Methods in Molecular Biology 1640

Springer Protocols

Alberto Carbonell *Editor*

Plant 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

Plant Argonaute Proteins

Methods and Protocols

Edited by

Alberto Carbonell

Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas–Universidad Politécnica de Valencia), Valencia, Spain

Humana Press

Editor
Alberto Carbonell
Instituto de Biología Molecular y Celular de Plantas
(Consejo Superior de Investigaciones Científicas–Universidad Politécnica de Valencia)
Valencia, Spain

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-7164-0 ISBN 978-1-4939-7165-7 (eBook)
DOI 10.1007/978-1-4939-7165-7

Library of Congress Control Number: 2017942942

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 almost two decades ago of gene silencing phenomena related to RNA interference (RNAi), intensive research has highlighted the importance of ARGONAUTE (AGO) proteins as central effectors of RNAi pathways. In eukaryotes, AGOs associate with small RNAs (sRNAs) to direct gene silencing and regulate key biological processes such as development, response to stress, epigenetics, and antiviral defense. In plants, *Arabidopsis thaliana* has been a particularly valuable model organism to study AGO roles because of the functional diversification of the ten AGOs encoded in its genome, among other reasons. During the last years, the main molecular and biological functions of plant AGOs have been characterized. However, as occurred in human and insects, it is possible that new functions will be discovered soon for plant AGOs.

The purpose of this book is to provide the reader with step-by-step methods to study plant AGO functions. After an introductory review chapter (Chapter 1), the book summarizes the main biochemical methods to study AGO–sRNA complexes (Chapters 2–5) and their interaction with target RNAs (Chapters 6–7), AGO subcellular localization (Chapter 8), AGO association with polysomes (Chapter 9), and AGO role in meiosis and DNA repair (Chapter 10). Next, methods for the identification, cloning, and characterization of *AGO* genes in different plant species are presented (Chapters 11–13), as well as nonradioactive protocols for sRNA detection (Chapters 14–15). Finally, a series of chapters describing computational methods to study plant AGO function and evolution are provided (Chapters 16–20).

My motivation to edit this *Methods in Molecular Biology* volume was to provide the most complete and updated list of protocols to study plant AGO function. Unfortunately, it has not been possible to cover all the contributions of plant AGOs nor to avoid some overlaps between chapters. Therefore, I deeply apologize to those readers who may regret such omissions or redundancies.

I am especially thankful to Prof. Jim Carrington for giving me the opportunity to study the biological roles of plant AGOs during my postdoc in his lab these last years. I would also like to thank all authors of the chapters for their effort and commitment to the project and for providing such high-quality manuscripts. Finally, I am also thankful to the series editor Prof. John Walker and the Springer staff for their support, help, and guidance.

Valencia, Spain *Alberto Carbonell*

Contents

Preface . . . . . *v*
Contributors . . . . . *ix*

1 Plant ARGONAUTEs: Features, Functions, and Unknowns 1
Alberto Carbonell

2 Analysis of the Uridylation of Both ARGONAUTE-Bound MiRNAs and 5′ Cleavage Products of Their Target RNAs in Plants 23
Guodong Ren, Xiaoyan Wang, and Bin Yu

3 In Vitro Formation of Plant RNA-Induced Silencing Complexes Using an Extract of Evacuolated Tobacco Protoplasts 39
Taichiro Iki, Masayuki Ishikawa, and Manabu Yoshikawa

4 In Vitro Analysis of ARGONAUTE-Mediated Target Cleavage and Translational Repression in Plants 55
Yukihide Tomari and Hiro-oki Iwakawa

5 Establishment of an In Vivo ARGONAUTE Reporter System in Plants 73
Károly Fátyol and József Burgyán

6 Immunoprecipitation and High-Throughput Sequencing of ARGONAUTE-Bound Target RNAs from Plants 93
Alberto Carbonell

7 Identification of ARGONAUTE/Small RNA Cleavage Sites by Degradome Sequencing 113
Ivett Baksa and György Szittya

8 Dissecting the Subnuclear Localization Patterns of Argonaute Proteins and Other Components of the RNA-Directed DNA Methylation Pathway in Plants 129
Cheng-Guo Duan and Jian-Kang Zhu

9 Detection of Argonaute 1 Association with Polysomes in *Arabidopsis thaliana* 137
Cécile Lecampion, Elodie Lanet, and Christophe Robaglia

10 Functional Analysis of Arabidopsis ARGONAUTEs in Meiosis and DNA Repair 145
Marina Martinez-Garcia and Mónica Pradillo

11 Isolation and Characterization of *ARGONAUTE* Mutants in *Chlamydomonas* 159
Tomohito Yamasaki

12 *ARGONAUTE* Genes in *Salvia miltiorrhiza*: Identification, Characterization, and Genetic Transformation 173
Meizhen Wang, Yuxing Deng, Fenjuan Shao, Miaomiao Liu, Yongqi Pang, Caili Li, and Shanfa Lu

13 Cloning and Characterization of *Argonaute* Genes in Tomato 191
Zhiqiang Xian, Fang Yan, and Zhengguo Li

14 Nonradioactive Detection of Small RNAs Using Digoxigenin-Labeled Probes 199
Ariel H. Tomassi, Delfina Gagliardi, Damian A. Cambiagno, and Pablo A. Manavella

15 Nonradioactive Plant Small RNA Detection Using Biotin-Labeled Probes 211
Jun Hu and Yingguo Zhu

16 Computational Analysis of Genome-Wide ARGONAUTE-Dependent DNA Methylation in Plants 219
Kai Tang, Cheng-Guo Duan, Huiming Zhang, and Jian-Kang Zhu

17 Structural and Functional Characterization of Plant ARGONAUTE MID Domains 227
Filipp Frank and Bhushan Nagar

18 Identification and Analysis of WG/GW ARGONAUTE-Binding Domains 241
Andrzej Zielezinski and Wojciech M. Karlowski

19 Identification and Evolutionary Characterization of ARGONAUTE-Binding Platforms 257
Joshua T. Trujillo and Rebecca A. Mosher

20 Phylogenetic and Evolutionary Analysis of Plant ARGONAUTES 267
Ravi K. Singh and Shree P. Pandey

Index *295*

Contributors

Ivett Baksa • *National Agricultural Research and Innovation Centre, Agricultural Biotechnology Institute, Epigenetics Group, Gödöllő, Hungary*

József Burgyán • *Agricultural Biotechnology Institute, National Agricultural Research and Innovation Centre, Gödöllő, Hungary*

Damian A. Cambiagno • *Instituto de Agrobiotecnología del Litoral (UNL–CONICET–FBCB), Santa Fe, Argentina*

Alberto Carbonell • *Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas–Universidad Politécnica de Valencia), Valencia, Spain*

Yuxing Deng • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*

Cheng-Guo Duan • *Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China; Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA*

Károly Fátyol • *Agricultural Biotechnology Institute, National Agricultural Research and Innovation Centre, Gödöllő, Hungary*

Filipp Frank • *Department of Biochemistry, Emory University, Atlanta, GA, USA; Department of Chemistry, Emory University, Atlanta, GA, USA*

Delfina Gagliardi • *Instituto de Agrobiotecnología del Litoral (UNL–CONICET–FBCB), Santa Fe, Argentina*

Jun Hu • *State Key Laboratory for Hybrid Rice, College of Life Sciences, Wuhan University, Hubei, China*

Taichiro Iki • *Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan*

Masayuki Ishikawa • *Division of Plant and Microbial Sciences, Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, Tsukuba, Ibaraki, Japan*

Hiro-oki Iwakawa • *Institute of 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, Tokyo, Japan*

Wojciech M. Karlowski • *Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland*

Elodie Lanet • *Laboratoire de Génétique et Biophysique des Plantes, Aix Marseille Univ, Marseille, France; CNRS, UMR 7265 Biologie Végétale et Microbiologie Environnementale, Marseille, France; CEA, BIAM, Marseille, France*

Cécile Lecampion • *Laboratoire de Génétique et Biophysique des Plantes, Aix Marseille Univ, Marseille, France; CNRS, UMR 7265 Biologie Végétale et Microbiologie Environnementale, Marseille, France; CEA, BIAM, Marseille, France*

Caili Li • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*

Zhengguo Li • *School of Life Sciences, Chongqing University, Chongqing, China*

Miaomiao Liu • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*
Shanfa Lu • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*
Pablo A. Manavella • *Instituto de Agrobiotecnología del Litoral (UNL–CONICET–FBCB), Santa Fe, Argentina*
Marina Martinez-Garcia • *School of Biosciences, University of Birmingham, Birmingham, UK*
Rebecca A. Mosher • *School of Plant Sciences, The University of Arizona, Tucson, AZ, USA*
Bhushan Nagar • *Department of Biochemistry and Groupe de Recherche Axé sur la Structure des Protéines, McGill University, Montreal, QC, Canada*
Shree P. Pandey • *Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Nadia, West Bengal, India*
Yongqi Pang • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*
Mónica Pradillo • *Departamento de Genética, Facultad de Biología, Universidad Complutense de Madrid, Madrid, Spain*
Guodong Ren • *State Key Laboratory of Genetic Engineering and Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China*
Christophe Robaglia • *Laboratoire de Génétique et Biophysique des Plantes, Aix Marseille Univ, Marseille, France; CNRS, UMR 7265 Biologie Végétale et Microbiologie Environnementale, Marseille, France; CEA, BIAM, Marseille, France*
Fenjuan Shao • *State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China*
Ravi K. Singh • *Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Nadia, West Bengal, India*
György Szittya • *National Agricultural Research and Innovation Centre, Agricultural Biotechnology Institute, Epigenetics Group, Gödöllő, Hungary*
Kai Tang • *Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China; Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA*
Yukihide Tomari • *Institute of 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, Tokyo, Japan*
Ariel H. Tomassi • *Instituto de Agrobiotecnología del Litoral (UNL–CONICET–FBCB), Santa Fe, Argentina*
Joshua T. Trujillo • *Biochemistry, Molecular and Cellular Biology Program, The University of Arizona, Tucson, AZ, USA*
Meizhen Wang • *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China*
Xiaoyan Wang • *State Key Laboratory of Genetic Engineering and Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China*
Zhiqiang Xian • *School of Life Sciences, Chongqing University, Chongqing, China*
Tomohito Yamasaki • *Science Department, Natural Science Cluster, Kochi University, Kochi, Japan*
Fang Yan • *School of Life Sciences, Chongqing University, Chongqing, China*

