains of Ago2 in vivo is of biological interest and critical physiologic importance, as a deeper understanding of their functions could provide us a novel mechanistic basis of RNA silencing and a design of improved therapeutics targeting miRNA-mediated events.

Mouse embryonic fibroblasts (MEFs) have been broadly used as an experimental tool for cellular- and molecular-based biology. Primary fibroblasts isolated from mouse embryos have limited capacity for proliferation. However, the murine fibroblasts can be immortalized by distinct procedures, so that it allows us to easily enable the desired genetic manipulations, reduce culture maintenance time and effort, and facilitate biological replicates in experiments without concern for the limitation of cell proliferation. There are basically two different ways to immortalize the primary mouse fibroblasts. One is to introduce an oncogene(s), such as SV40 large T antigen, into the primary fibroblasts in order to induce malignant transformation. However, the oncogene could disrupt proper cellular processes/events and the MEFs immortalized by the oncogene could manifest biological characteristics as cancer cells. Another method establishing immortalized MEFs is by means of the "3T3" protocol, which means that primary MEFs are transferred (the "T") every 3 days (the first "3"), and inoculated at the rigid density of 3×10^5 cells per 6 cm dish (the second "3") continuously (Fig. 1a) [14]. Upon successive passages with the defined 3T3 protocol, MEFs initially undergo decent doublings with each passage, and then go through a senescence crisis where the growth of MEFs virtually ceases (Fig. 1b). By continuing the 3T3 protocol even in the crisis phase, MEFs eventually show a sign of the immortalization with a constant or rising growth rate. There is also a potential issue that the MEFs immortalized by 3T3 protocol (3T3 MEFs) could spontaneously harbor a mutation(s) on endogenous oncogenes/tumor suppressor genes during overcoming the crisis phase. Nonetheless, 3T3 MEFs are regarded as one of the most conventional and useful cellular tools. Fundamental cellular and molecular mechanisms revealed by analysis of the 3T3 MEFs have been widely applied to other cell types in vitro and in vivo.

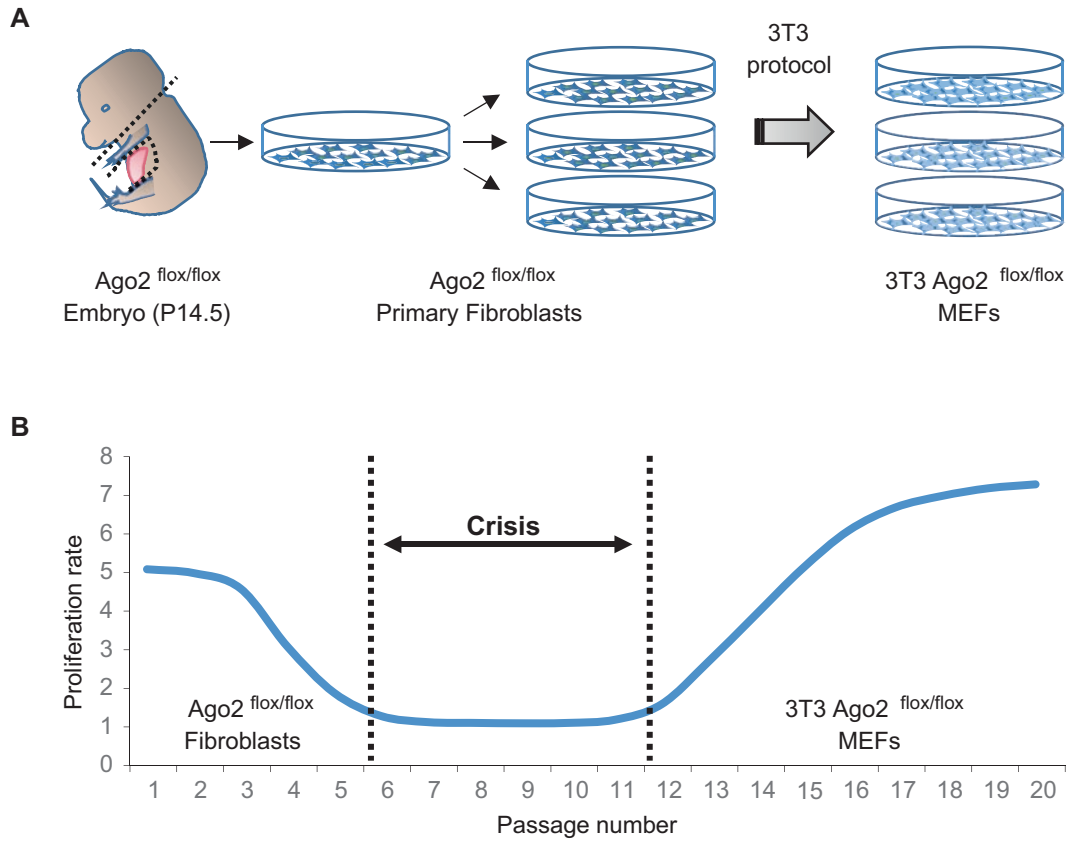


Fig. 1 Schemes of establishment of 3T3 MEFs. (**a**) A scheme of isolating mouse primary fibroblasts to generate immortalized MEFs through the 3T3 protocol. (**b**) A scheme showing the proliferation rate of MEFs during an immortalization process through 3T3 protocol. The cells go through a senescence crisis once and then become immortalized

In this chapter, we describe an establishment of 3T3 MEFs generated from an embryo carrying a floxed allele that enables conditional inactivation of Ago2 by utilizing adenovirus-mediated *cre* expression, and reconstitution of Ago2-deficient 3T3 MEFs with an Ago2 mutant by means of retrovirus-mediated gene transfer.

2 Materials

2.1 Immortalization of Ago2$^{flox/flox}$ MEFs by 3T3 Protocol

1. B6.129P2(129S4)-*Ago2$^{tm1.1Tara}$*/J (Ago2$^{flox/flox}$; Jackson Laboratory).
2. Culture medium: Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Hyclone) and 1× penicillin-streptomycin solution (Life Technologies).

3. 0.25% trypsin-EDTA (Life Technologies).
4. Phosphate-buffered saline (PBS) (Life Technologies).
5. Sterile Fine scissors and Fine forceps.
6. Cell/Tissue culture incubator with 5% CO_2 at 37 °C.
7. Biological safety cabinet.
8. Cell/Tissue culture supplies: Plastic tissue culture dishes (10-, 6-cm); sterile plastic serological pipettes; sterile disposable tubes (50 ml) cryovials.
9. BAMBANKER (WAKO).

2.2 Genetic Deletion of Ago2 by Adeno-Cre

1. Pre-packed Cre recombinase adenovirus (Vector Biolabs).
2. 293A cells (Invitrogen).
3. −80 °C deep freezer.
4. 37 °C water bath.
5. Anti-Mouse Ago2 Monoclonal Antibody (WAKO).

2.3 Genetic Reconstitution of Human Ago2 by Retrovirus

1. pMXs-Puro retroviral vector (Cell Biolabs).
2. PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies).
3. QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies).
4. DNA purification kit (Machery-Nagel).
5. BamHI (NEB), EcoRI (NEB).
6. DNA ligation kit (Takara).
7. Platinum-E (Plat-E) retroviral packaging cell line (Cell Biolabs).
8. Hexadimethrine bromide (Polybrene; Sigma-Aldrich).
9. Puromycin (Invitrogen).
10. pFRT/FLAG/HA-DEST EIF2C2 (Addgene).
11. Top10 competent cells (Invitrogen).
12. LB medium with/without ampicillin: dissolve 25 g of LB broth powder (Fisher) in about 800 ml of purified water. Heat the solution and frequently agitate until it is completely dissolved. Adjust the pH of the medium to 7.2 using 1 M NaOH and bring volume up to 1 L with purified water. Before use, sterilize by autoclaving for 15 min. If antibiotic is needed, allow the solution to cool to 55 °C or lower, and add ampicillin to the final concentration of 50 μg/ml. Store the solution at room temperature.
13. LB Agar plates with ampicillin: Suspend 40 g LB agar powder (Fisher) in 1 L of purified water. Heat the solution and

frequently agitate until it is completely dissolved. Before use, sterilize by autoclaving for 15 min. Allow the solution to cool to 55 °C, add ampicillin to the final concentration of 50 μg/ml. Mix the solution thoroughly, and pour about 25 ml in each Petri dish. Store the Petri dishes at 4 °C.

14. Plasmid miniprep kit (Machery-Nagel).
15. Lipofectamine 2000 (Invitrogen).
16. Opti-MEM (Invitrogen).
17. Sterilized 0.45-μm filter.
18. Hemagglutinin (HA) antibody (Cell Signaling).

3 Method

3.1 Isolate and Immortalize MEF Cells from Ago2$^{flox/flox}$ Mice

1. Set up breeding pairs with Ago2$^{flox/flox}$ males and females and check females the following day for vaginal plugs to determine if mating has occurred (*see* **Note 1**).
2. At 14.5 days of pregnancy, anesthetize the pregnant Ago2$^{flox/flox}$ female and place it on a dissecting board. Sterilize the fur of the mouse with 70% ethanol. Make a cut on the medioventral line with scissors, and expose the abdominal wall of the mouse.
3. Cut through the abdominal wall with sterile scissors. Lift up the uterine horns with sterile forceps (*see* **Note 2**), and cut off the uterus and place it in tissue culture dish containing PBS for wash.
4. Carefully dissect the uterus in the PBS, separate the embryos and transfer them to a new dish with PBS. Swirl the dish to remove blood from the embryos.
5. Draw an embryo and place it on an empty tissue culture dish. Remove the embryonic head, and tear out the red tissue (heart and liver). Place the rest of the embryo in the covered Petri dish containing 3 ml of 0.25% trypsin-EDTA.
6. Chop up the trimmed embryo with scissors repeatedly to mince the tissue into pieces as small as possible. Pipet the tissue pieces gently several times with a 3 ml syringe loaded with a 18G needle, then place the dish in the 37 °C tissue culture incubator for 5 min.
7. Remove the dish from the incubator, add 7 ml of DMEM culture medium to the dish in order to inactivate trypsin, and mix the cell suspension gently but thoroughly using a 10 ml serological pipet.
8. Transfer the cell suspension to a 15 ml conical centrifuge tube. By using an additional 3 ml of culture medium, wash the dish and transfer remaining cells to the tube.

9. Centrifuge the tube at 200 rcf for 5 min.
10. Carefully remove and discard approximately 10 ml of supernatant. Add 10 ml of fresh culture medium, and mix the cell suspension gently. Let the cell suspension sit for about 5 min to allow larger embryo fragments to sink to the bottom.
11. Transfer the supernatant, but not large embryo fragments, as much as possible to a 10 cm dish. Incubate the dish with 5% CO_2 at 37 °C overnight and change the culture medium the following morning.
12. Check the cells later in the day. If the cells reach confluency, make cell stocks with BAMBANKER or split the cells. At this time, the passage number of the Ago2$^{flox/flox}$ MEFs is "1."
13. To establish 3T3 Ago2$^{flox/flox}$ MEFs, seed the cells at the density of 3×10^5 cells per 6 cm dish, and split the cells with the same density every 3 days. The MEF cells will become immortalized after *about 15–20 times of passage*. Make cell stocks *in the* cryovial with an appropriate label including a passage number.

3.2 Genetic Deletion of Ago2 by Adeno-Cre

Expression levels of Ago2 in the 3T3 Ago2$^{flox/flox}$ MEFs are intact. Therefore, the 3T3 Ago2$^{flox/flox}$ MEFs can be used as wild-type. In order to generate Ago2-deficient 3T3 MEFs, a *cre* recombinase needs to be expressed in the cells (Fig. 2a). In this section, the *cre* gene is transferred into the cells by utilizing an adenovirus vector (*see* **Note 3**).

1. The 293A cell line, a subclone of the 293 cell line, is used to amplify a replication-incompetent adenovirus. Trypsinize and count the 293A cells, plate them to two 10 cm dishes with 3 × 106 cells per dish. On the following morning, refresh the medium and add a small amount (10–20 μl) of the pre-packed Adeno-Cre virus or a control adenovirus carrying green fluorescent protein (Adeno-GFP) to each dish. Swirl the dish gently to mix.
2. Keep culturing the cells in the incubator. When about 50% of 293A cells are detaching (about 1 day after transduction), squirt the cells off the dish with a 10 ml pipette. Transfer virus-containing cells and media to 15 ml sterile capped tubes.
3. Place the 15 ml tube containing harvested cells and media at −80 °C for at least 30 min. Remove tube and place in a 37 °C water bath for no longer than 15 min to thaw. Repeat the freezing and thawing steps twice. Centrifuge the cell lysate at 800 rcf for 10 min. Transfer the supernatant to cryovials in 1 ml aliquots. Store the viral stocks at −80 °C. Titer the virus stock if needed. This virus stock, the 1st amplification of the Adeno-Cre and Adeno-GFP virus, can be used for cell infection to

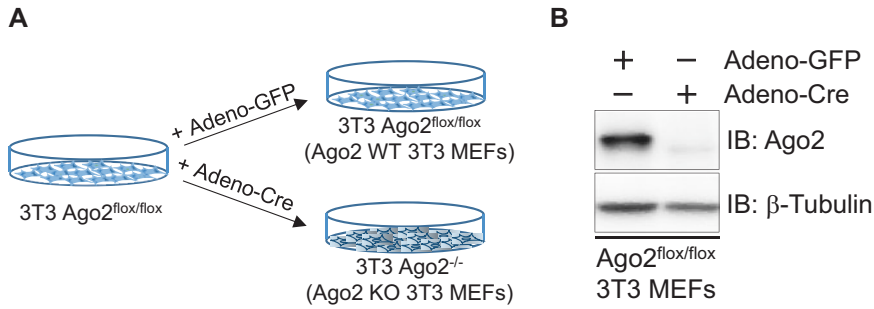


Fig. 2 Generation of Ago2-deficient 3T3 MEFs. (**a**) A scheme showing a strategy to delete Ago2 in 3T3 Ago2$^{flox/flox}$ MEFs by adenovirus-mediated cre recombinase (Adeno-Cre) expression. (**b**) A western blot analysis assessing expression levels Ago2 detected by an anti-Ago2 monoclonal antibody in MEFs treated with Adeno-GFP and Adeno-Cre. β-Tubulin is shown as a control

express transgene, further crude amplification, or large amplification and purification for in vivo study.

4. In order to infect Adeno-Cre and Adeno-GFP virus to 3T3 Ago2$^{flox/flox}$ MEFs, plate the cells to two 6 cm dishes at 5 x 105 cells per dish (Day 0). On the following morning (Day 1), refresh the medium and add approximately 100 μl of the crude viral stocks of Adeno-Cre or Adeno-GFP from **step 3** (*see* **Note 4**). Refresh the medium on the following morning (Day 2). Continue to culture the cells in the incubator.
5. When the infected 3T3 Ago2$^{flox/flox}$ MEFs reach confluency (1–2 days after transduction), plate the infected MEF cells in multiple 6 cm dishes at 5 × 10^5 cells per dish.
6. By using the cells on one of the plates, examine expression levels of Ago2 by western blot by detecting with anti-Ago2 antibody (Fig. 2b). If the titer of virus stocks is high enough, expression of Ago2 is completely lost in the 3T3 Ago2$^{flox/flox}$ MEFs infected with Adeno-Cre, but not in cells infected with Adeno-GFP, even after one-time virus infection.
7. If Ago2 expression remains in the 3T3 Ago2$^{flox/flox}$ MEFs infected with Adeno-Cre, repeat the infection step (**step 4**). Expression of Ago2 usually disappears within three times of Adeno-Cre infection. Genetic deletion of the Ago2 allele can be confirmed by genotyping using a specific primer set. After confirmation of Ago2 deletion, make stocks of the Ago2-deficient 3T3 MEFs (Ago2 KO 3T3 MEFs) and Adeno-GFP-treated 3T3 Ago2$^{flox/flox}$ MEFs (Ago2 WT 3T3 MEFs) (*see* **Note 5**).