Manabu Yoshikawa • *Division of Plant and Microbial Sciences, Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, Tsukuba, Ibaraki, Japan*

Bin Yu • *Center for Plant Science Innovation and School of Biological Sciences, University of Nebraska–Lincoln, Lincoln, NE, USA*

Huiming Zhang • *Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China*

Jian-Kang Zhu • *Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China; Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA*

Yingguo Zhu • *State Key Laboratory for Hybrid Rice, College of Life Sciences, Wuhan University, Hubei, China*

Andrzej Zielezinski • *Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland*

Chapter 1

Plant ARGONAUTEs: Features, Functions, and Unknowns

Alberto Carbonell

Abstract

ARGONAUTEs (AGOs) are the effector proteins in eukaryotic small RNA (sRNA)-based gene silencing pathways controlling gene expression and transposon activity. In plants, AGOs regulate key biological processes such as development, response to stress, genome structure and integrity, and pathogen defense. Canonical functions of plant AGO–sRNA complexes include the endonucleolytic cleavage or translational inhibition of target RNAs and the methylation of target DNAs. Here, I provide a brief update on the major features, molecular functions, and biological roles of plant AGOs. A special focus is given to the more recent discoveries related to emerging molecular or biological functions of plant AGOs, as well as to the major unknowns in the plant AGO field.

Key words ARGONAUTE, Small RNA, RNA silencing, MicroRNA, *Arabidopsis*

1 Introduction

In eukaryotes, ARGONAUTEs (AGOs) are the effector proteins functioning in small RNA (sRNA)-guided gene silencing pathways regulating gene expression and transposon activity [1]. AGO–sRNA complexes target and silence complementary DNA or RNA through posttranscriptional gene silencing (PTGS) or transcriptional gene silencing (TGS), respectively. Silencing of target transcripts occurs either through direct endonucleolytic cleavage (slicing) or through other cleavage-independent mechanisms such as target destabilization or translational repression [2].

AGOs have an ancient origin, as they are present in bacteria, archaea, and eukaryotes [3]. In plants, the AGO family includes a variable number of members depending on the plant species [4], with flowering plants encoding more AGOs. For example, *Arabidopsis thaliana* (*Arabidopsis*) and rice have 10 and 19 AGO members [5, 6], respectively, while the algae *Chlamydomonas reinhardtii* and the moss *Physcomitrella patens* have three and six [7–9], respectively. The expansion of the plant AGO family suggests a functional diversification of AGO proteins most likely during the

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_1,

specialization and evolution of endogenous sRNA-based RNA silencing pathways [10, 11]. Phylogenetically, flowering plant AGOs can be grouped in three major clades: AGO1/5/10, AGO2/3/7, and AGO4/6/8/9. In addition, grasses present an expanded AGO1/5/10 clade including AGO18 [10].

Crystallographic studies on eukaryotic AGOs have determined that AGOs present four functional domains: a variable N-domain and conserved PAZ, MID, and PIWI domains [12]. The MID and PAZ domains bind the 5′ monophosphorylated nucleotide and the 3′ nucleotide of the sRNA, respectively. The PIWI domain is the ribonucleolytic domain, with four metal-coordinating residues required for slicer activity [13, 14]. Plant AGOs associate with sRNA based on the identity of the 5′ nucleotide of the sRNA and/or other sequence and structural features of the sRNA duplex and the AGO PIWI domain [15–19]. Plant AGO–sRNA complexes can function through different modes to silence complementary DNA or RNA and exert their biological role.

I present next an updated overview on the known and emerging molecular and biological roles of plant AGOs. I also highlight the main unknowns in the plant AGO field.

2 Modes of Action of Plant AGOs

The main modes of action of plant AGOs are summarized in Fig. 1 and described next.

2.1 Endonucleolytic Cleavage

The PIWI domain of AGOs uses intrinsic RNAse H-like activity to cleave target RNA [20] and contains a metal-coordinating Asp–Glu–Asp–His/Asp catalytic tetrad [14, 21]. Slicing activity has been experimentally confirmed for *Arabidopsis* AGO1 [22, 23], AGO2 [24], AGO4 [25], AGO7 [15], and AGO10 [18, 26].

Since the initial observation that plant microRNAs (miRNAs) targeted and cleaved highly sequence complementary target RNAs [27], it was assumed that slicing was the predominant mode of action in miRNA-mediated PTGS in plants [28]. Indeed, the high degree of complementarity is a requirement for effective target slicing by plant AGOs [29]. However, later examples describing slicing-independent translational repression of certain miRNA targets (see below) have questioned this assumption. Still, growing evidence suggests that target RNA regulation by slicing is widespread in plants. First, sequencing of *Arabidopsis* mRNA degradome has revealed that most miRNA targets undergo slicing [30, 31]. Second, the slicing activity of *Arabidopsis* AGO1 (the primary miRNA-associating AGO), AGO2, and AGO7 is critical for plant development, antiviral activity, and juvenile to adult phase transition, respectively [24], while AGO4 and AGO10 exert their primary functions in a slicer-independent

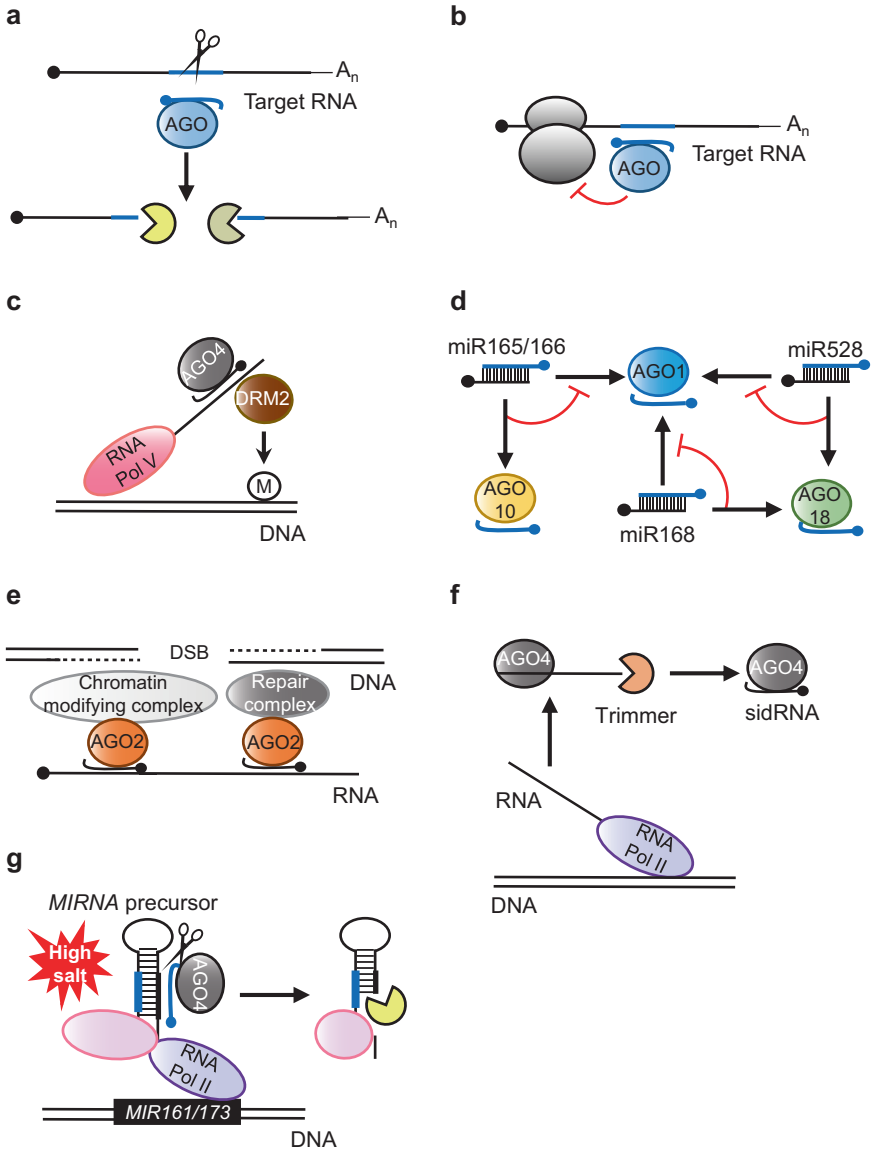


Fig. 1 Modes of action of plant AGOs. (**a**) Endonucleolytic cleavage. Several AGOs bind sRNAs and slice highly sequence complementary target RNAs. Cleavage products are degraded by components of endogenous degradation pathways. (**b**) Translational repression. Plant AGOs such as *Arabidopsis* AGO1 and AGO10 associate with miRNAs and target highly complementary RNAs to inhibit their translation. (**c**) Canonical RdDM pathway. AGO4–siRNA complexes bind to nascent Pol V transcripts. DRM2 is recruited to mediate DNA methylation. (**d**) MiRNA sequestration. *Arabidopsis* AGO10 sequesters miR165/166 from AGO1, while rice AGO18 sequesters miR168 and miR528 from AGO1. (**e**) Double-stranded break repair (DSB). *Arabidopsis* AGO2 binds to DSB-induced siRNAs (diRNAs) to mediate DSB repair. (**f**) DCL-independent siRNA biogenesis. AGO4 binds to nascent Pol II RNAs which are trimmed by a 3′–5′ exonuclease to produce sidRNAs. (**g**) Cotranscriptional regulation of *MIRNA* gene expression. Upon salinity stress AGO4 directs the slicing of nascent *MIR161* or *MIR173* precursors to cotranscriptionally regulate miRNA production

mode [18, 25]. And third, recent transcriptome profiling of *ago1* null and slicer-deficient *Arabidopsis* mutants confirmed that AGO1 slicer activity is necessary for the repression of the majority of miRNA targets [32].

Slicing activity of plant AGOs is also required for triggering the amplification of phased secondary small interfering RNAs (phasiRNAs) from certain target transcripts. For instance, *trans*-acting small interfering RNAs (tasiRNAs), a class of secondary small interfering RNAs (siRNAs) that forms through a refined mechanism, derive from four families of *TAS* transcripts that are initially cleaved by AGO1–miR173, AGO1–miR828, or AGO7–miR390 complexes to produce *TAS1-/TAS2-*, *TAS4-*, and *TAS3*-derived tasiRNAs, respectively [15, 24, 33–37]. Interestingly, a recent comparative analysis of tasiRNA generation in wild-type, *ago1* null, and *ago1* slicer-deficient *Arabidopsis* showed that slicing by AGO1 is required for the definition of the phase but not for the generation of *TAS*-derived tasiRNAs [38].

The subcellular location of miRNA-mediated target cleavage has been largely unknown. However, the observation that a reduced level of isoprenoids, which are essential for membrane sterols, blocks miRNA-mediated cleavage of several target transcripts suggests that membrane association of AGO1 is important for target cleavage [39]. In a recent work, *Arabidopsis* miRNAs were shown to associate with membrane-bound polysomes as opposed to polysomes in general, and this association was required for miRNA-triggered phasiRNA production [40]. Because slicing is required for phasiRNA production, it appears that, at least, part of the AGO-mediated target cleavage activity occurs in membrane-bound polysomes.

2.2 Translational Repression

Translational repression by miRNAs is common in animals where miRNA–target RNA interactions require limited sequence complementarity [41]. In plants, several evidences suggest that AGO–miRNA complexes can also translationally repress their target RNAs with almost perfect complementarity [39, 42–48]. Several *Arabidopsis* mutants impaired in miRNA-mediated gene repression at protein but not at mRNA levels have been described [39, 43, 49]. In particular, AGO1–miRNA-mediated translational repression in *Arabidopsis* occurs in the endoplasmic reticulum and requires the integral membrane protein ALTERED MERISTEM PROGRAM 1 to exclude target mRNAs from membrane-bound polysomes [43]. AGO10, another member of the AGO1 clade, also appears to translationally repress several *Arabidopsis* miRNA target genes, including AGO1 [50]. Very recently, it has been shown that AGO7–miR390 binding to a non-cleavable miR390 target site included in *TAS3a* noncoding transcripts cause ribosome stacking and subsequent inhibition of translation elongation [51]. However, the global contribution to plant miRNA-mediated

translational repression of direct blocking of ribosome movement through the binding of AGO–miRNA complexes appears to be limited [51].

2.3 Target mRNA Decay?

Target mRNA decay is a common outcome of AGO recruitment in nimals, where the complementarity between an amiRNA and its target mRNA is generally limited to the 5′ region of the miRNA and slicing is not common [52]. AGO–miRNA complexes destabilize target mRNAs in a process requiring both deadenylation and decapping [53, 54]. Recruitment of the two major deadenylases CCR4–NOT and PARN requires the adaptor protein GW182/TNRC6 that binds to hydrophobic pockets in AGOs [55, 56]. Because no homolog of GW182/TNRC6 is present in plant genomes, it is unlikely that such mechanism exists in plants. Three recent studies support this statement: (1) transcriptome analyses performed on either stable or conditional slicer-deficient AGO1 mutants did not show any substantial differences in gene expression between the two classes [32], (2) efficient translational repression by either wild-type or slicer-deficient AGO1 in lysates of tobacco protoplasts is not accompanied by any kind of reporter–mRNA degradation other than slicing [42], and (3) no deadenylation has been reported in sRNA-mediated translational repression in *Chlamydomonas reinhardtii* [57]. However, the possibility that a subset of plant miRNA targets could be regulated by AGO-mediated mRNA decay cannot be completely ruled out.

2.4 RNA-Dependent DNAMethylation

DNA methylation regulates gene expression, blocks transposon movement, and consequently maintains genome integrity. In plants, canonical DNA methylation is primarily mediated by AGO4–siRNA complexes functioning in RNA-dependent DNA methylation (RdDM) pathways [58]. These pathways are initiated by the synthesis of double-stranded RNA (dsRNA) by the concerted action of RNA polymerase IV (Pol IV) and RNA-dependent RNA polymerase 2 (RDR2) [59–63]. dsRNA processing in the nucleus by dicer-like 3 (DCL3) leads to the production of 24-nt siRNAs [59] that are exported to the cytoplasm where they are incorporated into AGO4. AGO4–siRNA complexes localize to the nucleus where they are recruited to target loci via base pairing with nascent Pol V transcripts and/or through their interaction with the glycine–tryptophane/tryptophane–glycine (GW/WG) AGO hook motifs present in both Pol V [64–67] and its associated factor suppressor of TY insertion 5 (SPT5) [68, 69]. Finally, AGO4–siRNA complexes recruit domain-rearranged methyltransferase 2 (DRM2) protein that methylates target DNA [70, 71]. Very recently, AGO4 interaction with DNA has been observed at RdDM targets. It appears that Pol V-dependent transcripts or their transcription are needed to lock Pol V into a stable DNA-bound configuration that allows AGO4 recruitment via Pol V and SPT5 AGO hook motifs [72].

AGO6 is also associated with RdDM in *Arabidopsis* and thought initially to have partially redundant functions with AGO4 [73]. Later studies assigned more specific features and functions for AGO6, such as its preferential association for a unique set of heterochromatic sRNAs [74] or its dominant expression in shoot and root apical meristems and not in mature leaves [75]. More recently, it has been proposed that AGO6 may indeed work sequentially with AGO4 in the methylation of most target loci [76]. In addition, AGO6 also associates with RDR6-dependent 21–22-nt sRNAs to direct the methylation of transcriptionally active transposons in *Arabidopsis* [77].

Unexpectedly, a combination of genetic, biochemical, and bioinformatic genome-wide analyses has recently showed that *Arabidopsis* AGO3, thought to function in PTGS as the other members of the AGO2/AGO7 clade, binds 24-nt sRNAs and can partially complement AGO4 function. The authors speculate with a role of AGO3 in RdDM in *Arabidopsis* [78], possibly under salinity stress when its expression is highly induced.

2.5 Emerging AGO Functions

Besides the well-characterized roles of plant AGOs in sRNA-mediated PTGS and TGS, new molecular functions have been described in the last years. First, *Arabidopsis* AGO10 and rice AGO18 sequester miR165/166 and miR168 from AGO1 to regulate shoot apical meristem (SAM) development [18] and antiviral defense [79], respectively. Rice AGO18 additionally sequesters miR528 from AGO1 upon viral infection to inhibit L-ascorbate oxidase (AO) mRNA cleavage by AGO1–miR528 complexes, thereby increasing AO-mediated accumulation of reactive oxygen species and enhancing antiviral defense [80]. Second, *Arabidopsis* AGO2 and AGO9 participate in the repair of double-strand break sites [81, 82]. Third, AGO4 participates in an alternative siRNA biogenesis pathway by binding precursor transcripts that are subsequently subjected to 3′–5′ exonucleolytic trimming for maturation and sidRNA (siRNA independent of DCLs) production [83]. AGO4–sidRNA complexes target Pol V transcripts to mediate DRM2-dependent DNA methylation. And fourth, a novel role for *Arabidopsis* AGO1 in the cotranscriptional regulation of *MIRNA* gene expression under salt stress conditions has been recently reported [84]. It seems that miRNA-loaded AGO1 interacts with chromatin at *MIR161* and *MIR173* loci, causing the disassembly of the transcriptional complex and the release of short and unpolyadenylated transcripts [84].

3 Biological Roles of Plant AGO Proteins

Plant AGOs have functionally diversified during evolution due to the expansion of the AGO family because of numerous duplications and losses [10, 85, 86]. The main biological roles of plant AGOs are listed in Table 1 and described next (*see* [4] for a recent review).

Table 1
Biological roles of plant ARGONAUTEs

Function	AGO involved[a]	References
Antibacterial immunity	AtAGO2 AtAGO4	[108] [109]
Antiviral defense	AtAGO1 AtAGO2 AtAGO4 AtAGO5 AtAGO7 AtAGO10 NbAGO1 NbAGO2 OsAGO1a/b OsAGO18	[6, 136–139] [17, 24, 136, 138, 140–143] [107, 142, 144–147] [136, 148] [136, 137] [136] [149] [150–152] [79] [79]
Cell specification Gamete Somatic	AtAGO9 ZmAGO9	[114] [113]
Chromosome segregation	ZmAGO9	[113]
Development	AtAGO1 OsAGO1a/b/c SiAGO1b	[22–24, 87, 153] [92] [154]
DNA methylation	AtAGO3 AtAGO4 AtAGO6 OsAGO4a OsAGO4b	[78] [25, 64, 65, 72, 155, 156] [73, 74, 76, 77, 156, 157] [158] [158]
DNA repair	AtAGO2 AtAGO9	[81] [82]
Germ cell development	ZmAGO18b	[159]
Leaf development	AtAGO1 AtAGO7 OsAGO10 ZmAGO7	[87, 88] [95, 160] [103] [100]
Meiosis	AtAGO4 OsAGO5c	[161] [112]
Megagametogenesis	AtAGO5	[111]
Phase transition	AtAGO7	[15, 24, 94, 95]
SAM development	AtAGO10 OsAGO7	[18, 26, 101, 102, 162, 163] [99]
SAM maintenance	OsAGO10	[103]

(continued)

Table 1
(continued)

Function	AGO involved[a]	References
Small RNA biogenesis		
miRNA	AtAGO1	[84]
siRNA	AtAGO4	[83]
tasiRNA	AtAGO1	[34]
	AtAGO7	[15]
	OsAGO7	[99]
	ZmAGO7	[100]
Stress response	AtAGO1	[164, 165]
	SiAGO1b	[154]
Tapetum development	ZmAGO18b	[159]

[a]At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Os, *Oryza sativa*; Si, *Setaria italica*; Zm, *Zea mays*

3.1 Plant AGOs and Development

The importance of AGOs in plant development became obvious after the characterization of the first *Arabidopsis ago1* mutants. These mutants—named "ARGONAUTE" because of the resemblance of their leaf defects with the tentacles of a small squid of the *Argonauta* genus—presented important pleiotropic developmental defects such as dwarfing and sterility [87]. Later, developmental screens in *Arabidopsis* identified a series of hypomorphic *ago1* alleles with reduced developmental defects. The characterization of such mutants highlighted AGO1 role in leaf polarity and lateral organ development [88–90]. The organ polarity defects exhibited by *ago1* mutants suggested that AGO1 plays a role in the miRNA pathway, as these defects were similar to those of *phabulosa* (*phb*) and *phavoluta* (*phv*) miRNA gain-of-function mutants [91]. *Ago1* mutants have also been characterized in rice and show obvious pleiotropic developmental defects such as severe dwarfism, tortuous shoots, narrow and rolled leaves, and low seed-setting rates [92].