3.3 Genetic Reconstitution of Human Ago2 by Retrovirus

Functional roles of specific amino acids and domains in Ago2 can be investigated in vivo by reintroducing Ago2 mutants in Ago2 KO 3T3 MEFs. In order to reintroduce the Ago2 mutant, one of the most conventional and convenient methods is to utilize a

retrovirus-mediated gene transfer that mediates stable genetic modification of treated cells by chromosomal integration of a gene(s) of interest. In this section, Ago2 KO 3T3 MEFs are reconstituted with human Ago2 WT or mutant by means of the retrovirus (Fig. 3a). Through this strategy, we can even replace Ago2 with other Argonaute members in MEFs (Fig. 3a).

Among the variety of retrovirus vectors, a pMXs-Puro retroviral vector based on the Moloney murine leukemia virus (MMLV) is used in this study. This pMXs-Puro vector can stably integrate and express a gene of interest with a puromycin selection. In order to insert human Ago2 into this vector, human Ago2 coding DNA sequence (CDS) is amplified by PCR with primers that contain a restriction enzyme digestion site, followed by the restriction enzyme digestion and a ligation of the human Ago2 CDS fragment with the pMXs-Puro backbone vector. Retrovirus vector with human Ago2 mutant can also be generated with a mutagenesis kit if necessary. After confirmation of the human Ago2 WT or mutant insertion, the retrovirus vector is ready for transfection to retrovirus packaging cells such as Plat-E cells to produce retrovirus, followed by infection to Ago2 KO 3T3 MEFs for reconstitution.

1. Generate a BamHI-human Ago2-EcoRI nucleotide fragment by PCR using PfuUltra II Fusion HS DNA Polymerase with an appropriate human Ago2 template (e.g., pFRT/FLAG/HA-DEST EIF2C2) (*see* **Note 6**). Digest pMXs-Puro retroviral vector and the purified human Ago2 DNA fragments with BamHI and EcoRI. Purify the digested human Ago2 fragments with a DNA purification kit, and the digested backbone vector by gel purification (*see* **Note 7**).

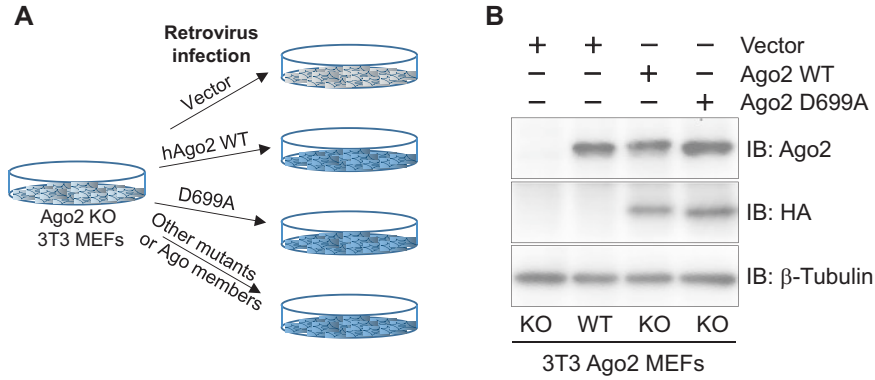


Fig. 3 Reconstitution of human Ago2 in Ago2-deficient 3T3 MEFs. (**a**) A scheme showing a strategy to reconstitute human Ago2 WT, mutant, or other Ago members by means of retrovirus-mediated gene transfer. (**b**) A western blot analysis assessing expression levels Ago2 detected by anti-Ago2 that detect both endogenous and reconstituted Ago2 and anti-HA monoclonal antibody that detects exogenous proteins. β-Tubulin is shown as a control

2. Ligate the purified human Ago2 and backbone fragments (follow manufacturer's protocol in the ligation kit). Transform the ligation product to Top10 competent cells and spread the cells on a LB agar plate containing 50 μg/ml ampicillin. Select 5–10 colonies from the plate after a 37 °C overnight incubation. Culture them overnight in LB medium with 50 μg/ml ampicillin, and isolate plasmid from each colony using a miniprep kit.
3. If needed, induce a mutation(s) into human Ago2 by QuikChange Lightning Site-Directed Mutagenesis Kit (follow manufacturer's protocol). Sequence the plasmids (use the PCR primers in **Note 6**) to confirm Ago2 insertion.
4. After confirmation, prepare enough plasmid for transfection (*see* **Note 8**), and seed Plat-E cells at 3×10^6 cells in a 6-cm dish (Day 0).
5. On Day 1, transfect 10 μg of pMXs-Puro retroviral empty vector, pMXs- human Ago2-Puro retroviral vector, or pMXs-human Ago2 mutant-Puro retroviral vector to the Plat-E cells in each 6-cm dish by Lipofectamine 2000.
6. On Day 2, refresh the culture medium and continue to incubate the cells in the incubator.
7. After 24 hours, harvest the virus-containing supernatant from the dish of transfected Plat-E cells and filtrate it with a 0.45-μm sterile syringe filter. Aliquot 1 ml of the supernatant into cryovials. This supernatant can be used for gene transfer immediately or stored at −80 °C. Titer the virus stock if needed.
8. Seed Ago2 KO and Ago2 WT 3T3 MEFs at 3×10^4 cells per well in a 24-well plate.
9. On the following morning, infect the MEFs with the different dilutions of retroviral supernatant in the presence of 6 μg/ml polybrene for 8–24 h in order to obtain cells expressing human Ago2 WT and mutant at similar levels (*see* **Note 9**). Remove and discard the medium from the wells and replace with fresh DMEM culture medium.
10. After additional 24 h incubation, trypsinize the infected MEFs and plate them in 6-cm dishes with fresh culture medium containing 2 μg/ml puromycin (*see* **Note 10**) for selection.
11. Continue to incubate the cells until the selection is completed (about 3 days). Harvest the selected cells and confirm the expression of human Ago2 WT or mutant protein in Ago2 KO MEF cells with anti-Ago2 antibody by western blot (Fig. 3b). If necessary, compare the expression levels of different groups of cells with Ago2 WT 3T3 MEF cells, and choose the cells expressing similar Ago2 levels as WT cells. Make

stocks of the Ago2 KO 3T3 MEFs reconstituted with human Ago2 WT or mutant, and pMXs-Puro retrovirus treated Ago2 KO 3T3 MEFs.

4 Notes

1. Please refer Information about "how to time mouse pregnancy" described in the website of the Jackson Laboratory.
2. Do not contaminate the uterus with the fur.
3. Every procedure in this part contains infectious virus. Please follow the Biological Safety Level 2 work guidelines.
4. Passaging and transduction can be conducted at the same time. Add the crude virus lysate to medium in the 6 cm dish when seeding the cells on Day 0. Swirl the dish gently to mix. Refresh the medium on the morning of Day 1.
5. Expression of genes induced by adenovirus is regulated in a temporal manner.
6. The primers for the PCR are 5′-CGATAGGGA TCC ACCATGTACTCGGGAGCCGGC-3′ (forward) and 5′-CGATAGGGAATTCCCTAGGCAGAGTCT GGGACGTCATATGGATAAGCAAAGTA CATGGTGCG-3′ (reverse). HA tag is introduced to the sequence as an optional marker.
7. BamHI and EcoRI locate separately in a multi-cloning site of the pMXs-Puro retroviral vector and there is about 1 kb between these two sites. So after digestion, two fragments with different sizes are observed in the gel. One of them is a fragment around 1 kb and the other one is about 5–6 kb, which is the backbone fragment.
8. If mini-prep is used, ethanol precipitation is needed for purification of the plasmid to ensure the transfection efficiency.
9. If it is necessary to adjust the expression level of restored human Ago2 WT similar to the endogenous Ago2 expression level in WT cells, you can transduce the Ago2 KO MEF cells with different titers of the retrovirus, such as 1, 10, 25, 50, or 100 μl of the crude virus to 500 μl of medium.
10. Ideally, optimized puromycin concentration for Ago2 MEF should be tested before selection by culture the Ago2 KO MEF cells in the presence of different concentration of puromycin, such as 1, 2, 4, 6 μg/ml and so on. Choose the lowest concentration of puromycin that can induce the death of more than 95% of cells, for the following selection experiment. It is always recommended to have Ago2 KO MEF cells (without transduction) as the positive control for selection.

Acknowledgments

We thank Kazutoshi Murakami, Elise Bernhard, and Vishnupriya Borra for discussions. This work was supported by NIH RO1 DK107530, Digestive Disease Research Core Center in Cincinnati (DK078392), and PRESTO from the Japan Science and Technology Agency.

References

1. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell 123(4):631–640. doi:10.1016/j.cell.2005.10.022
2. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature 465(7298):584–589. doi:10.1038/nature09092
3. Yang JS, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC (2010) Conserved vertebrate mir-451 provides a platform for dicer-independent, Ago2-mediated microRNA biogenesis. Proc Natl Acad Sci U S A 107(34):15163–15168. doi:10.1073/pnas.1006432107
4. O'Carroll D, Mecklenbrauker I, Das PP, Santana A, Koenig U, Enright AJ, Miska EA, Tarakhovsky A (2007) A slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. Genes Dev 21(16):1999–2004. doi:10.1101/gad.1565607
5. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
6. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305(5689):1437–1441. doi:10.1126/science.1102513
7. Cheng N, Li Y, Han ZG (2013) Argonaute2 promotes tumor metastasis by way of up-regulating focal adhesion kinase expression in hepatocellular carcinoma. Hepatology 57(5):1906–1918. doi:10.1002/hep.26202
8. Shen J, Xia W, Khotskaya YB, Huo L, Nakanishi K, Lim SO, Du Y, Wang Y, Chang WC, Chen CH, Hsu JL, Wu Y, Lam YC, James BP, Liu X, Liu CG, Patel DJ, Hung MC (2013) EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. Nature 497(7449):383–387. doi:10.1038/nature12080
9. Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, Wolfe SA, Giraldez AJ (2010) A novel miRNA processing pathway independent of dicer requires Argonaute2 catalytic activity. Science 328(5986):1694–1698. doi:10.1126/science.1190809
10. Tattikota SG, Rathjen T, McAnulty SJ, Wessels HH, Akerman I, van de Bunt M, Hausser J, Esguerra JL, Musahl A, Pandey AK, You X, Chen W, Herrera PL, Johnson PR, O'Carroll D, Eliasson L, Zavolan M, Gloyn AL, Ferrer J, Shalom-Feuerstein R, Aberdam D, Poy MN (2014) Argonaute2 mediates compensatory expansion of the pancreatic beta cell. Cell Metab 19(1):122–134. doi:10.1016/j.cmet.2013.11.015
11. Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ (2005) Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. Nature 434(7033):666–670. doi:10.1038/nature03514
12. Horman SR, Janas MM, Litterst C, Wang B, MacRae IJ, Sever MJ, Morrissey DV, Graves P, Luo B, Umesalma S, Qi HH, Miraglia LJ, Novina CD, Orth AP (2013) Akt-mediated phosphorylation of argonaute 2 downregulates cleavage and upregulates translational repression of MicroRNA targets. Mol Cell 50(3):356–367. doi:10.1016/j.molcel.2013.03.015
13. Jonas S, Izaurralde E (2015) Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet 16(7):421–433. doi:10.1038/nrg3965
14. Todaro GJ, Green H (1963) Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol 17:299–313

Chapter 15

Genomic Tagging of *AGO1* Using CRISPR/Cas9-Mediated Homologous Recombination

Sanjay Ghosh and Ji-Long Liu

Abstract

Tagging of genes at the endogenous loci is a powerful strategy for the analysis of protein function. We have developed a homologous recombination-based approach for inserting epitope tag into *Drosophila AGO1* locus by employing the CRISPR/Cas9 technology. The methodology involves co-expression of sgRNA (containing 20-nucleotide *AGO1* targeting sequence) and Cas9 protein, together with a donor template that has HA-AGO1 cassette flanked by sequences homologous to the *AGO1* locus. The integration is efficient and readily monitored by immunostaining of the transgenic cell line. This method facilitates rapid generation of stable cell lines and allows insertion of any tag sequence into endogenous loci, thus accelerating characterization of the tagged proteins.

Key words CRISPR/Cas9, *Drosophila*, AGO1, Homologous recombination, Epitope tagging, Genome engineering, Transfection vector

1 Introduction

DNA strand breaks are repaired by homologous recombination (HR) or non-homologous end joining repair (NHEJ) pathway within cells [1]. The homology-directed repair process is dependent on the presence of a donor DNA template with sequence homology flanking the site of lesion. This allows insertion of exogenous sequences at or near the site of repair and has been successfully co-opted for genome engineering [2]. Epitope tagging of genes has emerged as a powerful strategy to study the function of encoded proteins and has been widely used for analysis of protein function [3]. Compared with transgenic expression of recombinant proteins, targeted insertion of epitope tags into the endogenous chromosomal loci maintains the regulatory landscape of the gene and allows easy standardization for multiple analytical methods. However, the specificity and efficiency HR-mediated tagging is extremely low making the procedure laborious and demanding.

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_15,

Here, we describe the use of CRISPR/Cas9 method for seamless integration of HA-epitope sequence into the Argonaute-1 (*AGO1*) locus in *Drosophila* cells. The CRISPR/Cas9 system generates a specific targeted DNA double-strand break and consists of two components–a single guide RNA (sgRNA) with a 20 nucleotide targeting sequence at its 5′end and Cas9 nuclease that cleaves the DNA upon formation of the sgRNA-DNA complex in context of a 3 nucleotide protospacer adjacent motif (PAM) sequence [4, 5]. This RNA-guided endonuclease system is independent of host factors and has been widely used for gene editing and functional genomics studies across animal kingdom including *Drosophila* [6–12]. *Drosophila AGO1* binds microRNAs to form the miRNA-induced silencing complex (miRISC) that suppress gene expression by modulating mRNA stability and/or translation [13, 14]. For epitope tagging of AGO1, we targeted the exon common to all transcript variants and chose the sgRNA targeting sequence that showed high DNA cleavage efficiency [15]. The 20 nucleotide targeting sequence is cloned using a pair of complementary oligonucleotides containing overhangs for insertion into the *Drosophila* expression vector developed in our lab that allows constitutive co-expression of sgRNA and Cas9 protein thus reducing variability in transfection experiments (Fig. 1). The donor plasmid is generated in two steps. First, the homology arms (left and right of the cleavage site) are amplified from the genomic DNA using a gene-specific and a chimeric primer that contains the HA sequence. In the second step, both DNA fragments are joined together by overlap extension PCR [16] followed by subcloning into a standard cloning vector (Fig. 2). The sgRNA-Cas9 expression plasmid and the donor DNA are then combined for transfecting *Drosophila* S2R+ cells leading to DNA cleavage at the *AGO1* target site and followed by homologous recombination resulting in insertion of the HA tag in frame with *AGO1* coding sequence (Fig. 2). The transfected cells are subjected to immunostaining to detect expression of the HA-tagged AGO1 protein (Fig. 3). Furthermore, stable cell lines can be established by puromycin selection and genomic DNA sequenced to verify insertion of HA epitope.

2 Materials

Use autoclaved ultrapure water to prepare all solutions. All plastic wares (tips and tubes), containers for reagent storage and media for culture of bacterial cells should be sterilized by autoclaving before use.

2.1 Generation of AGO1-sgRNA-Cas9 Expression Plasmid

1. Plasmid DNA: pAc-sgRNA-Cas9 (Addgene plasmid #49330).
2. Oligonucleotide primers (*see* Table 1).

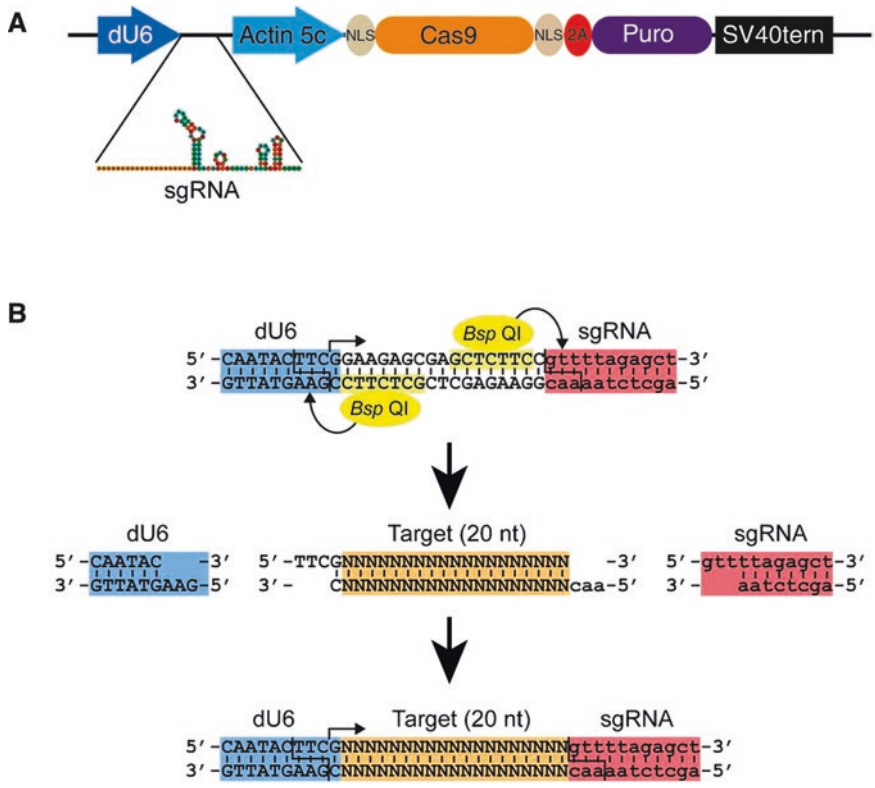


Fig. 1 CRISPR/Cas9 expression system for *Drosophila* cell lines (**a**) Schematic diagram showing the essential features of expression vector pAc-sgRNA-Cas9. The sgRNA cassette is under the control of *Drosophila* U6 snRNA promoter and includes a 20 nucleotide spacer sequence at the 5′ end that can be replaced with gene-specific targeting sequence by BspQ1 digestion. The human codon-optimized Cas9 coding sequence flanked by nuclear localization signal (NLS) is produced from actin5C promoter which also drives expression of the selectable marker puromycin *N*-acetyltransferase gene. The Cas9 and puromycin coding sequences are separated by "self-cleaving" T2A peptide that allows bicistronic expression. Transcription is terminated by polyadenylation signal from the SV40 virus. (**b**) Strategy for cloning the 20 nucleotide targeting sequence into the expression vector. BspQ1 digestion of the vector results in 5′ overhangs. A pair of 20 nucleotide target sequence containing oligonucleotides are designed such that upon annealing the ends provide the sequence complimentary to the vector overhangs allowing seamless cloning. This figure is adapted from [15]

3. Oligonucleotide annealing buffer (2×): 20 mM Tris pH 8.0, 2 mM EDTA pH 8.0, 100 mM NaCl.
4. QIAprep Spin Miniprep Kit (Qiagen).
5. QIAquick PCR purification kit (Qiagen).
6. BspQ1 (New England Biolabs).
7. Calf Intestinal Alkaline Phosphatase (New England Biolabs).
8. T4 polynucleotidyl kinase (New England Biolabs).
9. T4 DNA ligase (New England Biolabs).

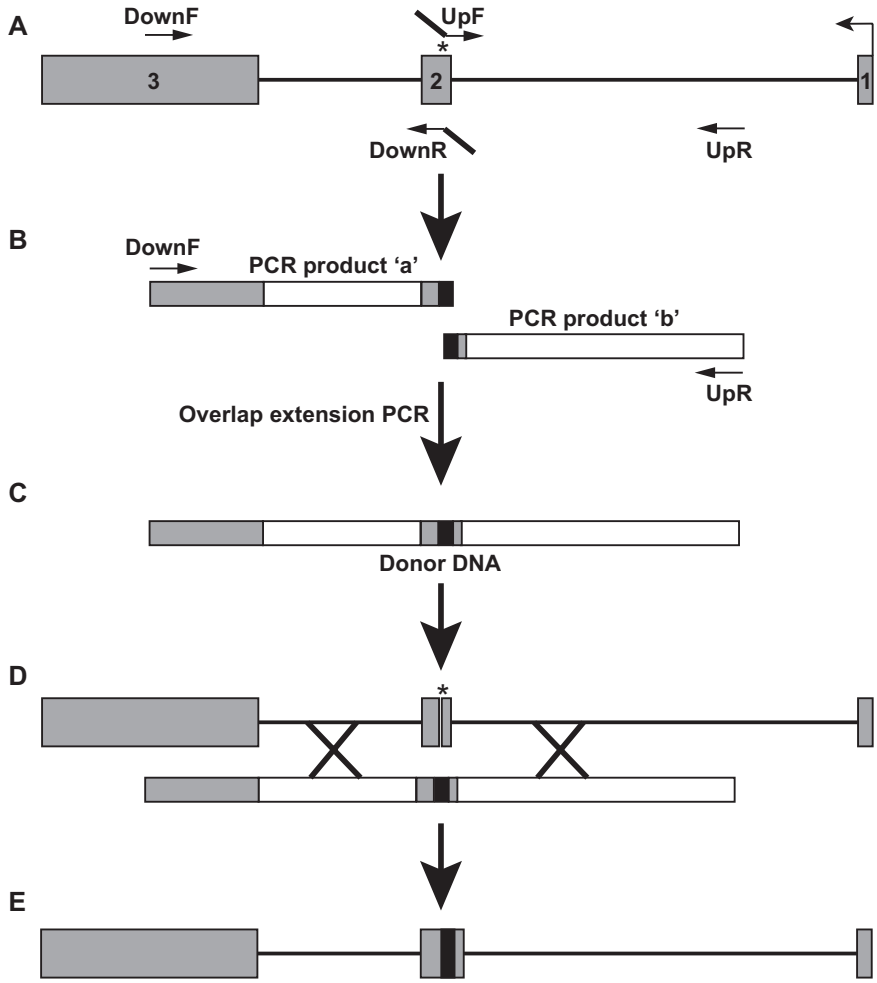


Fig. 2 Strategy for generation of donor plasmid and insertion of HA epitope by homologous recombination. (**a**) Schematic of *AGO1* locus showing a partial genomic region from exon 1 to exon 3. The region flanking the 20 nucleotide target sequence (marked by *) is PCR amplified using a pair of gene-specific and chimeric oligonucleotide containing the HA sequence (*black*, *bold*). (**b**) Both DNA fragments are used as template and joined together by overlap extension PCR using the gene specific primers. (**c**) The donor DNA contains HA sequence (*black box*) fused in-frame with the *AGO1* exon 2 sequence flanked by ~1 kb of homology arms. (**d**) Co-transfection of *AGO1*-sgRNA-Cas9 expression plasmid and donor DNA vector resulting in cleavage of DNA at the target site and subsequent integration of the HA tag (*black box*) in *AGO1* locus by homologous recombination (**e**). The transcription start site of *AGO1* is shown by an arrow while the exons are in *gray boxes*. The interconnecting line and the *white boxes* depict intron sequence. The sequences of the primers are shown in Table 1

10. Chemically competent *E. coli* DH5α cells (Thermo Fisher Scientific) (*see* **Note 1**).
11. 100 mg/ml Carbenicillin, disodium salt (Sigma) (*see* **Note 2**).
12. LB-Carbenicillin broth: Dissolve 10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 l deionized water and sterilize by autoclaving.

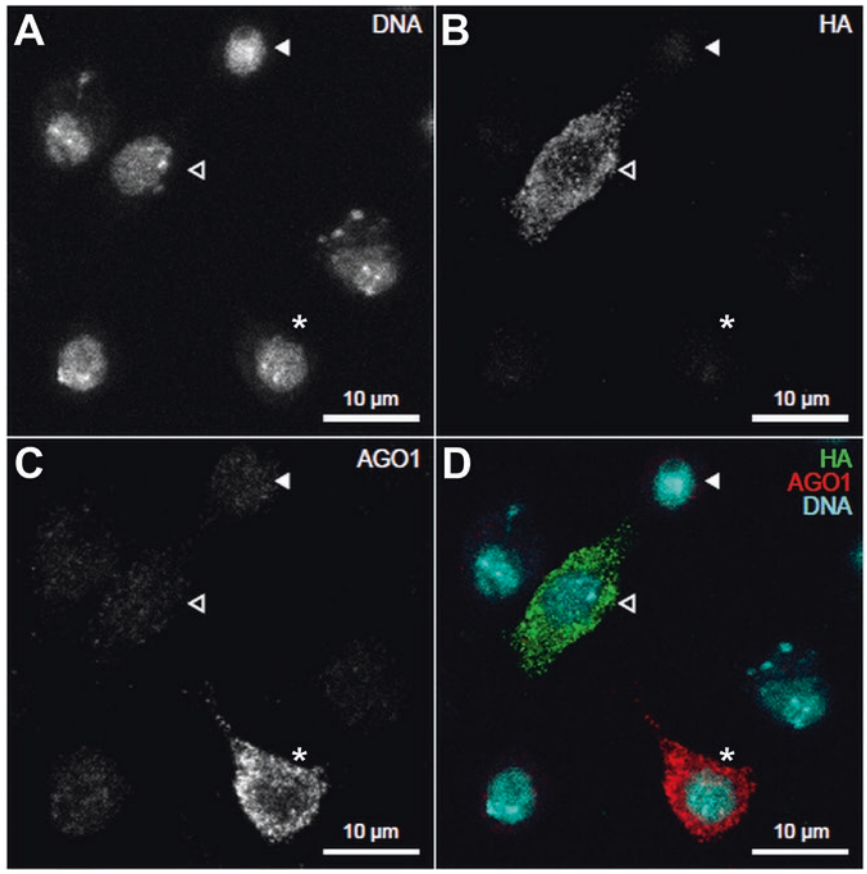


Fig. 3 Tagging of AGO1 gene with HA peptide. Cells co-transfected with expression vector and donor plasmid are co-stained with anti-HA (**b**, **d**, *green*) and anti-AGO1 antibody (**c**, **d**, *red*). The merged image is shown in (**d**). The epitope-tagged cell (*open arrowhead*) shows a cytoplasmic distribution of anti-HA signal similar to anti-AGO1 staining in an untransfected cell (*asterisk*). The absence of anti-HA and anti-AGO1 staining in cells (*closed arrowhead*) suggests a loss of AGO1 expression due to CRISPR/Cas9-mediated homozygous mutation. DNA stained with Hoechst is shown in cyan, (**a**). Scale bar 10 μm. This figure is modified from [15]

Add Carbenicillin (100 μg/ml final concentration) to the cooled media (~37 °C) and store at 4 °C.

13. LB-Carbenicillin agar: Add 15 g Agar to 1 l of LB media as described above and sterilize by autoclaving. Cool the media to ~40 °C, add Carbenicillin (100 μg/ml final concentration) and pour into petri dishes. Store at 4 °C.
14. S.O.C. medium (Thermo Fisher Scientific).
15. Nanodrop spectrophotometer.
16. Dry block heater.
17. Sterile 1.5 ml microcentrifuge tubes.

Table 1
List of oligonucleotides

Oligonucleotide	Sequence (5′ > 3′)
AGOsgF	TTCGCTGAGGCCGAGTGGGCGAGG
AGOsgR	AACCCTCGCCCACTCGGCCTCAGC
U63F	CATTTGCCAATTCTTATAATTC
AgoUpF	ACTCTGGCCGGCGTAGTCGGGCACGTCGTAGGGG-TATGACTCTATAAAAAGAAAAGGTAATC
AgoUpR	GGCCTGCGTCTCGAATTTATTC
AgoDownF	CTCGTCCGAGATTCGGGCG
AgoDownR	TACCCCTACGACGTGCCCGACTACGCCGGCCAGAGTCAGTGGAC-CCCCTCGCCC

2.2 Generation of AGO1-HA Donor Plasmid

2.2.1 Preparation of Genomic DNA from S2R+ Cells

1. *Drosophila* S2R+ cells (Drosophila Genomics Resource Centre, stock #150).
2. 75-cm^2 cell culture flask (Sigma).
3. Schneider's *Drosophila* medium (Sigma).
4. Fetal bovine serum (Thermo Fisher Scientific).
5. Penicillin-streptomycin (10,000 U/ml, Thermo Fisher Scientific).
6. Lysis buffer: 10 mM NaCl, 10 mM Tris pH 7.5, 10 mM EDTA pH 8.0, 0.5% SDS.
7. Proteinase K solution (20 mg/ml, Thermo Fisher Scientific).
8. Tris-equilibrated phenol:chloroform:isoamyl alcohol (Sigma).
9. 3 M sodium acetate solution pH 5.5 (Thermo Fisher Scientific).
10. 100% ethanol.
11. Laminar flow hood.
12. Cell culture incubator.
13. Table-top refrigerated microcentrifuge.
14. Sterile 1.5 ml microcentrifuge tubes.

2.2.2 Cloning of AGO1 Homology Arms

1. Oligonucleotide primers (*see* Table 1).
2. Phusion Flash High-Fidelity PCR master mix (Thermo Fisher Scientific).
3. Taq DNA polymerase (New England Biolabs).
4. 100 mM dATP solution.
5. 1 M Magnesium chloride solution.