Other *Arabidopsis ago* mutants such as *ago7* or *ago10* present limited developmental defects, and others like *ago2*, *ago3*, *ago4*, *ago5*, *ago6*, and *ago9* have no obvious growth-related phenotypes [93]. AGO7 was identified in a screen for mutants displaying accelerated juvenile to adult phase change [94]. AGO7 associates exclusively with miR390 to target *TAS3* transcripts and initiate *TAS3*-based tasiRNA biogenesis leading to the targeting of several *auxin response factor* genes involved in the regulation of developmental timing and lateral organ development in *Arabidopsis* [15, 35, 95–98]. The observation that AGO7 also participates in *TAS3*-dependent tasiRNA biogenesis in moss [97] and in monocot species such as rice [99] and maize [100] indicates that AGO7 function in tasiRNA biogenesis is deeply conserved in plants.

Arabidopsis AGO10 mutants (previously known as *phn* from "pinhead" and *zll* from "zwille") exhibit abnormal SAM development [101, 102]. Despite that early analyses of *Arabidopsis ago1ago10* double mutants revealed functional redundancies between the two AGOs in some aspects of development [101], later observations have assigned specific roles for AGO10. Contrary to AGO1 which is expressed ubiquitously, AGO10 is predominantly expressed in the provasculature, the adaxial leaf primordia, and the meristem [101, 102]. AGO10 expression pattern is consistent with its roles in the maintenance of SAM development and leaf development in *Arabidopsis* [101, 102] and rice [103]. More recent observations indicate that AGO10 sequesters miR165/miR166 from AGO1 to regulate SAM development [18] and associates with miR172 to favor floral determinacy [26].

3.2 Plant AGOs and Pathogen Defense

Plant AGOs play a key role in antiviral defense (for a recent review, *see* [104]). Inantiviral silencing, highly structured RNAs and/or dsRNAs of viral origin are processed by plant DCLs into 21–24-nt virus-derived siRNAs (vsiRNAs). vsiRNAs associate with specific AGOs to target and repress cognate viral RNA through endonucleolytic cleavage or translational repression or cognate viral DNA through hypermethylation or by regulating host gene expression to enhance antiviral defense [105]. Plant AGOs with roles in antiviral silencing include *Arabidopsis* AGO1, AGO2, AGO4, AGO5, AGO7, and AGO10, *N. benthamiana* AGO1 and AGO2, and rice AGO1 and AGO18. Plant AGOs can also bind sRNAs derived from viroids to attenuate viroid accumulation in vivo [106]. Interestingly, a recent report suggests that *Arabidopsis* AGO4 has direct antiviral activity against *Plantago asiatica mosaic virus* independent of its RdDM function [107].

In addition to their well-known role in antiviral defense, several *Arabidopsis* AGOs have antibacterial activity. In particular, AGO2 binds miR393b* to translationally repress the Golgi-localized *MEMB12* gene, resulting in the exocytosis of the pathogenesis-related protein PR1 with high antibacterial activity [108]. AGO4 is required for *Arabidopsis* resistance to *Pseudomonas syringae*, in a mode independent of other components of the RdDM pathway [109].

3.3 Plant AGO Functions in Meiosis and Gametogenesis

Plant AGOs have a key role during sexual reproduction, with specific AGOs being preferentially expressed in reproductive tissues and enriched in germline cells [110]. For instance, *Arabidopsis* AGO5 is expressed in the somatic cells around megaspore mother cells and in the megaspores, and *ago5* mutants are impaired in megagametogenesis initiation [111]. In rice, mutations in *meiosis arrested at leptotene 1* (*MEL1*)—one of the five AGO5 homologs in rice—induce precocious meiotic arrest and male sterility, with abnormal tapetums and aberrant pollen mother cells [112]. In maize, AGO9

is expressed in ovule somatic cells surrounding female meiocytes and contributes to non-CG DNA methylation in heterochromatin, and chromosome segregation is arrested during meiosis in *ago9* mutants [113]. Both *Arabidopsis* and maize AGO9 act in somatic cells to regulate cell fate specification in a non-cell autonomous manner. However, *Arabidopsis* AGO9 represses germ cell fate in somatic cells [114], while maize AGO9 inhibits somatic cell fate in germ cells [113].

4 Major Unknowns in the Plant AGO Field

4.1 AGO Protein Interactors

In principle, plant AGOs likely need cofactors to exert their functions. However, in contrast to the situation observed in other organisms, a limited number of proteins interacting with plant AGOs have been described to date. Known plant AGO interactors are (1) cyclophilin 40 (CYP40), heat shock protein 90 (HSP90), and transportin 1 (TRN1) which interact with AGO1 and facilitate miRNA loading [115–118] and (2) the Pol V NRPE subunit [66], the transcription elongation factor SPT5 [68, 69], and the putative oxidoreductase WGRP1 [119] which interact with AGO4 via their GW/WG AGO hook motifs. Systematic genome-wide scans for AGO protein interactors through more refined co-immunoprecipitation coupled with mass spectrometry analyses should identify a larger number of AGO partners, especially in response to abiotic or biotic stresses.

4.2 AGO Target RNAs

A fundamental requisite to understand AGO function is the identification of the whole spectrum of cellular target RNAs regulated by plant AGOs. Contrary to animal miRNAs, the majority of plant miRNAs regulate highly sequence complementary mRNAs [120]. This strict complementary feature of functionally relevant miRNA–target interactions made early bioinformatic studies highly successful in predicting miRNA targets in plants [28]. Molecular validation of numerous plant miRNA targets has relied on the amplification by 5′ rapid amplification of cDNA ends (RACE) of 3′ cleavage products from cell extracts [121]. Because the isolation of loss-of-function miRNA mutants is difficult due mainly to the genetic redundancy in most miRNA families, the biological significance of individual miRNA–target interactions has been explored by other genetic approaches. These include the overexpression of miRNAs, miRNA-resistant targets, or target mimicries [122]. The first genome-wide assessment of the repertoire of miRNA target RNAs regulated by cleavage corresponds to degradome sequencing analyses [30]. It appears that many conserved canonical targets have consistently strong degradome signatures, suggesting that this approach may be more likely to detect functionally relevant targets. Unfortunately, weak signatures are also recovered from several

conserved canonical targets, and new potential targets do not follow the canonical parameters of base pairing. Therefore, the functional significance of degradome signatures is still not always clear [122]. All these approaches are useful to confirm or discover miRNA targets but do not reveal which specific AGO member mediates their regulation.

Genome-wide analysis of AGO-bound target RNAs has been reported in animals by applying a step of in vivocross-linking (generally using ultraviolet light) in intact cells of tissues before immunoprecipitating the AGO of interest and analyzing by high-throughput sequencing the co-immunoprecipitated AGO-bound RNAs [123]. Such AGO cross-linkingimmunoprecipitation followed by sequencing (CLIP–Seq) approaches have not been reported in plants. This could be due because, in contrast with animals where the majority of miRNA targets are not sliced, AGO–sRNA–target RNA interactions are ephemeral for the majority of plant target RNAs that might be immediately sliced upon AGO–miRNA recognition. Indeed, recent AGO RNA immunoprecipitation followed by high-throughput sequencing (RIP–Seq) analysis of AGO1-bound RNAs in *Arabidopsis* revealed that target RNAs are more efficiently co-immunoprecipitated with slicer-deficient AGO1 forms [24]. This suggests that AGO1 ternary complexes including miRNAs and target RNAs are more stable when AGO1 is catalytically inactive. By comparing the pool of target RNAs recovered from immunoprecipitates containing catalytically active or inactive AGO1 forms, it is possible to identify the repertoire of AGO1 target RNAs regulated by slicing and those regulated in a slicing-independent mode. Moreover, the application of this methodology to the different *Arabidopsis* AGOs could reveal the specific pool of target RNAs regulated by each specific AGO in different stress conditions or cell types. Understanding AGO–sRNA–target RNA dynamics is crucial to better understand sRNA-mediated gene silencing in plants.

4.3 AGO Transcriptional Regulators

While some plant AGOs such as *Arabidopsis* AGO1 and AGO4 are ubiquitously expressed, others have a more restricted expression. This is the case of *Arabidopsis* AGO9 and AGO10, which are expressed in female gamete and their accessory cells [114] or in provasculature, adaxial leaf primordia, and the meristem [101, 102], respectively. Moreover, several AGOs are induced upon abiotic or biotic stress. For instance, rice AGO18 accumulation is induced upon viral infection [79], while AGO2 and AGO3 accumulation is induced by gamma irradiation and bacterial infection [108] and salt stress [78], respectively. The differential spatiotemporal expression of the distinct AGO members as well as the induction of certain AGOs upon stress suggests that AGO transcription may be regulated. However, transcriptional regulators of plant AGOs are largely unknown. Only recently, it was shown that

Arabidopsis AGO10 expression is activated by at least one homeodomain–leucine zipper (HD–ZIP) transcription factor [124] and inhibited by the LBD12-1 transcription factor that directly binds to AGO10 promoter [125].

5 Conclusions and Future Challenges

Intensive research over the past two decades has elucidated the main functions of plant AGOs. However, future research should identify new functions for plant AGOs, as occurred for AGOs from other organisms. Emerging functions of non-plant AGOs include nonsense-mediated mRNA decay regulation in humans [126], alternative splicing in humans [127, 128] and *Drosophila* [129], sRNA-independent association with full-length introns (called "agotrons") to control gene expression in humans and probably in other mammals [130], nucleosome occupancy at human transcription start sites [131], and quality control of human proteins entering the secretory pathway [132]. Remarkably, DNA-guided genome editing has been recently reported in human cells using *Natronobacterium gregoryi* AGO [133], although failure to replicate these results by other groups [134] has questioned the general applicability of this approach.

Several outstanding questions remain to be answered in the plant AGO field. At the molecular level, more structural work is needed to better understand the formation of AGO ternary complexes. In particular, how AGOs scan and find target transcripts? Or how ternary complexes dissociate? Indeed, to date no crystal structure for a complete plant AGO has been solved. Also, besides AGO4 binding to sidRNA precursors, can other AGOs regulate target RNAs in a sRNA guide-independent mode? Regarding sRNA-mediated translational repression of target RNAs, what is the degree of miRNA–target RNA complementarity necessary to support the translational inhibition activity of plant miRNAs? Can AGOs other than AGO1 or AGO10 be programmed to function in a translational repression mode? And for those target RNAs regulated by slicing and translational repression [39, 46, 135], what mechanism(s) underlie the choice between these two modes of action? At a cellular level, how AGO ternary complexes are programmed in different cell types and tissues? Cell-type-specific profiling of AGO–small RNA–target RNA dynamics in different cell types and tissues should shed light on the role of the different AGO modules in the large regulatory networks established during development and stress response. Because of the broad interest of these fundamental questions, I anticipate that at least some of them will be answered soon.

Acknowledgments

I would like to thank James C. Carrington for intellectual and funding support, as well as members of his laboratory. This work was supported by an Individual Fellowship from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 655841 to A.C.

References

1. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
2. Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 12(2):99–110. doi:10.1038/nrg2936
3. Cerutti H, Casas-Mollano JA (2006) On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet 50(2):81–99. doi:10.1007/s00294-006-0078-x
4. Fang X, Qi Y (2016) RNAi in plants: an argonaute-centered view. Plant Cell 28(2):272–285. doi:10.1105/tpc1500920
5. Kapoor M, Arora R, Lama T, Nijhawan A, Khurana JP, Tyagi AK, Kapoor S (2008) Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analysis during reproductive development and stress in rice. BMC Genomics 9:451. doi:10.1186/1471-2164-9-451
6. Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H (2002) Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. Plant Cell 14(3):629–639. doi:10.1105/tpc010358
7. Yamasaki T, Kim EJ, Cerutti H, Ohama T (2016) Argonaute3 is a key player in miRNA-mediated target cleavage and translational repression in Chlamydomonas. Plant J 85(2):258–268. doi:10.1111/tpj13107
8. Schroda M (2006) RNA silencing in Chlamydomonas: mechanisms and tools. Curr Genet 49(2):69–84. doi:10.1007/s00294-005-0042-1
9. Arif MA, Frank W, Khraiwesh B (2013) Role of RNA interference (RNAi) in the moss Physcomitrella patens. Int J Mol Sci 14(1):1516–1540. doi:10.3390/ijms14011516
10. Zhang H, Xia R, Meyers BC, Walbot V (2015) Evolution, functions, and mysteries of plant ARGONAUTE proteins. Curr Opin Plant Biol 27:84–90. doi:10.1016/jpbi201506011
11. Chapman EJ, Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. Nat Rev Genet 8(11):884–896. doi:10.1038/nrg2179
12. Tolia NH, Joshua-Tor L (2007) Slicer and the argonautes. Nat Chem Biol 3(1):36–43. doi:10.1038/nchembio848
13. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305(5689):1434–1437. doi:10.1126/science1102514
14. 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
15. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133(1):128–141. doi:10.1016/jcell200802033
16. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis Argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/jcell200802034
17. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol 49(4):493–500. doi:10.1093/pcp/pcn043
18. Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X (2011) Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot

apical meristem development. Cell 145(2):242–256. doi:10.1016/jcell201103024
19. Zhang X, Niu D, Carbonell A, Wang A, Lee A, Tun V, Wang Z, Carrington JC, Chang CE, Jin H (2014) ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. Nat Commun 5:5468. doi:10.1038/ncomms6468
20. 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/science1102513
21. Sheng G, Zhao H, Wang J, Rao Y, Tian W, Swarts DC, van der Oost J, Patel DJ, Wang Y (2014) Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. Proc Natl Acad Sci U S A 111(2):652–657. doi:10.1073/pnas1321032111
22. Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A 102(33):11928–11933. doi:10.1073/pnas0505461102
23. Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19(3):421–428. doi:10.1016/jmolcel200506014
24. Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC (2012) Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. Plant Cell 24(9):3613–3629. doi:10.1105/tpc112099945
25. Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature 443(7114):1008–1012. doi:10.1038/nature05198
26. Ji L, Liu X, Yan J, Wang W, Yumul RE, Kim YJ, Dinh TT, Liu J, Cui X, Zheng B, Agarwal M, Liu C, Cao X, Tang G, Chen X (2011) ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in Arabidopsis. PLoS Genet 7(3):e1001358. doi:10.1371/journalpgen1001358
27. Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science 297(5589):2053–2056. doi:10.1126/science1076311
28. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110(4):513–520. doi:10.1016/S0092-8674(02)00863-2
29. Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP (2004) MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5′ region. EMBO J 23(16):3356–3364. doi:10.1038/sjemboj7600340
30. German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, De Paoli E, Lu C, Schroth G, Meyers BC, Green PJ (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26(8):941–946. doi:10.1038/nbt1417
31. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. Curr Biol 18(10):758–762. doi:10.1016/jcub200804042
32. Arribas-Hernandez L, Kielpinski LJ, Brodersen P (2016) mRNA decay of most Arabidopsis miRNA targets requires slicer activity of AGO1. Plant Physiol 171(4):2620–2632. doi:10.1104/pp.16.00231
33. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC (2010) Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat Struct Mol Biol 17(8):997–1003. doi:10.1038/nsmb1866
34. Montgomery TA, Yoo SJ, Fahlgren N, Gilbert SD, Howell MD, Sullivan CM, Alexander A, Nguyen G, Allen E, Ahn JH, Carrington JC (2008) AGO1-miR173 complex initiates phased siRNA formation in plants. Proc Natl Acad Sci U S A 105(51):20055–20062. doi:10.1073/pnas0810241105
35. Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121(2):207–221. doi:10.1016/jcell200504004
36. Yoshikawa M, Peragine A, Park MY, Poethig RS (2005) A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes Dev 19(18):2164–2175. doi:10.1101/gad1352605
37. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes Dev 20(24):3407–3425. doi:10.1101/gad1476406
38. Arribas-Hernandez L, Marchais A, Poulsen C, Haase B, Hauptmann J, Benes V, Meister G, Brodersen P (2016) The slicer activity of

ARGONAUTE1 Is required specifically for the phasing, not production, of trans-acting short interfering RNAs in Arabidopsis. Plant Cell 28(7):1563–1580. doi:10.1105/tpc1600121
39. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190. doi:10.1126/science1159151
40. Li S, Le B, Ma X, Li S, You C, Yu Y, Zhang B, Liu L, Gao L, Shi T, Zhao Y, Mo B, Cao X, Chen X (2016) Biogenesis of phased siRNAs on membrane-bound polysomes in Arabidopsis. Elife 5:e22750. doi:10.7554/eLife22750
41. Zeng Y, Yi R, Cullen BR (2003) MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A 100(17):9779–9784. doi:10.1073/pnas1630797100
42. Iwakawa HO, Tomari Y (2013) Molecular insights into microRNA-mediated translational repression in plants. Mol Cell 52(4):591–601. doi:10.1016/jmolcel201310033
43. Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, Zhang F, Raikhel N, Jiang L, Chen X (2013) MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. Cell 153(3):562–574. doi:10.1016/jcell201304005
44. Li JF, Chung HS, Niu Y, Bush J, McCormack M, Sheen J (2013) Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. Plant Cell 25(5):1507–1522. doi:10.1105/tpc113112235
45. Liu MJ, SH W, JF W, Lin WD, YC W, Tsai TY, Tsai HL, SH W (2013) Translational landscape of photomorphogenic Arabidopsis. Plant Cell 25(10):3699–3710. doi:10.1105/tpc113114769
46. Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. Plant Cell 15(11):2730–2741. doi:10.1105/tpc016238
47. Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science 303(5666):2022–2025. doi:10.1126/science1088060
48. Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J 49(4):683–693. doi:10.1111/j1365-313X200602983x
49. Yang L, Wu G, Poethig RS (2012) Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. Proc Natl Acad Sci U S A 109(1):315–320. doi:10.1073/pnas1114673109
50. Mallory AC, Hinze A, Tucker MR, Bouche N, Gasciolli V, Elmayan T, Lauressergues D, Jauvion V, Vaucheret H, Laux T (2009) Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. PLoS Genet 5(9):e1000646. doi:10.1371/journalpgen1000646
51. Hou CY, Lee WC, Chou HC, Chen AP, Chou SJ, Chen HM (2016) Global analysis of truncated RNA ends reveals new insights into ribosome stalling in plants. Plant Cell 28(10):2398–2416. doi:10.1105/tpc1600295
52. Rogers K, Chen X (2013) Biogenesis, turnover, and mode of action of plant microRNAs. Plant Cell 25(7):2383–2399. doi:10.1105/tpc113113159
53. Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E (2006) mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev 20(14):1885–1898. doi:10.1101/gad1424106
54. Wu L, Fan J, Belasco JG (2006) MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A 103(11):4034–4039. doi:10.1073/pnas0510928103
55. Schirle NT, MacRae IJ (2012) The crystal structure of human Argonaute2. Science 336(6084):1037–1040. doi:10.1126/science1221551
56. Pfaff J, Hennig J, Herzog F, Aebersold R, Sattler M, Niessing D, Meister G (2013) Structural features of Argonaute-GW182 protein interactions. Proc Natl Acad Sci U S A 110(40):E3770–E3779. doi:10.1073/pnas1308510110
57. Ma X, Kim EJ, Kook I, Ma F, Voshall A, Moriyama E, Cerutti H (2013) Small interfering RNA-mediated translation repression alters ribosome sensitivity to inhibition by cycloheximide in Chlamydomonas reinhardtii. Plant Cell 25(3):985–998. doi:10.1105/tpc113109256
58. Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11(3):204–220. doi:10.1038/nrg2719
59. Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional

diversification of small RNA pathways in plants. PLoS Biol 2(5):E104. doi:10.1371/journalpbio0020104
60. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. Science 308(5718):118–120. doi:10.1126/science1106910
61. Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet 37(7):761–765. doi:10.1038/ng1580
62. Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell 120(5):613–622. doi:10.1016/jcell200502007
63. Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, Pikaard CS (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol Cell 48(5):811–818. doi:10.1016/jmolcel201209027
64. Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS (2006) The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell 126(1):79–92. doi:10.1016/jcell200605031
65. Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell 126(1):93–106. doi:10.1016/jcell200605032
66. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T (2007) Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev 21(20):2539–2544. doi:10.1101/gad451207
67. Li CF, Henderson IR, Song L, Fedoroff N, Lagrange T, Jacobsen SE (2008) Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in Arabidopsis thaliana. PLoS Genet 4(2):e27. doi:10.1371/journalpgen0040027
68. Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, Lagrange T (2009) RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. EMBO Rep 10(6):649–654. doi:10.1038/embor200931
69. He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu HL, Wang CS, Jin H, Zhu JK (2009) An effector of RNA-directed DNA methylation in Arabidopsis is an ARGONAUTE 4- and RNA-binding protein. Cell 137(3):498–508. doi:10.1016/jcell200904028
70. Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE (2014) Molecular mechanism of action of plant DRM de novo DNA methyltransferases. Cell 157(5):1050–1060. doi:10.1016/jcell201403056
71. Cao X, Jacobsen SE (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci U S A 99(Suppl 4):16491–16498. doi:10.1073/pnas162371599
72. Lahmy S, Pontier D, Bies-Etheve N, Laudie M, Feng S, Jobet E, Hale CJ, Cooke R, Hakimi MA, Angelov D, Jacobsen SE, Lagrange T (2016) Evidence for ARGONAUTE4-DNA interactions in RNA-directed DNA methylation in plants. Genes Dev 30(23):2565–2570. doi:10.1101/gad289553116
73. Zheng X, Zhu J, Kapoor A, Zhu JK (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J 26(6):1691–1701. doi:10.1038/sjemboj7601603
74. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC (2010) The Arabidopsis RNA-directed DNA methylation Argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22(2):321–334. doi:10.1105/tpc109072199
75. Eun C, Lorkovic ZJ, Naumann U, Long Q, Havecker ER, Simon SA, Meyers BC, Matzke AJ, Matzke M (2011) AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in Arabidopsis thaliana. PLoS One 6(10):e25730. doi:10.1371/journalpone0025730
76. Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, Wang P, Zhou J, Liang G, Liu N, Wang C, Zhu JK (2015) Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. EMBO J 34(5):581–592. doi:10.15252/embj201489453
77. McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK (2015) ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. EMBO J 34(1):20–35. doi:10.15252/embj201489499

78. Zhang Z, Liu X, Guo X, Wang XJ, Zhang X (2016) Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing. Nat Plants 2(5):16049. doi:10.1038/nplants201649
79. Wu J, Yang Z, Wang Y, Zheng L, Ye R, Ji Y, Zhao S, Ji S, Liu R, Xu L, Zheng H, Zhou Y, Zhang X, Cao X, Xie L, Wu Z, Qi Y, Li Y (2015) Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA. Elife 4:05733. doi:10.7554/eLife05733
80. Wu J, Yang R, Yang Z, Yao S, Zhao S, Wang Y, Li P, Song X, Jin L, Zhou T, Lan Y, Xie L, Zhou X, Chu C, Qi Y, Cao X, Li Y (2017) ROS accumulation and antiviral defence control by microRNA528 in rice. Nat Plants 3:16203. doi:10.1038/nplants2016203
81. Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, White CI, Rendtlew Danielsen JM, Yang YG, Qi Y (2012) A role for small RNAs in DNA double-strand break repair. Cell 149(1):101–112. doi:10.1016/jcell201203002
82. Oliver C, Santos JL, Pradillo M (2014) On the role of some ARGONAUTE proteins in meiosis and DNA repair in Arabidopsis thaliana. Front Plant Sci 5:177. doi:10.3389/fpls201400177
83. Ye R, Chen Z, Lian B, Rowley MJ, Xia N, Chai J, Li Y, He XJ, Wierzbicki AT, Qi Y (2016) A Dicer-independent route for biogenesis of siRNAs that direct DNA methylation in Arabidopsis. Mol Cell 61(2):222–235. doi:10.1016/jmolcel201511015
84. Dolata J, Bajczyk M, Bielewicz D, Niedojadlo K, Niedojadlo J, Pietrykowska H, Walczak W, Szweykowska-Kulinska Z, Jarmolowski A (2016) Salt stress reveals a new role for ARGONAUTE1 in miRNA biogenesis at the transcriptional and posttranscriptional levels. Plant Physiol 172(1):297–312. doi:10.1104/pp1600830
85. Singh RK, Gase K, Baldwin IT, Pandey SP (2015) Molecular evolution and diversification of the Argonaute family of proteins in plants. BMC Plant Biol 15(1):23. doi:10.1186/s12870-014-0364-6
86. Singh RK, Pandey SP (2015) Evolution of structural and functional diversification among plant Argonautes. Plant Signal Behav 10(10):e1069455. doi:10.1080/1559232420151069455
87. Bohmert K, Camus I, Bellini C, Bouchez D, Caboche M, Benning C (1998) AGO1 defines a novel locus of Arabidopsis controlling leaf development. EMBO J 17(1):170–180. doi:10.1093/emboj/171170
88. Kidner CA, Martienssen RA (2004) Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. Nature 428(6978):81–84. doi:10.1038/nature02366
89. Sorin C, Bussell JD, Camus I, Ljung K, Kowalczyk M, Geiss G, McKhann H, Garcion C, Vaucheret H, Sandberg G, Bellini C (2005) Auxin and light control of adventitious rooting in Arabidopsis require ARGONAUTE1. Plant Cell 17(5):1343–1359. doi:10.1105/tpc105031625
90. Yang L, Huang W, Wang H, Cai R, Xu Y, Huang H (2006) Characterizations of a hypomorphic argonaute1 mutant reveal novel AGO1 functions in Arabidopsis lateral organ development. Plant Mol Biol 61(1-2):63–78. doi:10.1007/s11103-005-5992-7
91. Kidner CA, Martienssen RA (2005) The developmental role of microRNA in plants. Curr Opin Plant Biol 8(1):38–44. doi:10.1016/jpbi200411008
92. Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y (2009) Rice microRNA effector complexes and targets. Plant Cell 21(11):3421–3435. doi:10.1105/tpc109070938
93. Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13(7):350–358. doi:10.1016/jtplants200804007
94. Hunter C, Sun H, Poethig RS (2003) The Arabidopsis heterochronic gene ZIPPY is an ARGONAUTE family member. Curr Biol 13(19):1734–1739
95. Adenot X, Elmayan T, Lauressergues D, Boutet S, Bouche N, Gasciolli V, Vaucheret H (2006) DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Curr Biol 16(9):927–932. doi:10.1016/jcub200603035
96. Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, Carrington JC (2006) Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. Curr Biol 16(9):939–944. doi:10.1016/jcub200603065
97. Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006) A two-hit trigger for siRNA biogenesis in plants. Cell 127(3):565–577. doi:10.1016/jcell200609032
98. Hunter C, Willmann MR, Wu G, Yoshikawa M, de la Luz G-NM, Poethig SR (2006) Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. Development 133(15):2973–2981. doi:10.1242/dev02491
99. Nagasaki H, Itoh J, Hayashi K, Hibara K, Satoh-Nagasawa N, Nosaka M, Mukouhata M, Ashikari M, Kitano H, Matsuoka M, Nagato Y, Sato Y (2007) The small interfer-