6. pGEM-T Easy kit (Promega).
7. 100 mg/ml Carbenicillin, disodium salt (Sigma).
8. LB-Carbenicillin broth: *see* Subheading 2.1.
9. LB-Carbenicillin agar: *see* Subheading 2.1.
10. QIAquick Gel Extraction Kit (Qiagen).
11. QIAprep Spin Miniprep Kit (Qiagen).
12. 6× DNA loading dye.
13. Thermocycler.
14. Sterile 1.5 ml microcentrifuge tubes.
15. 0.2 ml nuclease-free PCR tubes.
16. SYBR Safe DNA gel stain.
17. Agarose gel electrophoresis apparatus with power supply.
18. 0.5 × TBE buffer: 40 mM Tris, 45 mM Boric acid, 1 mM EDTA pH 8.
19. 1% Agarose gel: Dissolve 1 g Agarose in 100 ml 0.5 × TBE buffer by boiling, cool to ~50 °C, add SYBR Safe DNA gel stain (1×) and pour in casting tray with the gel comb.
20. UV transilluminator.
21. Nanodrop spectrophotometer.

2.3 Transfection of S2R+ Cells

1. *Drosophila* S2R+ cells (Drosophila Genomics Resource Centre, stock #150).
2. 75 cm^2 cell culture flask (Sigma).
3. 6-well cell culture plates (Sigma).
4. Schneider's *Drosophila* medium (Sigma).
5. Fetal bovine serum (Thermo Fisher Scientific).
6. Penilicillin-streptomycin (10,000 U/ml, Thermo Fisher Scientific) as in 2.1.1 item 5 for consistency.
7. Fugene HD transfection reagent (Promega).
8. Sterile 1.5 ml microcentrifuge tubes.
9. Cell culture incubator.
10. Laminar flow hood.

2.4 Detection of HA-Tagged AGO1 by Immunostaining

1. 24-well cell culture plates.
2. Phosphate buffered saline (PBS) solution: Dissolve one PBS tablet (Sigma) in 200 ml ultrapure water. Store at room temperature.
3. Paraformaldehyde, methanol-free (VWR).
4. PBT: PBS, 0.1% TritonX-100, store at room temperature.
5. PBTS: PBT, 5% Normal goat serum.

6. Anti-AGO1 mouse monoclonal antibody (* gift of Mikiko C. Siomi, Keio University School of Medicine, Japan).
7. Anti-HA rat monoclonal antibody (clone 3F10).
8. Goat anti-Mouse IgG (H + L) Cy5 (Thermo Fisher Scientific).
9. Goat anti-Mouse IgG (H + L) Alexa Fluor 488 (Thermo Fisher Scientific).
10. PBT-Hoechst: add 1 μl 20 mM Hoechst 33342 solution (Thermo Fisher Scientific) to 1230 μl PBT (final concentration of Hoechst is 10 μg/ml).
11. Mounting medium (Dako).
12. Glass slides and round coverslips.

3 Methods

Perform all procedures at room temperature unless otherwise specified and use autoclaved ultrapure water in reactions throughout. Assemble reactions in sterile 1.5 ml microcentrifuge tubes except for PCR amplifications where 0.2 ml tubes are used. Add the reaction components in the order as listed in the Tables. After assembly of the reaction, mix well and centrifuge briefly to collect the reaction at the bottom of the tube.

3.1 Generation of AGO1-sgRNA-Cas9 Expression Plasmid

1. Prepare pAc-sgRNA-Cas9 plasmid DNA (Fig. 1a) using QIAprep Spin Miniprep Kit. Measure the concentration of the sample using Nanodrop spectrophotometer and dilute to 200 ng/μl (*see* **Note 3**).
2. Digest the plasmid DNA with restriction enzyme BspQ1 (Fig. 1b). Assemble the reaction on ice as shown in Table 2. Incubate the reaction tube at 50 °C on block heater for 2 h (*see* **Note 4**).
3. Dephosphorylate the 5′ ends of the linearized plasmid DNA by adding 1 μl Calf Intestinal Alkaline Phosphatase (10 U/μl) to the reaction. Incubate at 37 °C for 10 min.
4. Purify the DNA using QIAquick PCR purification kit. Elute in 30 μl Elution Buffer. Determine the concentration of the DNA using Nanodrop spectrophotometer and dilute to 50 ng/μl (*see* **Note 3**).
5. Order oligonucleotides AGOsgF and AGOsgR (sequences in Table 1) containing the *AGO1* targeting sequence with the overhangs for cloning into BspQ1 restriction site (underlined). Prepare 100 μM stock solutions by resuspending the lyophilized oligonucleotides in ultrapure water.
6. Assemble the oligonucleotide annealing reaction on ice as shown in Table 3. Mix and transfer the tube to 98 °C on a

Table 2
Restriction digestion of plasmid

Component	Volume (μl)	Final concentration
Ultrapure water	33	
pAc-sgRNA-Cas9 (200 ng/μl)	10	40 ng/μl
NEB buffer 3.1 (10×)	5	1×
BspQ1 (10 U/μl)	2	0.4 U/μl
Total volume	50	

Table 3
Oligonucleotide annealing reaction

Component	Volume (μl)	Final concentration
AGOsgF (100 μM)	10	25 μM
AGOsgR (100 μM)	10	25 μM
Annealing buffer (2×)	20	1×
Total volume	40	

Table 4
Phosphorylation reaction

Component	Volume (μl)	Final concentration
Ultrapure water	7	
T4 DNA ligase buffer (10×)	1	1×
Annealed oligo mix	1	
T4 Polynucleotidyl kinase (10 U/ul)	1	1 U/μl
Total volume	10	

block heater. Incubate for 2 min. Switch off the block heater and allow it to cool to room temperature. Centrifuge the tube briefly and keep on ice (*see* **Note 5**).

7. Phosphorylate the annealed oligonucleotides by preparing the reaction mix as shown in Table 4. Incubate at 37 °C for 30 min. Dilute the reaction 10× by adding 90 μl ultrapure water. Transfer the tube on ice (*see* **Note 6**).

Table 5
Ligation reaction

Component	Volume (μl)	Final concentration
Ultrapure water	4	
T4 DNA ligase buffer (10×)	1	1×
BspQ1 digested vector (50 ng)	2	10 ng/μl
Phosphorylated oligos (10× diluted)	2	
T4 DNA ligase (400 U/μl)	1	40 U/μl
Total volume	10	

8. Assemble the ligation reaction with the linearized vector (from **step 4**) and annealed oligonucleotide as in Table 5. Incubate at 16 °C for 2 h (*see* **Note 7**).
9. Transform 50 μl aliquots of competent *E. coli* cells with 2 μl of ligation reaction as well as dephosphorylated BspQ1 digested vector (50 ng), incubate both tubes on ice for 30 min followed by heat shock at 42 °C for 30 s. Return the tube on ice and keep for an additional 2 min. Add 150 μl SOC media to the cells, mix by pipeting and plate 50 μl on LB-carbenicillin agar plates. Incubate the plates in 37 °C incubator overnight (*see* **Note 8**).
10. Pick 5 single colonies from the plate containing cells transformed with the ligation reaction (*see* **Note 9**). Inoculate 5 ml of LB-carbenicillin broth with each colony and culture overnight at 37 °C with shaking.
11. Prepare plasmid DNA using QIAprep Spin Miniprep Kit and verify the insertion of *AGO1* targeting sequence at the BspQ1 restriction site of pAc-sgRNA-Cas9 by sequencing with U63F primer (*see* Table 1). Dilute DNA (pAc-AGO1sgRNA-Cas9) to 100 ng/μl (*see* **Note 3**).

3.2 Generation of AGO1-HA Donor Plasmid

3.2.1 Preparation of Genomic DNA from Drosophila S2R+ Cells

1. Prepare complete culture medium as shown in Table 6 (*see* **Note 10**). Grow S2R+ cells in 75 cm^2 cell culture flask at 25 °C to ~80% confluency.
2. Harvest 1 ml culture by centrifugation at 2000 × *g* for 2 min. Wash twice in 1 ml PBS (*see* **Note 11**).
3. Resuspend cell pellet in 0.2 ml Lysis Buffer containing 200 μg/ml proteinase K and incubate at 55 °C for 2 h (*see* **Note 12**).

Table 6
***Drosophila* S2R+ complete culture medium**

Component	Volume (ml)	Final concentration
Schneider's *Drosophila* Medium	500	
Fetal bovine serum	50	10%
Penicillin-streptomycin (10,000 U/ml)	5	100 U/ml

4. Mix lysate with equal volume of phenol:chloroform:isoamyl alcohol and leave the tube horizontally on a shaker with gentle rocking for 1 h (*see* **Note 13**).
5. Centrifuge at 10,000 × *g* for 3 min. Transfer the aqueous phase to a new 1.5 ml tube.
6. Precipitate DNA by adding 0.1 volume 3 M sodium acetate solution pH 5.5 and 2.5 volume 100% ethanol. Centrifuge at 13,000 × *g* for 15 min at 4 °C.
7. Wash pellet with ice-cold 70% ethanol, air dry for 5 min and suspend in 50 µl of Elution Buffer. Dilute DNA to 10 ng/µl (*see* **Note 3**).

3.2.2 Cloning of AGO1 Homology Arms

1. Order primers (AgoUpF, AgoUpR, AgoDownF, AgoDownR) for amplification of *AGO1* left and right homologous arms carrying HA sequence (*see* Table 1 for the sequences, the sequence encoding HA peptide is underlined, Fig. 2a).
2. Prepare a master mix for two PCR reactions as shown in Table 7 in 0.2 ml PCR tubes. Add the forward and reverse primers pairs in separate tubes followed by equal volumes of the master mix in each tube. Mix by pipeting.
3. Transfer both tubes in thermal cycler and perform the PCR reaction with the program as shown in Table 8.
4. Add 10 µl of 6× DNA loading dye to each PCR sample and load into separate wells in 1% agarose gel containing SYBR Safe DNA gel stain (1×). Run gel at 10 V/cm for 30 min (*see* **Note 14**).
5. Visualize the gel under green light and cut out the DNA band with a clean scalpel (the expected size of the left homology arm amplicon is 1049 bp while that for the left homology arm is 1036 bp). Purify the DNA using QIAquick Gel Extraction kit. Elute in 30 µl Elution Buffer. Determine the DNA concentration using Nanodrop spectrophotometer and dilute each DNA sample to 100 ng/µl.

Table 7
PCR reaction

Component	Volume (μl)	Master mix (for two reactions)	Final concentration
Ultrapure water	19	38 μl	
Phusion flash master mix (2×)	25	50 μl	1×
Forward primer[a] (10 μM)	2.5	–	0.5 μM
Reverse primer[a] (10 μM)	2.5	–	0.5 μM
Genomic DNA (10 ng/μl)	1	2 μl	0.2 ng/μl
Total volume	50	90 μl = 45 μl/reaction	

[a]Use primer combinations AgoUpF and AgoUpR for amplification of left homology arm, and AgoDownF and AgoDownR for the right homology arm

Table 8
PCR program

Initial denaturation	98 °C × 10 s	1 × cycle
Denaturation	98 °C × 1 s	30 × cycles
Annealing*	T °C × 5 s	
Extension	72 °C × 15 s	
Final extension	72 °C × 1 min	1 × cycle
Hold	4 °C	∞

[a]Use 60 °C annealing temperature for the reaction with primer pair AgoUpF and AgoUpR and 70 °C for AgoDownF and AgoDownR pair

6. Join the DNA fragments by performing overlap extension PCR in a single 0.2 ml tube (Fig. 2b, c). Assemble the reaction as shown in Table 9 and perform the PCR reaction on a thermocycler using the program in Table 10.
7. Perform agarose gel electrophoresis as described in **step 4**. Cut out the DNA band corresponding to 2 kb and purify the DNA using QIAquick Gel Extraction kit. Elute in 30 μl Elution Buffer (*see* **Note 15**).
8. Assemble the DNA tailing reaction as shown in Table 11 on ice. Incubate at 72 °C for 30 min. Return the tube on ice (*see* **Note 16**).

Table 9
Overlap extension PCR reaction

Component	Volume (μl)	Final concentration
Ultrapure water	18	
Phusion flash master mix (2×)	25	1×
AgoUpR (10 μM)	2.5	0.5 μM
AgoDownF (10 μM)	2.5	0.5 μM
AGO1 left arm PCR product (100 ng/μl)	1	2 ng/μl
AGO1 right arm PCR product (100 ng/μl)	1	2 ng/μl
Total volume	50	

Table 10
PCR program for overlap extension PCR

Initial denaturation	98 °C × 10 s	1 × cycle
Denaturation	98 °C × 1 s	30 × cycles
Annealing	60 °C × 5 s	
Extension	72 °C × 30 s	
Final extension	72 °C × 1 min	1 × cycle
Hold	4 °C	∞

Table 11
DNA tailing reaction

Component	Volume (μl)	Final concentration
Ultrapure water	4	
Taq DNA polymerase buffer (10×)	5	1×
Purified PCR product	30	
$MgCl_2$ (25 mM)	5	2.5 mM
dATP (10 mM)	1	0.2 mM
Taq DNA polymerase (5 U/μl)	5	0.5 U/μl
Total volume	50	

Table 12
Transfection reaction

Component	1 (Empty vector)	2 (Ctrl CRISPR)	3 (Ctrl HA)	4 (Expt)
Pre-warmed culture media	54 µl	54 µl	54 µl	54 µl
Elution buffer	20 µl	20 µl	20 µl	–
pAc-sgRNA-Cas9 plasmid (100 ng/µl)	20 µl			
pAc-*AGO1*sgRNA-Cas9 plasmid (100 ng/µl)	–	20 µl	–	20 µl
pGEM-T easy *AGO1-HA* plasmid (100 ng/µl)	–	–	20 µl	20 µl
FuGENE HD	6 µl	6 µl	6 µl	6 µl

9. Use 5 µl tailing reaction to ligate with pGEM-T Easy vector. Perform blue-white selection and prepare plasmid DNA from LB-carbenicillin culture inoculated with a single white colony using QIAprep Spin Miniprep Kit. Verify the donor vector by sequencing the DNA using T7 and SP6 primers. Dilute DNA (pGEM-AGO1-HA) to 100 ng/µl.

3.3 Transfection of S2R+ Cells

1. Grow S2R+ cells in 75 cm^2 cell culture flask at 25 °C to ~70% confluency (*see* **Note 17**).
2. Dilute cells to 10^6 cells/ml in 20 ml culture media.
3. Plate 2 ml of the diluted cells into 4 wells of a 6-well cell culture plate. Leave cells at 25 °C for a minimum of 3 h (*see* **Note 18**).
4. Aliquot reagents and DNA samples in 1.5 ml tubes as shown in Table 12. Mix the reaction by flicking and leave for 20 min for the formation of DNA-FuGENE complex. Drop the mixture evenly onto appropriate wells of the plate, mix by gentle swirling and keep the plate in 25 °C for 72 h (*see* **Note 19**).
5. Remove the transfection media and put fresh pre-warmed 2 ml media into each well. Incubate at 25 °C (*see* **Note 20**).

3.4 Detection of HA-Tagged AGO1 by Immunostaining

1. Transfer transfected cells (2×10^6 cells/ml) to 24-well culture plate with sterile round coverslips. Leave cells for at least 3 h (*see* **Note 18**).
2. Carefully remove the media and wash once with 500 µl PBS (*see* **Note 21**).
3. Fix cells in each well with freshly prepared 500 µl PBS-4% paraformaldehyde for 10 min.