ing RNA production pathway is required for shoot meristem initiation in rice. Proc Natl Acad Sci U S A 104(37):14867–14871. doi:10.1073/pnas0704339104
100. Douglas RN, Wiley D, Sarkar A, Springer N, Timmermans MC, Scanlon MJ (2010) Ragged seedling2 encodes an ARGONAUTE7-like protein required for mediolateral expansion, but not dorsiventrality, of maize leaves. Plant Cell 22(5):1441–1451. doi:10.1105/tpc109071613
101. Lynn K, Fernandez A, Aida M, Sedbrook J, Tasaka M, Masson P, Barton MK (1999) The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. Development 126(3):469–481
102. Moussian B, Schoof H, Haecker A, Jurgens G, Laux T (1998) Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. EMBO J 17(6):1799–1809. doi:10.1093/emboj/1761799
103. Nishimura A, Ito M, Kamiya N, Sato Y, Matsuoka M (2002) OsPNH1 regulates leaf development and maintenance of the shoot apical meristem in rice. Plant J 30(2):189–201
104. Carbonell A, Carrington JC (2015) Antiviral roles of plant ARGONAUTES. Curr Opin Plant Biol 27:111–117. doi:10.1016/jpbi201506013
105. Szittya G, Burgyan J (2013) RNA interference-mediated intrinsic antiviral immunity in plants. Curr Top Microbiol Immunol 371:153–181. doi:10.1007/978-3-642-37765-5_6
106. Minoia S, Carbonell A, Di Serio F, Gisel A, Carrington JC, Navarro B, Flores R (2014) Specific argonautes selectively bind small RNAs derived from potato spindle tuber viroid and attenuate viroid accumulation in vivo. J Virol 88(20):11933–11945. doi:10.1128/JVI01404-14
107. Brosseau C, El Oirdi M, Adurogbangba A, Ma X, Moffett P (2016) Antiviral defense involves AGO4 in an Arabidopsis-Potexvirus interaction. Mol Plant Microbe Interact 29(11):878–888. doi:10.1094/MPMI-09-16-0188-R
108. Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang HD, Raikhel N, Jin H (2011) Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. Mol Cell 42(3):356–366. doi:10.1016/jmolcel201104010
109. Agorio A, Vera P (2007) ARGONAUTE4 is required for resistance to Pseudomonas syringae in Arabidopsis. Plant Cell 19(11):3778–3790. doi:10.1105/tpc107054494
110. Borges F, Martienssen RA (2015) The expanding world of small RNAs in plants. Nat Rev Mol Cell Biol 16(12):727–741. doi:10.1038/nrm4085
111. Tucker MR, Okada T, Hu Y, Scholefield A, Taylor JM, Koltunow AM (2012) Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in Arabidopsis. Development 139(8):1399–1404. doi:10.1242/dev075390
112. Nonomura K, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, Kurata N (2007) A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. Plant Cell 19(8):2583–2594. doi:10.1105/tpc107053199
113. Singh M, Goel S, Meeley RB, Dantec C, Parrinello H, Michaud C, Leblanc O, Grimanelli D (2011) Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. Plant Cell 23(2):443–458. doi:10.1105/tpc110079020
114. Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature 464(7288):628–632. doi:10.1038/nature08828
115. 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. doi:10.1016/jmolcel201005014
116. Iki T, Yoshikawa M, Meshi T, Ishikawa M (2012) Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. EMBO J 31(2):267–278. doi:10.1038/emboj2011395
117. Smith MR, Willmann MR, Wu G, Berardini TZ, Moller B, Weijers D, Poethig RS (2009) Cyclophilin 40 is required for microRNA activity in Arabidopsis. Proc Natl Acad Sci U S A 106(13):5424–5429. doi:10.1073/pnas0812729106
118. Cui Y, Fang X, Qi Y (2016) TRANSPORTIN1 promotes the association of microRNA with ARGONAUTE1 in Arabidopsis. Plant Cell 28(10):2576–2585. doi:10.1105/tpc1600384
119. Karlowski WM, Zielezinski A, Carrere J, Pontier D, Lagrange T, Cooke R (2010)

Genome-wide computational identification of WG/GW Argonaute-binding proteins in Arabidopsis. Nucleic Acids Res 38(13):4231–4245. doi:10.1093/nar/gkq162
120. Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159. doi:10.1146/annurev-arplant-050312-120043
121. Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53. doi:10.1146/annurevarplant57032905105218
122. Li J, Reichel M, Li Y, Millar AA (2014) The functional scope of plant microRNA-mediated silencing. Trends Plant Sci 19(12):750–756. doi:10.1016/jtplants201408006
123. Mittal N, Zavolan M (2014) Seq and CLIP through the miRNA world. Genome Biol 15(1):202. doi:10.1186/gb4151
124. Brandt R, Xie Y, Musielak T, Graeff M, Stierhof YD, Huang H, Liu CM, Wenkel S (2013) Control of stem cell homeostasis via interlocking microRNA and microProtein feedback loops. Mech Dev 130(1):25–33. doi:10.1016/jmod201206007
125. Ma W, Wu F, Sheng P, Wang X, Zhang Z, Zhou K, Zhang H, Hu J, Lin Q, Cheng Z, Wang J, Zhu S, Zhang X, Guo X, Wang H, Wu C, Zhai H, Wan J (2017) The LBD12-1 transcription factor suppresses apical meristem size by repressing Argonaute 10 expression. Plant Physiol 173(1):801–811. doi:10.1104/pp1601699
126. Choe J, Cho H, Lee HC, Kim YK (2010) microRNA/Argonaute 2 regulates nonsense-mediated messenger RNA decay. EMBO Rep 11(5):380–386. doi:10.1038/embor201044
127. Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, Morozova N, Fenouil R, Descostes N, Andrau JC, Mathieu J, Hamiche A, Ait-Si-Ali S, Muchardt C, Batsche E, Harel-Bellan A (2012) Argonaute proteins couple chromatin silencing to alternative splicing. Nat Struct Mol Biol 19(10):998–1004. doi:10.1038/nsmb2373
128. Allo M, Agirre E, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V, Bellora N, Singh B, Petrillo E, Blaustein M, Minana B, Dujardin G, Pozzi B, Pelisch F, Bechara E, Agafonov DE, Srebrow A, Luhrmann R, Valcarcel J, Eyras E, Kornblihtt AR (2014) Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. Proc Natl Acad Sci U S A 111(44):15622–15629. doi:10.1073/pnas1416858111
129. Taliaferro JM, Aspden JL, Bradley T, Marwha D, Blanchette M, Rio DC (2013) Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression. Genes Dev 27(4):378–389. doi:10.1101/gad210708112
130. Hansen TB, Veno MT, Jensen TI, Schaefer A, Damgaard CK, Kjems J (2016) Argonaute-associated short introns are a novel class of gene regulators. Nat Commun 7:11538. doi:10.1038/ncomms11538
131. Carissimi C, Laudadio I, Cipolletta E, Gioiosa S, Mihailovich M, Bonaldi T, Macino G, Fulci V (2015) ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome occupancy at human Transcription Start Sites. Nucleic Acids Res 43(3):1498–1512. doi:10.1093/nar/gku1387
132. Karamyshev AL, Patrick AE, Karamysheva ZN, Griesemer DS, Hudson H, Tjon-Kon-Sang S, Nilsson I, Otto H, Liu Q, Rospert S, von Heijne G, Johnson AE, Thomas PJ (2014) Inefficient SRP interaction with a nascent chain triggers a mRNA quality control pathway. Cell 156(1-2):146–157. doi:10.1016/jcell201312017
133. Gao F, Shen XZ, Jiang F, Wu Y, Han C (2016) DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. Nat Biotechnol 34(7):768–773. doi:10.1038/nbt3547
134. Lee SH, Turchiano G, Ata H, Nowsheen S, Romito M, Lou Z, Ryu SM, Ekker SC, Cathomen T, Kim JS (2016) Failure to detect DNA-guided genome editing using Natronobacterium gregoryi Argonaute. Nat Biotechnol 35(1):17–18. doi:10.1038/nbt3753
135. Beauclair L, Yu A, Bouche N (2010) microRNA-directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in Arabidopsis. Plant J 62(3):454–462. doi:10.1111/j1365-313X201004162x
136. Garcia-Ruiz H, Carbonell A, Hoyer JS, Fahlgren N, Gilbert KB, Takeda A, Giampetruzzi A, Garcia Ruiz MT, McGinn MG, Lowery N, Martinez Baladejo MT, Carrington JC (2015) Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection. PLoS Pathog 11(3):e1004755. doi:10.1371/journalppat1004755
137. Qu F, Ye X, Morris TJ (2008) Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc Natl Acad Sci U S A 105(38):14732–14737. doi:10.1073/pnas0805760105
138. Wang XB, Jovel J, Udomporn P, Wang Y, Wu Q, Li WX, Gasciolli V, Vaucheret H, Ding SW