Table 13
Primary antibody solutions

Antibody	Volume	Dilution
Mouse anti-AGO1*	1 µl	1:1000
Rat anti-HA	1 µl	1:1000
PBTS	998 µl	
Total volume	1 ml	

Table 14
Secondary antibody solutions

Antibody	Volume	Dilution
Goat anti- mouse Cy5	1 µl	1:1000
Goat anti-rat Alexa Fluor 488	1 µl	1:1000
PBTS	998 µl	
Total volume	1 ml	

4. Wash twice in 1 ml PBT for 5 min each. Remove PBT completely.
5. Block in 200 µl PBTS for 1 hour. Remove blocking solution fully.
6. Prepare primary antibody mix in PBTS (*see* Table 13) and add 200 µl to each well. Seal the lid with Parafilm and incubate at room temperature overnight.
7. Wash twice with 500 µl PBTS for 10 min each.
8. Prepare the secondary antibodies mix as shown in Table 14 and add 200 µl to each well. Leave the culture plate in dark for 2–4 h at room temperature (*see* **Note 22**).
9. Remove the secondary antibody mix and add 200 µl PBT-Hoechst (10 µg/ml final concentration) solution. Incubate for 5 min in dark.
10. Wash twice with 500 µl PBT for 10 min each and once with PBS briefly.
11. Carefully pick the coverslip with fine forceps and mount on a glass slide with a drop of mounting medium. Allow the mounting medium to dry and seal the coverslip edges with nail polish (*see* **Note 23**).

12. Image the slides using a confocal scanning laser microscope using blue laser (excitation wavelength 488 nm), red laser (excitation wavelength 640 nm), and 405 laser diode. Collect images in separate channels and visualize in an appropriate software. A representative image is shown in Fig. 3. Cytoplasmic staining with anti-HA antibody shows successful tagging of the *AGO1* locus.

4 Notes

1. Other *E. coli* strains (such as XL1-Blue) can also be purchased. Alternatively, chemically competent cells with high transformation efficiency (>10^7cfu/μg DNA) can be prepared using the protocol described in [17] which is sufficient for most cloning procedures.
2. Prepare 100 mg/ml carbenicillin in water (1000 × stock) and freeze at −20 °C. Keep the working aliquot at 4 °C. Ampicillin can be substituted for carbenicillin and used at same concentration for both liquid media and solid agar plates. We routinely use carbenicillin as this tends to be more stable than ampicillin and reduces the formation of satellite colonies upon long-term incubation of agar plates at 37 °C.
3. The DNA samples can be stored at this stage. We recommend 4 °C for short-term (up to a week) while they should be stored frozen at −20 °C for longer times. The cleanliness and purity of the DNA is critical for the efficiency of the protocol. Therefore, the DNA should be prepared using recommended kits and the DNA quality assessed by measuring the ratio of $Absorbance_{260}$ to $Absorbance_{280}$—a value of 1.8 indicates good quality.
4. The volume of restriction enzyme in a reaction should not exceed 5% (v/v) while the final concentration of DNA should be below 100 ng/μl. Water bath set at the recommended temperature can be used instead of the block heater.
5. This step takes up to 2 h to complete. Do not transfer the reaction tube directly from block heater to room temperature or ice as annealing efficiency is dependent on slow cooling of the reaction. Alternatively, this step can be performed using a thermocycler with a stepwise temperature program with low ramping.
6. Unless specified while ordering, oligonucleotides are supplied non-phosphorylated and should be 5′phosphorylated before using them for ligation reaction with a de-phosphorylated DNA vector. Store oligonucleotide solutions frozen at −20 °C.
7. Alternatively, the ligation reaction can be performed at room temperature for 5 min using Quick ligation kit (New England Biolabs).

8. The transformation efficiency can be increased by incubating the SOC media-cell mix at 37°C with shaking for 1 h before plating.
9. The plates containing dephosphorylated BspQ1 digested vector should be devoid of bacterial colonies. Otherwise (1) check that the vector is linearized upon BspQ1 digestion by agarose gel electrophoresis, (2) repeat the dephosphorylation step after purifying the linearized vector DNA using QIAquick PCR purification kit (Qiagen).
10. Store complete culture media at 4 °C. Aliquot Fetal Bovine Serum and Penicillin-Streptomycin and store frozen at −20 °C. Always pre-warm the complete media to 25 °C before adding to the cells.
11. The cell pellet can be stored at this step at −20 °C if required.
12. Cell lysis can be continued for overnight at 37 °C if necessary.
13. Rigorous mixing or vortexing should be avoided at this step.
14. Prepare and run agarose gel in 0.5 × TBE. Load appropriate DNA marker for estimation of the molecular weight of the PCR product. Alternatively, ethidium bromide (final concentration 0.5 μg/ml) can be used to visualize DNA bands under UV light. In this case, always wear protective eyeglasses and stand behind protective shield when looking at the gel under the UV light. Minimize UV exposure to the gel.
15. In absence of PCR amplification, Taq polymerases and PCR programs dedicated to amplification of longer templates (e.g., Expand High-Fidelity PCR system (Roche)) can be used.
16. Phusion DNA polymerase generates blunt end PCR products. To facilitate TA cloning, it is necessary to add a deoxyadenosine at the 3′end of the blunt-end DNA which is achieved by the tailing reaction. To avoid this step, the overlap extension PCR reaction can be carried out using Taq polymerases that add deoxyadenosine overhangs at the 3′ends (e.g., Expand High-Fidelity PCR system (Roche)). Alternatively, the blunt end PCR product can be directly subcloned using Blunt-end PCR cloning kits.
17. Assess cell viability using Trypan blue staining. Use cells for transfection only if the viability is >80%.
18. The cells can be left overnight at 25 °C to allow better adherence to the culture plate.
19. Do not exceed the recommended DNA amount or alter the ratio of DNA to FuGENE (1 μg DNA:3 μl FuGENE ratio). Add FuGENE at the center of the mix avoiding contact with the side of the tube.

20. Stable cell lines with homozygous insertion can be generated by adding complete culture media supplemented with puromycin (final concentration 5 μg/ml). Passage the cells in fresh selection media as confluency reaches ~70%. It takes 3–4 weeks to select stable cell lines using puromycin selection. For quicker selection and an additional marker (GFP), the vector pGTL-1 can also be used for this protocol [18]. pGTL-1 contains blasticidin resistance gene which allows selection of stable cell lines by culturing cells in complete media containing 25 μg/ml Blasticidin.
21. Tilt the culture dish slightly and pipette from the edges avoiding the cells on coverslip. The cells should never be allowed to dry at any stage of the procedure.
22. The cells can be incubated at 4 °C for overnight. Cover the culture plate in aluminum foil during the incubation with fluorescent antibodies.
23. For analyses with an inverted microscope, place the coverslip on a slide with the cell side facing up, add a drop of mounting medium and gently place another coverslip on top of it. Alternatively for upright microscopes put a drop of mounting media on the slide and place the coverslip with the cell side facing down. Avoid bubbles during mounting.

References

1. Chapman JR, Taylor MRG, Boulton SJ (2012) Playing the end game: DNA double-strand break repair pathway choice. Mol Cell 47(4):497–510
2. Capecchi MR (1989) Altering the genome by homologous recombination. Science 244(4910):1288–1292
3. Brizzard B (2008) Epitope tagging. BioTechniques 44(5):693–695
4. Jinek M et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821
5. Wang H, Russa ML, Qi LS (2016) CRISPR/Cas9 in genome editing and beyond. Annu Rev Biochem 85(1):227–264
6. Harrison MM et al (2014) A CRISPR view of development. Genes Dev 28(17):1859–1872
7. Port F et al (2014) Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in drosophila. Proc Natl Acad Sci U S A 111(29):E2967–E2976
8. Yu Z et al (2013) Highly efficient genome modifications mediated by CRISPR/Cas9 in drosophila. Genetics 195(1):289–291
9. Bassett AR, Liu JL (2014) CRISPR/Cas9 and genome editing in drosophila. J Genet Genomics 41(1):7–19
10. Bassett AR et al (2013) Highly efficient targeted mutagenesis of drosophila with the CRISPR/Cas9 system. Cell Rep 4(1):220–228
11. Gratz SJ et al (2013) Genome engineering of drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194(4):1029–1035
12. Ren X et al (2014) Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in drosophila. Cell Rep 9(3):1151–1162
13. Förstemann K et al (2007) Drosophila microRNAs are sorted into functionally distinct Argonaute complexes after production by dicer-1. Cell 130(2):287–297

14. Smibert P et al (2013) Homeostatic control of Argonaute stability by microRNA availability. Nat Struct Mol Biol 20(7):789–795
15. Bassett AR et al (2014) Mutagenesis and homologous recombination in drosophila cell lines using CRISPR/Cas9. Biol Open 3(1):42–49
16. Horton RM et al (2013) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. BioTechniques 8(5):528–535
17. Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96(1):23–28
18. Ghosh S, Tibbit C, Liu JL (2016) Effective knockdown of *Drosophila* long non-coding RNAs by CRISPR interference. Nucl Acids Res 44(9):e84

Chapter 16

Accurate Profiling and Quantification of tRNA Fragments from RNA-Seq Data: A *Vade Mecum* for MINTmap

Phillipe Loher, Aristeidis G. Telonis, and Isidore Rigoutsos

Abstract

There is an increasing interest within the scientific community in identifying tRNA-derived fragments (tRFs) and elucidating the roles they play in the cell. Such endeavors can be greatly facilitated by mining the numerous datasets from many cellular contexts that exist publicly. However, the standard mapping tools cannot be used for the purpose. Several factors complicate this endeavor including: the presence of multiple identical or nearly identical isodecoders at various genomic locations; the presence of identical sequence segments that are shared by isodecoders of the same or even different anticodons; the existence of numerous partial tRNA sequences across the genome; the existence of hundreds of "lookalike" sequences that resemble true tRNAs; and others. This is generating a need for specialized tools that can mine deep sequencing data to identify and quantify tRFs. We discuss the various complicating factors and their ramifications, and how to use and run MINTmap, a tool that addresses these considerations.

Key words Transfer RNA, tRNA, tRNA-derived fragments, tRFs, tRF license plate, internal tRFs, i-tRFs, 5′-halves, 3′-halves, 5′-tRFs, 3'-tRFs tRNA space, tRNA-lookalikes, MINTmap, MINTsubmit, MINTbase, MINTcodes

1 Introduction

An integral component of translation is transfer RNA (tRNA). These molecules serve as the actual translators and linkers between the genetic code and the amino acids. However, the functions and roles of tRNAs extend well beyond delivering amino acids to the ribosome. In fact, their intracellular levels and their physiological variation, or variations due to diseases, have been shown to be of significance [1–4]. Equally importantly, tRNAs serve as the template for the production of smaller RNA molecules, termed tRNA-derived fragments (tRFs), which are functional molecules and regulatory in nature, rather than degradation products [5, 6]. The roles of tRFs that have been elucidated so far range from protein-coding gene expression regulation via microRNA (miRNA)-like mechanisms [7, 8] to the decoying of RNA-binding proteins [9] and the displacement of messenger RNAs (mRNAs) from the

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2_16,

ribosome [10]. There is currently an increasing research interest in the field of tRFs that concerns biological processes and diseases in humans and in a wide variety of "leaves" in the tree of life [11].

To accelerate tRF research, a reliable method for detecting and quantifying the molecule(s) under consideration is needed. Because of the partial sequence overlaps exhibited by tRFs from the same isodecoder, conventional qPCR methods are not directly applicable [12, 13]. Instead, approaches that are specific to both the 5′ and 3′ endpoints of a tRF are needed [14]. Currently, RNA sequencing (RNA-seq) technologies of molecules shorter than ~50 nucleotides (nt) remain the best option for the unbiased assessment of tRF abundance levels. As a matter of fact, RNA-seq can easily distinguish molecules that differ by a single nucleotide, and does not require a priori knowledge of the sequences that are being sought.

The product of an RNA-seq experiment is a big text file containing the sequenced reads that one can readily mine. To "make sense" of the reads, i.e., to identify and distinguish the different classes of molecules that were captured, specific tools (mappers) are used that *map* the reads on a reference genome. In the post-genomic era, this is a fairly straightforward task. However, read mapping is one of the bottlenecks in the field of tRF research. In fact, before mapped reads can be flagged as tRFs, one needs to properly establish and characterize the specific regions in the genome that code for tRFs, the *tRF space* [13, 15, 16]. In what follows, we present and discuss the potential problems that arise from the idiosyncrasies of tRNA space and the details of how to run the MINTmap [16] tool.

The task of identifying genomic loci that code for tRNAs, the *tRNA space* [13, 16] has been a topic of research activity for nearly two decades. Consequently, there is a reasonably good understanding of what constitutes the tRNA space. On the other hand, the definition of tRF space is not as well defined.

First, it is reasonable to assume that the tRF space and the tRNA space should have a very high overlap. Therefore, one possibility is to claim as "tRFs" all the sequenced reads that get mapped inside the tRNA space. The potential problem in this scenario is that some shorter sequences inside tRNA space also happen to exist outside of it. Characteristic examples are the many partial tRNAs [17] that are identified by RepeatMasker [18].

A second problem stems from the variability within the tRNA space and the decisions of the algorithms that recognize tRNAs. Such algorithms draw conclusions by using the 2-D folding of the sequence. For example, tRNAscan-SE [19] would mark as *pseudo-tRNAs* those tRNAs that cannot be predicted to have a proper cloverleaf structure, even if they had high sequence similarity with molecules that are known to fold properly. Although this is a valid and reasonable discrimination of tRNAs, it has inevitable

consequences with regard to the definition of a tRF: can one characterize as "tRFs" molecules that may originate from a misfolded tRNA, i.e., an RNA molecule that is not functioning as a tRNA? This question raises a more general issue of post-transcriptional tRNA modifications: should molecules that emerge from sources outside of the biogenetic pathway and life cycle of a tRNA be labeled tRFs? For example, the most commonly discussed modification in tRNAs that affects tRFs is the post-transcriptional addition of a CCA tail at the 3′ end of the mature tRNAs [20]. In fact, there is a specific category of tRFs that contains this trinucleotide addition: these tRFs are therefore considered to emerge from the mature, properly folded tRNA, because the enzymatic CCA addition requires proper folding of the tRNA [20]. Other common modifications are the non-templated addition of a nucleotide at the −1 position of tRNAs that code for histidine (His tRNAs) and base modifications, e.g., the methylation at the base of the stem of the D loop (position 8 or 9 of the mature tRNA) [21, 22].

Although the above concerns may seem negligible when considering the *function* of tRFs, instead of their *biogenesis*, they have a significant impact on the setup of the mapping algorithm. By default, most options of the commonly used searching [23–26] and mapping [27–30] algorithms would not be good choices for tRF analysis from RNA-seq datasets [7, 13]. For example, the short length of tRFs and the repeat nature of tRNAs can lead to statistically insignificant alignment scores and a bias in the expression profiles as these genomic mappings would not be reported.

In addition to these considerations, without proper editing, tRFs that end with the post-transcriptionally added CCA tail or other modifications will not map to the genome. Trimming the CCA trinucleotide before processing is not a viable alternative either: doing so would increase the probabilities that the trimmed read(s) are mapped to non-relevant genomic positions. Other concerns include intron editing, multiple-counting of sequenced reads that can be mapped to more than one distinct tRNAs, or the trade-off between allowing mismatches to accommodate errors due to base modifications and accurately locating the tRNA source of a tRF.

According to version v1.0 of the gtRNAdb database [31], the human nuclear genome contains 610 loci that encode for tRNAs [16]. The mitochondrial genome encodes an additional 22 tRNAs. We also reported 497 tRNA-lookalikes in the nuclear genome [17]. The total sequence space that is spanned by the these 1,129 loci is relatively small when considering the size of each locus (~72 nt) compared to the size of the human genome, and the fact that some of the sequences repeat among different tRNAs. Hence, it is reasonable to investigate the possibility of creating a fast and exhaustive mapping pipeline that can be built once per genomic assembly and can address the computationally most-demanding part, namely, the identification of all possible

tRFs. One could then use the output from such a pipeline to thoroughly study the tRFs contained in any short RNA-seq dataset [7]. We recently developed such an approach. MINTmap, for *MI*tochondrial and *N*uclear *T*RNA fragment *map*per, tackles the idiosyncrasies of the tRNA space, maps the sequencing reads exhaustively and comprehensively, and accurately quantifies the present tRFs [16]. MINTmap achieves very high speeds while having only minimal resource requirements.

2 Materials

The MINTmap tool and source code are available for download at https://cm.jefferson.edu/MINTcodes/. Prior to running MINTmap, the dataset of sequenced reads must already be quality- and adaptor-trimmed with a tool such as cutadapt [32]. MINTmap was designed to be very efficient. As a matter of fact, mining of the tRFs contained in a given short RNA-seq dataset takes only a few minutes. To do so for a user's short RNA-seq dataset, we recommend an Intel Core i5/i7, or equivalent, with at least 2 Gb of main memory. Additionally, the installation of Java v1.8 and of Perl v5 or greater is required to run the MINTmap script. Here, we assume the most typical scenario in which the user makes use of the prebuilt lookup tables that are distributed with the MINTmap codes and which are derived from both nuclear tRNAs contained in gtRNAdb [31] and mitochondrial tRNAs contained in NCBI. Should a user need to build lookup tables for their own collection of tRNAs (a one-time event) we recommend a compute-node with a minimum of 16-cores, e.g., Intel Xeon, Intel Core i5/i7, or equivalent. The sequenced reads [33] available under GEO accession GSE16579 that are included in MINTmap's example dataset as well as in this paper can be downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16579.

3 Methods

3.1 The Algorithm

In what follows, we summarize the steps that MINTmap takes when mapping a short RNA-seq dataset. More details and the series of steps used to create the lookup tables can be found in the original MINTmap publication [16].

By default, MINTmap will profile tRFs that overlap the mature tRNA sequences from 640 tRNA loci. These loci consist of 508 true and 102 pseudo human tRNAs from v1.0 of gtRNAdb [31], the 22 known human mitochondrial tRNAs from NCBI, and 8 exact human tRNA lookalikes [17] that are present in the human nuclear genome [16] for a total of 640 loci [16]. This reference set was carefully chosen to capture both

nuclear and mitochondrial tRFs and allow for the enumeration of all isodecoders to which a tRF sequence could be mapped [16]. We selected to use v1.0 of gtRNAdb, instead of the more recent release v2.0 [34], because we found that the latter contains inconsistencies that would directly affect the quality of the reported results [35].

When a user invokes MINTmap to generate a dataset's tRF profile, frequency counts are generated for each distinct sequence segment that is supported by the sequenced reads. During this process, the algorithm excludes sequenced reads that could not possibly be tRF sequences if exact matching were used. MINTmap comes with a prebuilt lookup table for *H. sapiens* that includes *all* possible sequences with lengths between 16 and 50 nt that overlap the 640 tRNAs being considered. This table was pre-generated through exhaustive enumeration of all possible sub-sequences that can be generated by these 640 tRNA mature sequences and incorporates provisions to accommodate a single nucleotide post-transcriptional addition to the 5′ end (e.g., case of G at the 5′ end of $tRNA^{His}$ [36]) as well as the non-templated CCA. Any sequenced reads that are not present in this lookup table are discarded as they do not exist in tRNA space. For each unique tRF sequence that survives, its frequency is also preserved. For tRFs that are present in multiple isodecoders, additional bookkeeping ensures that no double-counting takes place.

Once the tRFs that are present among the sequences have been identified, MINTmap further separates the tRFs into two categories: *exclusive* tRFs and *ambiguous* tRFs. Exclusive tRFs are those whose sequences cannot be found in any genomic locations other than the 640 tRNAs being considered. On the other hand, ambiguous tRFs have identical instances within the 640 tRNAs and other genomic locations as well [7, 13, 16]. In the output of MINTmap, we list separately tRFs that are exclusive to tRNA space and tRFs that are of ambiguous genomic origin. Also included in the exclusive and ambiguous tRF profiles are the corresponding tRF abundances. By distinguishing between these two categories, we give users the option to decide whether to analyze tRF sequences that might not necessarily originate in tRNA space.

Each reported tRF sequence is accompanied by additional metadata (Fig. 1) that includes the following:

1. *Reads-per-million (RPM) values* for each tRF. These normalized abundance values are generated and shown alongside the raw counts of sequenced reads. Customized RPM values are also permitted (*see* Subheading 3.2.2 below).
2. Each reported tRF is associated with standardized tRF names that include the tRF's "license plate," a unique genome-independent label [37] (Fig. 1, top—first column) as well as genome-centric label(s) [7] (Fig. 1, middle). For tRFs

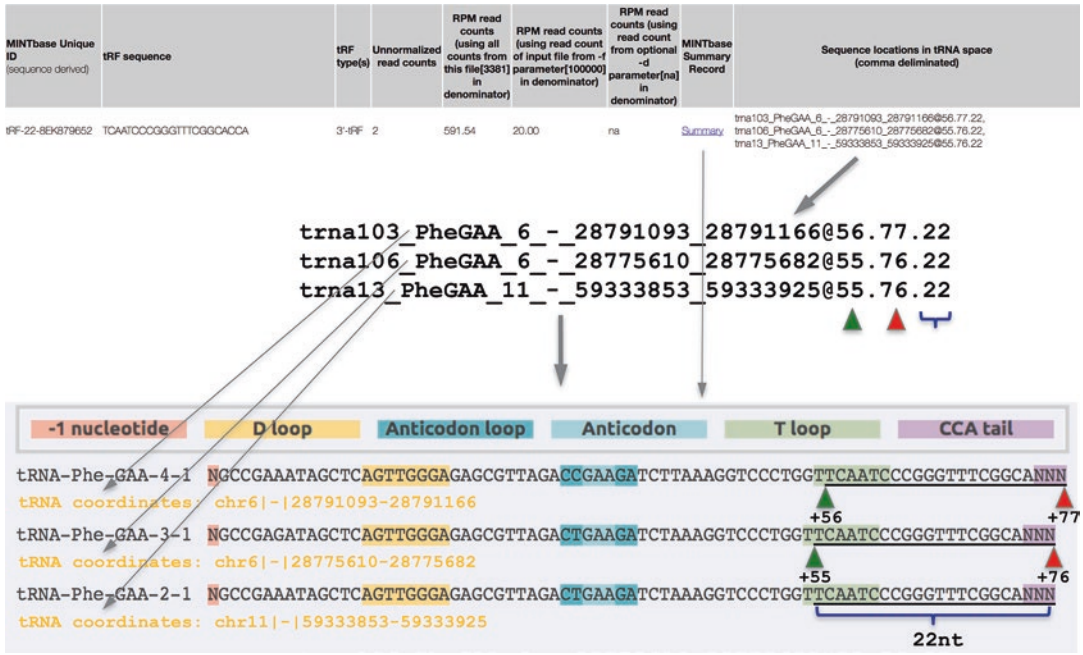


Fig. 1 *Top*: Example of MINTmap's tRF profile output (in HTML format) from MINTmap's included example which uses randomly selected reads from datasets listed under GEO accession GSE16579. Only one tRF, the one with license plate tRF-22-8EK879652, was selected for this example. Columns include: the tRF's license plate, the tRF's sequence, all possible structural types associated with the tRF, un-normalized counts of the already trimmed reads that support the tRF, the tRF's normalized (RPM) abundance using as a denominator the number of reads that map either exclusively or non-exclusively to tRNA space, the tRF's RPM abundance using as a denominator the total number of trimmed reads found in the processed input file, the tRF's RPM abundance using as a denominator the user-definable denominator, link to the MINTbase summary for the tRF, and all possible sequence locations in tRNA space where the tRF can be found. *Middle*: Genome-centric labels for tRF-22-8EK879652. The genomic coordinates for all three tRNAs that contain tRF sequence TCAATCCCGGGTTTCGGCACCA are listed before the "@" symbol. *Bottom*: MINTbase summary output to illustrate the tRF position with respect to the mature tRNA. The genome-centric name contains additional information before the "@" symbol. The *first number* indicates the starting position of the tRF with respect to the 5′ end of the mature tRNA (*green arrow*), the *second number* indicates the ending position of the tRF with respect to the 5′ end of the mature tRNA (*red arrow*), and the *last number* is the length of the tRF

that can have multiple origins, the listed genome-centric labels include all possible sources within tRNA-space from which the tRF can arise.

3. Additionally, each tRF can be tagged with the tRF's structural type (5′-half, 3′-half, 5′-tRF, i-tRF, or 3′-tRF) as we described elsewhere [7, 13, 16, 37]. This is achieved through an optional lookup table that has pre-recorded all possible types associated with each tRF sequence. This optional lookup table for the human genome is provided currently as part of the MINTmap code distribution.

4. The output files are generated in two formats: plaintext and HTML. The HTML-formatted output files include hyperlinks to each tRF's record in MINTbase [37]—all possible tRFs, whether exclusive to tRNA space or ambiguous, have a MINTbase record that summarizes what is currently known for the tRF including a list of all isodecoders that contain it, a list of public datasets in which the tRF can be found, information about the tRF's abundance, etc. By linking each identified tRF to MINTbase, we also provide users with additional data exploration capabilities (*see* also below about data confidentiality).

3.2 Running MINTmap

Figure 2 illustrates the general flow of MINTmap from the standpoint of the user. In its simplest incarnation, the command that will generate tRF profiles looks like.

```
./MINTmap.pl -f <filename>
```

where <filename> is the name of the file that contains short RNA-seq reads that have already been subjected to quality trimming and adapter removal. Several command-line options allow for extended control by the user. Next, we examine these in detail.

3.2.1 Input

A user typically begins with a file that contains the sequenced short reads. This file would be generated by a sequencing facility and be in one of two formats: either a plaintext FASTQ file (ending in ".fastq"), or a gzip-compressed FASTQ file (ending in "fastq.gz").

Before such a file can be analyzed by MINTmap, the reads it contains will need to be preprocessed by quality and adaptor trimming. Quality trimming removes bases from either the 5′ or the 3′ end of a sequenced read if they are deemed to be of "low quality:" In such cases, the actual nucleotide identity cannot be stated with high confidence. Next, adaptor removal is the part of the mapping pipeline that removes any remnants of the adaptors that were ligated during library preparation. These adaptors are *not* part of the transcribed target molecule ("payload"). A tool such as "cutadapt" [32] can be used for both quality and adaptor trimming. To avoid misrepresenting the 3′ endpoint of a molecule, we highly recommend discarding reads in which the 3′ adaptor was not cut (e.g., this can be achieved by using the "–discard-untrimmed" option if using "cutadapt"). As an example, we list here an entry depicting a sequenced read in FASTQ format spanning four lines. The first line starts with a "@" and contains a sequence identifier with an optional description. The second line lists the nucleotide sequence of the read and may include snippets of any ligated adaptors (e.g., 3′ adaptor, shown underlined) that were used during library preparation. The third line starts with a "+" and is followed by an optional identifier. Finally, the fourth line contains the quality values for each nucleotide in the sequence.

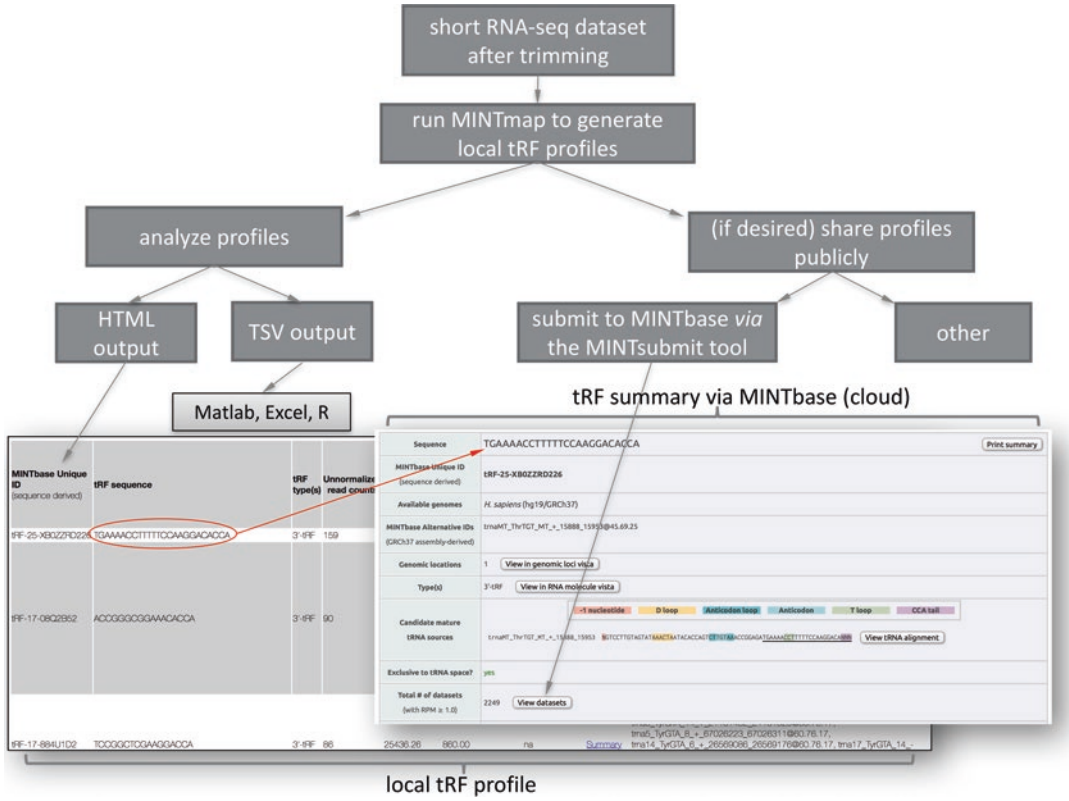


Fig. 2 Logical flow of how the processing of short RNA-seq datasets proceeds when using MINTmap. The MINTmap tool will generate locally tRF profiles after processing quality- and adaptor-trimmed sequenced reads. The profiles can be viewed, studied and analyzed in the user's local environment using a web-browser (HTML) or via standard software packages such as Matlab, Excel, and R. Should the user decide to investigate how the tRFs in the generated profiles compare to the profiles of other public datasets, the user can follow the hyperlinks contained in the HTML-formatted output: these links connect to individual records that we have built for all possible tRFs of lengths 16–50 nt and are populated with what is currently known publicly for each tRF. User-generated profiles can be optionally shared via the MINTsubmit tool (*see* text and Fig. 2) for subsequent inclusion in MINTbase

```
   @SRR029124.3603492    WICMT-SOLEXA_309ETAAXX:
2:76:1574:879 length=36
ATTCCCCGACGGGGAGCCATCGTATGCCGTCTTTTT
+SRR029124.3603492 WICMT-
SOLEXA_309ETAAXX:2:76:1574:879 length=36
   IIIIIIIIIIIIIIIIIIIIIIIIIIII+II$I;
```

Following adapter removal and quality trimming the entry will look like:

```
   @SRR029124.3603492WICMT-SOLEXA_309ETAAXX:2:76
:1574:879length=36
ATTCCCCGACGGGGAGCCA
```

```
+SRR029124.3603492WICMT-SOLEXA_309ETAAXX:2:76:1
574:879length=36
   IIIIIIIIIIIIIIIIII
```

In this particular example, TCGTATGCCGTCTTCTGCTTG was the ligated 3′ adaptor. After quality and adaptor trimming, the sequence TCGTATGCCGTCTTTTT was removed, resulting in a "payload" sequence of ATTCCCCGACGGGGAGCCA.

3.2.2 User-Definable Parameters

We have chosen default settings for MINTmap that facilitate the analysis of *H. sapiens* short RNA-seq datasets using the tRNA reference set (640 tRNA loci) and associated lookup tables that are supplied with the MINTmap codes. A full-fledged command line invocation looks like

./MINTmap.pl –f trimmedfastqfile [–p outputprefix] [–l lookuptable] [–s tRNAsequences] [–o tRFtypes] [–d custom-RPM] [–a assembly] [–h]

and can be used to modify the various settings. Items surrounded by square brackets square brackets "["and"]" indicate arguments with optional or pre-defined arguments. The meaning of each component in the above command line is as follows:

–f trimmedfastqfile

This is a required argument. The string "trimmedfastqfile" is meant to be the filename of the file that contains the sequenced reads to be analyzed. If the file referred to by "trimmedfastqfile" ends in ".gz" it will be treated as a gzip-ped FASTQ file. Nota bene: the file to be used here must already have been quality- and adaptor-trimmed (*see* Subheading 3.2.1 above).

–p outputprefix

If not specified by the user, "outputprefix" will be set automatically to "output" and all output files will be generated in the current working directory. The string "outputprefix" will be prepended in the output files that MINTmap will generate. The string is meant to serve as a mnemonic for the user.

Nota bene: if the prefix includes an absolute or relative path (e.g., "/Users/myuserid/output_of_25Dec2016"), the output directory will be changed accordingly (in this case to "/Users/myuserid/") and the output files with the requested prefix (in this case "output_of_25Dec2016") will be generated therein.

–l lookuptable

Unless specified by the user, this argument will be set automatically to "LookupTable.tRFs.MINTmap_v[MINTmapVersion].txt"—in **Note 2** below we explain in more detail when a user might wish to specify a file other than the provided default one. The string "lookuptable" is meant to be the filename of the file containing all possible tRFs of lengths 16–50 nt for the tRNA

space currently in use. Every tRF sequence listed in this file is associated with a "Y"(es) or "N"(o) value that is the answer to the question "is the tRF sequence present in tRNA space *exclusively*?" The first two lines of the lookup table should include file-level md5 checksum values for the files that are specified with the –s and –o flags to ensure that concordant files are used. For the lookup table that is provided with version v1 of the MINTmap codes these two lines look like:

```
#TRNASEQUENCES:tRNAspace.Spliced.Sequences.
MINTmap_v1.fa
MD5SUM:90396bee2706039122cdd0077ca43f2e
#OTHERANNOTATIONS:OtherAnnotations.MINTmap_
v1.txt
MD5SUM:0a6d60b0c0108885a67fa4a4e9a10807
```

The first line states that the lookup table is meant to be used in conjunction with the tRNA FASTA file "tRNAspace.Spliced.Sequences.MINTmap_v1.fa," which has an MD5 checksum value of 90396bee2706039122cdd0077ca43f2e. MD5 checksums can be generated, if needed, using commands such as "md5sum" for Redhat Linux or "md5" for MacOSX. The second line states that the lookup table is meant to be used in conjunction with the tRF-type-lookup-file "OtherAnnotations.MINTmap_v1.txt," which has an MD5 checksum value of 0a6d60b0c0108885a67fa4a4e9a10807. Below, we explain how one can use the –s and the –o flags to change these files. If either of the two checksum values in this lookup table does not match the actual checksum of the corresponding file (whether the default files are used or user-defined files are passed using options –s or –o) MINTmap will generate an error message and halt.

–s tRNAsequences

Unless specified by the user, this argument will be set automatically to "tRNAspace.Spliced.Sequences.MINTmap_v[MINTmapVersion].fa"—in **Note 2** below we explain in more detail when a user might wish to specify a file other than the default. The string "tRNAsequences" is meant to be the filename of the file that contains the sequences that comprise the true tRNA space and can serve as potential source(s) of tRFs. Only *spliced* tRNA sequences with *no* post-transcriptional modifications should be included in this file. The label of each sequence in this file should be unique and, ideally, should contain the tRNA name and tRNA-locus coordinates that will be used in the output of MINTmap. An example entry comprises two lines and looks like this:

```
>trnaMT_ProTGG_MT_-_15956_16023
CAGAGAATAGTTTAAATTAGAATCTTAGCTT
TGGGTGCTAATGGTGGAGTTAAAGACTTTTTCTCTGA
```

The first of each two-line pair is the "label" line whereas the second is the "data" line. The actual tRNA sequence should be listed on a single "data" line. This is a modified version of the popular FASTA format and differs from the latter in that it does not allow the data to be split over multiple lines. To help users avoid using incorrect parameters, this file's MD5 checksum information is stored on the first line of the lookup table (*see* the description for the –l parameter above).

–o tRFtypes

This is an optional argument. Unless specified by the user, this argument will be set automatically to "OtherAnnotations. MINTmap_v[MINTmapVersion].txt"—*see* **Note 2** for more detail on when one should specify a file other than the default. If the user specifies a file that does not exist, the structural type column in the output files will be set to "na" (*N*ot *A*pplicable). The string "tRFtypes" is meant to be the filename of the file that contains the lookup table which associates tRFs with their structural type. This file contains two columns separated by a "tab." The first column lists the sequences of a tRFs. The second column lists the structural type for the corresponding tRF sequence. If a tRF can have multiple structural types (because it appears at different locations in different isodecoders), the second column will list all structural types, separated by commas. Example entries from the included table include:

```
   AAACTTAACACTTTATAATCA
i-tRF
AAAGGTATTAGAAAAACCATTTCATAACT
5´-tRF
AAAGGTATTAGAAAAACCATTTCATAACTT
5´-half
AAAGGTCCCCGGTTCGAAACCGGGCAGAAACAC
3´-tRF
AAAACTTTACAGTCAGAGGTTCAATTCCTCTTCTTAACACCA
3´-half
   CCCTGTGGTCTAGTGG
5´-tRF, i-tRF
```

Nota bene: to help users avoid using incorrect parameters, this file's MD5 checksum information is stored on the second line of the lookup table (*see* the description for the –l parameter above).

–d customRPM

This is an optional argument. The value "customRPM" is meant to be used as an alternative denominator when computing the RPM abundance of a tRF. When this parameter is defined by the user, an additional column in the output files will be populated with RPM values that have been calculated using this value in the denominator—i.e.,

these values will be equal to raw reads/<customRPM > *1,000,000. A common value to use here is the *original* number of sequenced reads *prior* to quality- and adaptor-trimming. Since MINTmap is only given the sequenced reads after they have been quality filtered and trimmed, use of this parameter enables custom normalization of the tRF abundances in preparation for cross-dataset comparisons.

Nota bene: even when this parameter is not specified, multiple RPM values are reported in the output; however, these RPM values are the result of normalizing using the number of reads that remain after the contents of the original dataset have been trimmed.

–a assembly

Unless specified by the user, this argument is set automatically to "GRCh37." The value of "assembly" refers to the genome assembly version used. The standard notation (e.g., "GRCh37") is expected. This value is listed in the HTML output files in the table headers and is also part of the hyperlink to the MINTbase record of a tRF.

–h

This is an optional argument that presents the user with a list of the various parameters that the user can pass to MINTmap.

3.2.3 Usage Example and Generated Output

The current distribution of MINTmap includes the dataset ExampleRun/exampleInput.fastq.gz, which we compiled by randomly selecting 100,000 reads (after quality- and adaptor-trimming) from the datasets listed under GEO accession GSE16579 [33]. This GEO entry comprises deep-sequencing data from multiple human cell lines. By typing the following command, a user can map the reads in the example dataset using MINTmap's default GRCh37 assembly settings:

```
./MINTmap.pl -f ExampleRun/exampleInput.fastq.
gz
```

Upon completion, MINTmap will generate two kinds of output: abundances for tRFs with exclusive instances in tRNA space (indicated by the infix "-exclusive-," and, abudances for tRFs with instances both inside and outside of tRNA space (indicated by the infix "-ambiguous-." The user can then continue by first examining the example's tRFs that are exclusive to tRNA space by opening the file output-MINTmap_v1-exclusive-tRFs.expression.html with the help of a web browser (note the infix *-exclusive-tRFs* in the filename). The most abundant exclusive tRF is a 3′-tRF with sequence TGAAAACCTTTTTCCAAGGACACCA and is supported by 159 reads (Fig. 2, bottom left). Inspecting the output file further, we see that the tRF is annotated with trnaMT_ThrTGT_MT_+_15888_15953@45.69.25, which means that the tRF overlaps with the mature tRNA sequence of the

mitochondrial $tRNA^{ThrTGT}$, the latter being located on the positive strand of the MT genome between positions 15,888 and 15,953 inclusive. Within the mature tRNA sequence, the tRF spans positions 45 through 69, and is 25 nt in length. More details (Fig. 1) regarding this tRF naming convention appear in our earlier work [7, 16, 37].

MINTmap will generate the following information for every short RNA-seq dataset that is mapped (Fig. 1):

1. tRF profiles, separately for exclusive and for ambiguous tRFs

   For each of these two categories of tRFs, as illustrated in Fig. 2, two files are created, containing "expression.html" and "expression.txt" in their filenames. These profiles are generated in HTML and TSV (tab-separated values) formats, respectively. For each expressed tRF, the tRF profiles include the tRF's "license plate" [16, 37], the tRF's sequence, all possible structural types associated with the tRF, un-normalized counts of the already-*trimmed* reads that support the tRF, the tRF's normalized (RPM) abundance using as a denominator the number of reads that map either exclusively or non-exclusively to tRNA space, the tRF's RPM abundance using as a denominator the *total* number of trimmed reads found in the processed input file, the tRF's RPM abundance using as a denominator the user-definable denominator (*see* –d option above—this option can be used to pass the *total* number of *sequenced* reads to MINTmap), and all possible sequence locations in tRNA space where the tRF can be found. The HTML variant of the tRF profile includes one more item for each tRF, namely a hyperlink to the tRF's MINTbase record [37]. Each such record lists that lists a plethora of other details for the tRF.

   Nota bene: we wish to highlight here the value of the plain-text TSV output files. These files are provided intentionally because the TSV format makes it easy for the user to import, sort, and analyze locally the tRF profiles using standard programs like Microsoft's Excel. Additionally, these files can be readily imported into programs written in Matlab, R, Python, Perl, etc., should the user wish to carry out more detailed analyses.

2. High-level mapping statistics, independently for exclusive and ambiguous tRFs.

   This information is provided in the file ending in ".countsmeta.txt." To illustrate this, we note the line.

   100000 3381 3.38%

   from the above example's file output-MINTmap_v1-exclusive-tRFs.countsmeta.txt file. The line indicates that 100 k reads were mapped, of which 3381 (or 3.38% were identified as belonging to exclusive tRFs).

3.3 "Not-Sharing" Versus "Publicly Sharing" MINTmap's Output

MINTmap is unlike previously reported tools that are used for mapping tRFs [38, 39] in that it can do all of the processing of deep-sequencing data in the user's local environment: no Internet connection is required for the purpose. Moreover, MINTmap generates all of its output, i.e., the tRF profiles, locally. This was an intentional design requirement that we imposed on MINTmap. By generating the tRF profiles locally and by providing the results in a format (TSV) that can be readily analyzed by users with standard software, we give users the ability to study their data at their own pace and determine what they wish to do with the generated results. If and when a user decides to publicly share the output of MINTmap with others in the community, we have made provisions that will allow the deposition of the user's data on MINTbase using a tool called MINTsubmit (*see* below). Both the database and the tool were described in our recent work [37].

MINTbase is a repository and a web-based interactive framework that complements MINTmap and is available at https://cm.jefferson.edu/MINTbase/. MINTbase allows users to study in detail individual tRFs, analyze a collection of tRFs, study the tRF-generation potential of specific isodecoders and isoacceptors, visualize tRF alignments relative to mature tRNAs, report a tRF's position relative to known tRNA features such as loops and the anticodon, etc. Equally importantly, MINTbase allows user to interrogate interactively public datasets and determine for a specific tRF which of the publicly available datasets contain it, in which tissues the tRF has been found expressed, etc. Version (v1.0) of MINTbase contains tRF profiles from 832 public datasets (the vast majority of them derived from actual human tissues, as opposed to cell lines) and detailed records for 7129 expressed tRFs (current as of March 07, 2017). A version v2.0 of MINTbase is being planned for release. The new version will include tRFs from more than 10,000 public datasets.

The MINTsubmit tool (Fig. 3) is available at https://cm.jefferson.edu/MINTsubmit/ and allows a researcher to submit MINTmap profiles for inclusion in the public MINTbase database. Should a researcher choose to share their MINTmap-generated profiles with the wider community, she or he can use MINTsubmit to upload their data. To maximize usefulness, avoid overlaps, and properly track the provenance of deposited data, submitters also need to provide information about the cell/tissue type from which the dataset was generated, etc. Only tRF profiles whose matching RNA-seq datasets have been deposited in a public repository (e.g., NIH's GEO, EBI's ArrayExpress, etc.) and described in a published or accepted manuscript will be considered for inclusion in MINTbase.

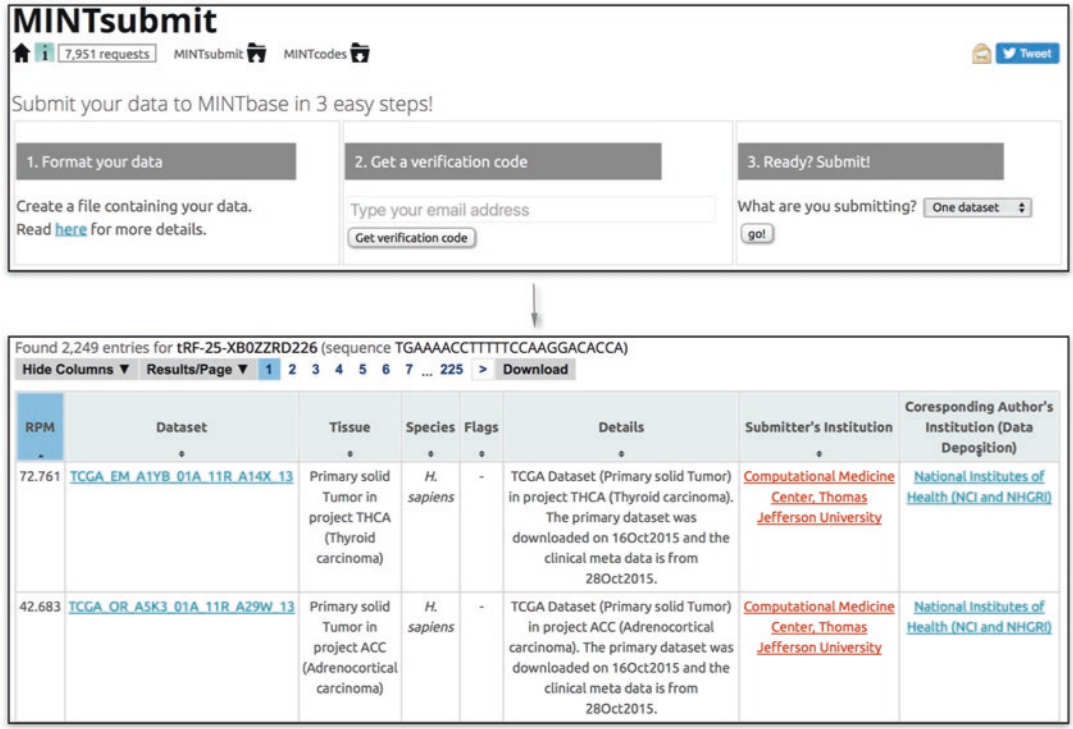


Fig. 3 A user can optionally share publicly their MINTmap-generated tRF profiles by depositing them on MINTbase (https://cm.jefferson.edu/MINTbase/), a cloud-based repository and interactive framework for studying tRFs. The user can use the MINTsubmit tool (*Top panel*) that can be accessed by visiting https://cm.jefferson.edu/MINTsubmit/. Following acceptance of the submission, the submitted tRF profiles become public and can be studied in relation to all other available datasets using MINTbase's various "vistas" (e.g., the "expression" vista shown in the *bottom panel*). Vistas allow users to search MINTbase using a variety of tRNA and tRF attributes such as specific isodecoders or isoacceptors, specific tRF structural type, minimum RPM expression levels, tissue(s) in which a tRF has been found, etc.

4 Notes

4.1 Dealing with Multiple Alignments and Potentially False-Positive tRFs

MINTmap profiles are created at the level of a tRF molecule. That is, the abundance of each tRF (and the tRF's respective RPM values) are reported independent of the number of locations within tRNA space where the tRF can be found. A molecule-centered approach prevents multi-counting and the misrepresentation of read counts. For a tRF molecule that may have multiple potential origins, all instances in tRNA space are reported for it. For example, referring again to the output-MINTmap_v1-exclusive-tRFs.expression.html (*see* above), the 3′-tRF with sequence ACCGGGCGGAAACACCA is the second most abundant tRF. This sequence is exclusive to tRNA space but could have

arisen from one or more of 14 different genomic locations all of which are enumerated in the last column.

Analogously, in the generated output, we include the same type of information for ambiguous tRFs in the file output-MINTmap_v1-ambiguous-tRFs.expression.html. For the above example, the most abundant ambiguous tRF is a 5′-tRF with sequence GCATTGGTGGTTCAGTGGTAGA. It is important to note here that the final column of the ambiguous-tRF table lists only the locations *within* tRNA space (some ambiguous tRFs can have numerous genomic instances, which would make their inclusion unwieldy). Specifically, this sequence appears in the true tRNA space a total of 10 times (listed in the Table) as well as has additional instances outside of tRNA space that are *not* listed in the Table.

4.2 When Will a User Need Custom Annotations?

MINTmap comes with prebuilt lookup tables and sequence annotations for mapping human short RNA-seq datasets to the 640 tRNA loci mentioned above. Thus, in most cases, the default choices for the –s, –l, and –o parameters will suffice and need not be explicitly specified when using MINTmap. These 640 loci are based on the GRCh37 assembly. This choice is dictated by the fact that 610 of these loci are drawn from version v1.0 (also based on GRCh37) of the gtRNAdb repository. As we described above, v2.0 of gtRNAdb, which is based on GRCh38, is not adequately stable for the purposes of analyzing tRF profiles.

It is conceivable that there may be cases in which the user wishes to specify her or his own lookup tables and sequences. Such cases could arise in situations where:

- The user wishes to leverage a different set of tRNAs, such as newly discovered tRNAs.
- The user wishes to use a different assembly of the human genome.
- The user wishes to create tRF profiles for a species other than *H. sapiens*, e.g., *M. musculus*.

MINTmap provides users with the ability to pass user-specified lookup and tRNA sequence files, all of which were described in the previous section. We already published very detailed steps on how to create a tRF lookup table for use with MINTmap [16]. Here, we summarize the key points.

As is evident from the previous section, the –s, –l, and –o parameters are closely intertwined. Thus, if one switches, e.g., the tRNA annotations (by using the parameter –s), the files pointed to by the other two parameters (–l and –o) must also be modified. As an example, let us consider the simple case where a newly discovered tRNA is added to the true tRNA space. Such an addition is likely to give rise to

more candidate tRF sequences that are not already in the default tRF lookup table that is used when the –l argument is not defined. These new candidate tRFs will need to be added to the lookup table and the new lookup table passed to MINTmap using the "–l" argument.

These additions to the space of possible tRFs would also require the regeneration of the exclusivity flags (Y/N) *across all candidate tRFs*, even those previously contained in the lookup table. Indeed, it is entirely possible that some tRFs were previously listed as "ambiguous" because they only had a non-tRNA space hit inside the sequence that is now the newly discovered tRNA. Similarly, if tRNA sequences are removed or edited (e.g., because of updated endpoints), the total list of candidate tRFs could change; this would prompt a regeneration of all exclusivity flags due to changes of what is considered tRNA space.

As mentioned in the previous section, the structural type lookup table (−o parameter) is optional and if the specified file does not exist, tRF profiles will still be generated (but will lack associated structural type annotations). If such information is desired, then any changes to the tRNA space would have to be reflected in changes to the structural type lookup table as well: indeed, the addition or removal of tRNA loci has the potential to change the structural type(s) of those tRFs that are present in the tRNA that was changed. For any newly added tRNAs, feature annotations (e.g., anticodon location) are required to properly characterize a tRF's structural type and generate the relevant entries in the structural type lookup table. If anticodon location information is not available, a tool such as the ViennaRNA package [40] could be used to predict the tRNA's secondary structure and subsequently estimate the location and identity of the anticodon.

5 Conclusion

We presented a detailed discussion of how one can use MINTmap, a tool we developed for generating tRF profiles. MINTmap takes as its input a file containing quality- and adaptor-trimmed reads generated by a deep-sequencing experiment and produces exhaustive profiles of the tRFs that were present in the original RNA pool. MINTmap does not need to access the Internet and operates locally thereby ensuring the confidentiality of the user data. The output of MINTmap includes a plaintext file and an HTML-formatted file. The plaintext file is tab-separated and is meant to enable further studies of the mined tRFs by the user with the help of standard software packages such as Excel, Matlab, or R. We hope that MINTmap will prove helpful to researchers who are involved in tRF studies.

Acknowledgments

The project was supported in part by a William Keck Foundation grant (IR), by Institutional Funds, and by a Commonwealth of Pennsylvania grant SAP#4100062221 (IR).

References

1. Grewal SS (2015) Why should cancer biologists care about tRNAs? tRNA synthesis, mRNA translation and the control of growth. Biochim Biophys Acta 1849(7):898–907. doi:10.1016/j.bbagrm.2014.12.005
2. Huang HY, Hopper AK (2016) Multiple layers of stress-induced regulation in tRNA biology. Life 6(2). doi:10.3390/life6020016
3. Orioli A (2017) tRNA biology in the omics era: stress signalling dynamics and cancer progression. Bioessays 39(3). doi:10.1002/bies.201600158
4. Phizicky EM, Hopper AK (2010) tRNA biology charges to the front. Genes Dev 24(17):1832–1860. doi:10.1101/gad.1956510
5. Shigematsu M, Kirino Y (2015) tRNA-derived short non-coding RNA as interacting partners of argonaute proteins. Gene Regul Syst Biol 9:27–33. doi:10.4137/GRSB.S29411
6. Anderson P, Ivanov P (2014) tRNA fragments in human health and disease. FEBS Lett 588(23):4297–4304. doi:10.1016/j.febslet.2014.09.001
7. Telonis AG, Loher P, Honda S, Jing Y, Palazzo J, Kirino Y, Rigoutsos I (2015) Dissecting tRNA-derived fragment complexities using personalized transcriptomes reveals novel fragment classes and unexpected dependencies. Oncotarget 6(28):24797–24822. doi:10.18632/oncotarget.4695
8. Hasler D, Meister G (2016) From tRNA to miRNA: RNA-folding contributes to correct entry into noncoding RNA pathways. FEBS Lett 590(15):2354–2363. doi:10.1002/1873-3468.12294
9. Goodarzi H, Liu X, Nguyen HC, Zhang S, Fish L, Tavazoie SF (2015) Endogenous tRNA-derived fragments suppress breast cancer progression via YBX1 displacement. Cell 161(4):790–802. doi:10.1016/j.cell.2015.02.053
10. Gebetsberger J, Wyss L, Mleczko AM, Reuther J, Polacek N (2016) A tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress. RNA Biol 0:1–10. doi:10.1080/15476286.2016.1257470
11. Alves CS, Vicentini R, Duarte GT, Pinoti VF, Vincentz M, Nogueira FT (2017) Genome-wide identification and characterization of tRNA-derived RNA fragments in land plants. Plant Mol Biol 93(1-2):35–48. doi:10.1007/s11103-016-0545-9
12. Olvedy M, Scaravilli M, Hoogstrate Y, Visakorpi T, Jenster G, Martens-Uzunova ES (2016) A comprehensive repertoire of tRNA-derived fragments in prostate cancer. Oncotarget 7(17):24766–24777. doi:10.18632/oncotarget.8293
13. Telonis AG, Loher P, Kirino Y, Rigoutsos I (2016) Consequential considerations when mapping tRNA fragments. BMC Bioinformatics 17:123. doi:10.1186/s12859-016-0921-0
14. Honda S, Kirino Y (2015) Dumbbell-PCR: a method to quantify specific small RNA variants with a single nucleotide resolution at terminal sequences. Nucleic Acids Res 43(12):e77. doi:10.1093/nar/gkv218
15. Veneziano D, Di Bella S, Nigita G, Lagana A, Ferro A, Croce CM (2016) Noncoding RNA: current deep sequencing data analysis approaches and challenges. Hum Mutat 37(12):1283–1298. doi:10.1002/humu.23066
16. Loher P, Telonis AG, Rigoutsos I (2017) MINTmap: fast and exhaustive profiling of nuclear and mitochondrial tRNA fragments from short RNA-seq data. Sci Rep 7:41184. doi:10.1038/srep41184
17. Telonis AG, Loher P, Kirino Y, Rigoutsos I (2014) Nuclear and mitochondrial tRNA-lookalikes in the human genome. Front Genet 5:344. doi:10.3389/fgene.2014.00344
18. Bao W, Kojima KK, Kohany O (2015) Repbase update, a database of repetitive elements in eukaryotic genomes. Mob DNA 6(1):11. doi:10.1186/s13100-015-0041-9
19. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25(5):955–964
20. Hou YM (2010) CCA addition to tRNA: implications for tRNA quality control. IUBMB Life 62(4):251–260. doi:10.1002/iub.301
21. Heinemann IU, Nakamura A, O'Donoghue P, Eiler D, Soll D (2012) tRNAHis-guanylyl-

transferase establishes tRNAHis identity. Nucleic Acids Res 40(1):333–344. doi:10.1093/nar/gkr696

22. Heinemann IU, Randau L, Tomko RJ Jr, Soll D (2010) 3′-5′ tRNAHis guanylyltransferase in bacteria. FEBS Lett 584(16):3567–3572. doi:10.1016/j.febslet.2010.07.023
23. Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A 85(8):2444–2448
24. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2
25. States DJ, Gish W (1994) Combined use of sequence similarity and codon bias for coding region identification. J Comput Biol 1(1):39–50. doi:10.1089/cmb.1994.1.39
26. Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12(4):656–664. doi:10.1101/gr.229202. Article published online before March 2002
27. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25. doi:10.1186/gb-2009-10-3-r25
28. Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25(14):1754–1760. doi:10.1093/bioinformatics/btp324
29. Homer N, Merriman B, Nelson SF (2009) BFAST: an alignment tool for large scale genome resequencing. PLoS One 4(11):e7767. doi:10.1371/journal.pone.0007767
30. David M, Dzamba M, Lister D, Ilie L, Brudno M (2011) SHRiMP2: sensitive yet practical SHort read mapping. Bioinformatics 27(7):1011–1012. doi:10.1093/bioinformatics/btr046
31. Chan PP, Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res 37(Database issue):D93–D97. doi:10.1093/nar/gkn787
32. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17(1):10–12. doi:10.14806/ej.17.1.200
33. Mayr C, Bartel DP (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138(4):673–684. doi:10.1016/j.cell.2009.06.016
34. Chan PP, Lowe TM (2016) GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic Acids Res 44(D1):D184–D189. doi:10.1093/nar/gkv1309
35. Rigoutsos I (2016) Comment on PMID 26673694:GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. In: PubMed Commons [Internet]. National Library of Medicine, Bethesda, MD. http://www.ncbi.nlm.nih.gov/pubmed/26673694-cm26673694_13813
36. Rao BS, Jackman JE (2015) Life without post-transcriptional addition of G-1: two alternatives for tRNAHis identity in Eukarya. RNA 21(2):243–253. doi:10.1261/rna.048389.114
37. Pliatsika V, Loher P, Telonis AG, Rigoutsos I (2016) MINTbase: a framework for the interactive exploration of mitochondrial and nuclear tRNA fragments. Bioinformatics 32(16):2481–2489. doi:10.1093/bioinformatics/btw194
38. Kumar P, Mudunuri SB, Anaya J, Dutta A (2015) tRFdb: a database for transfer RNA fragments. Nucleic Acids Res 43(Database issue):D141–D145. doi:10.1093/nar/gku1138
39. Zheng LL, Xu WL, Liu S, Sun WJ, Li JH, Wu J, Yang JH, Qu LH (2016) tRF2Cancer: a web server to detect tRNA-derived small RNA fragments (tRFs) and their expression in multiple cancers. Nucleic Acids Res 44(W1):W185–W193. doi:10.1093/nar/gkw414
40. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA package 2.0. Algorithms Mol Biol 6:26. doi:10.1186/1748-7188-6-26

INDEX

Katsutomo Okamura and Kotaro Nakanishi (eds.), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1680, DOI 10.1007/978-1-4939-7339-2, © Springer Science+Business Media LLC 2018

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

W

X

Y

Z