(2011) The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in Arabidopsis thaliana. Plant Cell 23(4):1625–1638. doi:10.1105/tpc110082305
139. Dzianott A, Sztuba-Solinska J, Bujarski JJ (2012) Mutations in the antiviral RNAi defense pathway modify Brome mosaic virus RNA recombinant profiles. Mol Plant Microbe Interact 25(1):97–106. doi:10.1094/MPMI-05-11-0137
140. Harvey JJ, Lewsey MG, Patel K, Westwood J, Heimstadt S, Carr JP, Baulcombe DC (2011) An antiviral defense role of AGO2 in plants. PLoS One 6(1):e14639. doi:10.1371/journalpone0014639
141. Jaubert M, Bhattacharjee S, Mello AF, Perry KL, Moffett P (2011) ARGONAUTE2 mediates RNA-silencing antiviral defenses against Potato virus X in Arabidopsis. Plant Physiol 156(3):1556–1564. doi:10.1104/pp111178012
142. Ma X, Nicole MC, Meteignier LV, Hong N, Wang G, Moffett P (2015) Different roles for RNA silencing and RNA processing components in virus recovery and virus-induced gene silencing in plants. J Exp Bot 66(3):919–932. doi:10.1093/jxb/eru447
143. Cao M, Du P, Wang X, Yu YQ, Qiu YH, Li W, Gal-On A, Zhou C, Li Y, Ding SW (2014) Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. Proc Natl Acad Sci U S A 111(40):14613–14618. doi:10.1073/pnas1407131111
144. Hamera S, Song X, Su L, Chen X, Fang R (2012) Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. Plant J 69(1):104–115. doi:10.1111/j1365-313X201104774x
145. Bhattacharjee S, Zamora A, Azhar MT, Sacco MA, Lambert LH, Moffett P (2009) Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. Plant J 58(6):940–951. doi:10.1111/j1365-313X200903832x
146. Raja P, Sanville BC, Buchmann RC, Bisaro DM (2008) Viral genome methylation as an epigenetic defense against geminiviruses. J Virol 82(18):8997–9007. doi:10.1128/JVI00719-08
147. Raja P, Jackel JN, Li S, Heard IM, Bisaro DM (2014) Arabidopsis double-stranded RNA binding protein DRB3 participates in methylation-mediated defense against geminiviruses. J Virol 88(5):2611–2622. doi:10.1128/JVI02305-13
148. Brosseau C, Moffett P (2015) Functional and genetic analysis identify a role for Arabidopsis ARGONAUTE5 in antiviral RNA silencing. Plant Cell 27(6):1742–1754. doi:10.1105/tpc1500264
149. Ghoshal B, Sanfacon H (2014) Temperature-dependent symptom recovery in Nicotiana benthamiana plants infected with tomato ringspot virus is associated with reduced translation of viral RNA2 and requires ARGONAUTE 1. Virology 456-457:188–197. doi:10.1016/jvirol201403026
150. Scholthof HB, Alvarado VY, Vega-Arreguin JC, Ciomperlik J, Odokonyero D, Brosseau C, Jaubert M, Zamora A, Moffett P (2011) Identification of an ARGONAUTE for antiviral RNA silencing in Nicotiana benthamiana. Plant Physiol 156(3):1548–1555. doi:10.1104/pp111178764
151. Fatyol K, Ludman M, Burgyan J (2016) Functional dissection of a plant Argonaute. Nucleic Acids Res 44(3):1384–1397. doi:10.1093/nar/gkv1371
152. Odokonyero D, Mendoza MR, Alvarado VY, Zhang J, Wang X, Scholthof HB (2015) Transgenic down-regulation of ARGONAUTE2 expression in Nicotiana benthamiana interferes with several layers of antiviral defenses. Virology 486:209–218. doi:10.1016/jvirol201509008
153. Vaucheret H, Vazquez F, Crete P, Bartel DP (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev 18(10):1187–1197. doi:10.1101/gad1201404
154. Liu X, Tang S, Jia G, Schnable JC, Su H, Tang C, Zhi H, Diao X (2016) The C-terminal motif of SiAGO1b is required for the regulation of growth, development and stress responses in foxtail millet (Setaria italica (L) P Beauv). J Exp Bot 67(11):3237–3249. doi:10.1093/jxb/erw135
155. Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299(5607):716–719. doi:10.1126/science1079695
156. Wu L, Mao L, Qi Y (2012) Roles of dicer-like and argonaute proteins in TAS-derived small interfering RNA-triggered DNA methylation. Plant Physiol 160(2):990–999. doi:10.1104/pp112200279
157. Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK (2013) The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. Plant Physiol 162(1):116–131. doi:10.1104/pp113216481
158. Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA methylation mediated

by a microRNA pathway. Mol Cell 38(3):465–475. doi:10.1016/jmolcel201003008
159. Zhai J, Zhang H, Arikit S, Huang K, Nan GL, Walbot V, Meyers BC (2015) Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. Proc Natl Acad Sci U S A 112(10):3146–3151. doi:10.1073/pnas1418918112
160. Xu L, Yang L, Pi L, Liu Q, Ling Q, Wang H, Poethig RS, Huang H (2006) Genetic interaction between the AS1-AS2 and RDR6-SGS3-AGO7 pathways for leaf morpho-genesis. Plant Cell Physiol 47(7):853–863. doi:10.1093/pcp/pcj057
161. Oliver C, Santos JL, Pradillo M (2016) Accurate chromosome segregation at first meiotic division requires AGO4, a protein involved in RNA-dependent DNA methylation in Arabidopsis thaliana. Genetics 204(2):543–553. doi:10.1534/genetics116189217
162. Liu Q, Yao X, Pi L, Wang H, Cui X, Huang H (2009) The ARGONAUTE10 gene modulates shoot apical meristem maintenance and establishment of leaf polarity by repressing miR165/166 in Arabidopsis. Plant J 58(1):27–40. doi:10.1111/j1365-313X200803757x
163. Zhou Y, Honda M, Zhu H, Zhang Z, Guo X, Li T, Li Z, Peng X, Nakajima K, Duan L, Zhang X (2015) Spatiotemporal sequestration of miR165/166 by Arabidopsis Argonaute10 promotes shoot apical meristem maintenance. Cell Rep 10(11):1819–1827. doi:10.1016/jcelrep201502047
164. Li W, Cui X, Meng Z, Huang X, Xie Q, Wu H, Jin H, Zhang D, Liang W (2012) Transcriptional regulation of Arabidopsis MIR168a and argonaute1 homeostasis in abscisic acid and abiotic stress responses. Plant Physiol 158(3):1279–1292. doi:10.1104/pp111188789
165. Earley K, Smith M, Weber R, Gregory B, Poethig R (2010) An endogenous F-box protein regulates ARGONAUTE1 in Arabidopsis thaliana. Silence 1(1):15. doi:10.1186/1758-907X-1-15

Chapter 2

Analysis of the Uridylation of Both ARGONAUTE-Bound MiRNAs and 5′ Cleavage Products of Their Target RNAs in Plants

Guodong Ren, Xiaoyan Wang, and Bin Yu

Abstract

Uridylation (3′ untemplated uridine addition) provides a mechanism to trigger the degradation of miRNAs and the 5′ cleavage products (5′ CP) that are produced from miRNA-directed ARGONAUTE (AGO) cleavage of target RNAs. We have recently shown that HEN1 SUPPRESSOR 1 (HESO1), a terminal uridylyltransferase, and its homolog UTP:RNA uridylyltransferase 1 (URT1) catalyze the uridylation of miRNAs and 5′ CPs within the AGO complex in higher plants. In this chapter, we describe detailed protocols for analyzing 3′ end uridylation of both AGO-bound miRNAs and 5′ CP.

Key words miRNA, ARGONAUTE, 5′ Cleavage product, Uridylation, HESO1, Rapid amplification of cDNA ends (RACE)

1 Introduction

MicroRNAs (miRNAs) are ~21-nucleotide-long noncoding RNAs that mainly regulate gene expression at posttranscriptional levels. They play crucial roles in many biological processes such as development, physiology, metabolism, and immunity [1, 2]. miRNAs are first excised as duplexes with 2 nt 3′ overhang from their imperfect stem–loop precursors called primary miRNA transcripts (pri-miRNAs) [3, 4]. Unlike metazoans, plant miRNA duplex is subject to 2′-O-methylation at 3′ end of each strand by the small RNA methyltransferase HUA1 ENHANCER1 (HEN1) [5]. Upon production, miRNA is loaded into the effector protein called ARGONAUTE (AGO) and then guide it to interact with target RNAs through base pairing with the complementary sequence(s) within targets [6, 7]. AGO suppresses target expression through deadenylation-mediated RNA decay, translational inhibition, or target cleavage [8]. Target RNA cleavage by AGO happens at a position opposite to between 10th and 11th nucleotides of the

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_2,

guide miRNA, which produces a 5′ RNA fragment (5′ cleavage product, 5′ CP) and a 3′ RNA fragment (3′ CP) [9]. In plants, extensive complementarity between miRNAs and targets results in predominant target cleavage by AGO1 (the plant miRNA effector). In contrast, less complementarity in metazoans leads to predominant deadenylation-mediated RNA decay and/or translational inhibition, while target cleavage also exists [10, 11]. The AGO cleavage products need to be further removed. Otherwise, they may serve as template for secondary small interfering RNA (siRNA) formation that may cause lethality of organisms [12].

In plants lacking *hen1*, miRNAs often contain untemplated uridines at 3′ end (uridylation) and become less stable [13]. Subsequently, studies have shown that uridylation regulates the stability and activity of some metazoan miRNAs [14, 15]. Interestingly, 5′ CPs can also be uridylated, which triggers 5′ CP degradation [16, 17]. We recently show that both miRNAs and 5′ CP are uridylated by the terminal uridine transferase HEN1 SUPRESSOR 1 (HESO1) and its homolog UTP:RNA uridylyltransferase 1 (URT1) [16, 18–21]. Further studies have shown that HESO1 and URT1 bind AGO1, demonstrating that HESO1 and URT1 act on miRNAs and 5′ CP within the AGO1 complex [16, 19]. Here, we describe the in vitro and in vivo protocols used to analyze uridylation of AGO1-bound miRNAs and 5′ CP, respectively (Fig. 1).

2 Materials and Key Equipment

2.1 Analysis of the 3′ Uridylation of ARGONAUTE-Bound MiRNAs

2.1.1 Transient Expression of GFP–AGO1 in Nicotiana benthamiana

1. *N. benthamiana* plants.
2. *Agrobacterium tumefaciens* strain GV3101 carrying a GFP–AGO1 or P19 plasmid [22].
3. YEB medium: 5 g/L beef extract, 1 g/L yeast extract, 5 g/L peptone, 5 g/L sucrose, 0.5 g/L magnesium chloride ($MgCl_2$), 1% agar (plate only), with appropriate antibiotics.
4. Infiltration medium: 10 mM $MgCl_2$, 10 mM MES pH 5.6, 200 μM acetosyringone (add freshly).
5. 3 mL syringe.

Fig. 1 (continued) PAGE gel and autoradiography. (**b**) A schematic diagram for detecting 3′ end uridylation of 5′ CP in vivo. Either RLM-3′ RACE or cRACE strategy was used to retrieve the 3′ end signatures of 5′ CP. For RLM-3′ RACE, total RNA was ligated to an RNA adaptor followed by reverse transcription and two rounds of semi-nested PCR. For cRACE, total RNA was directly subject to self-ligation (for uncapped products) or successively treated with CIP and TAP before self-ligation (for capped products). After reverse transcription using R1 primer, two rounds of nested PCR were performed. For both methods, PCR products were cloned, and end signatures were obtained by Sanger sequencing

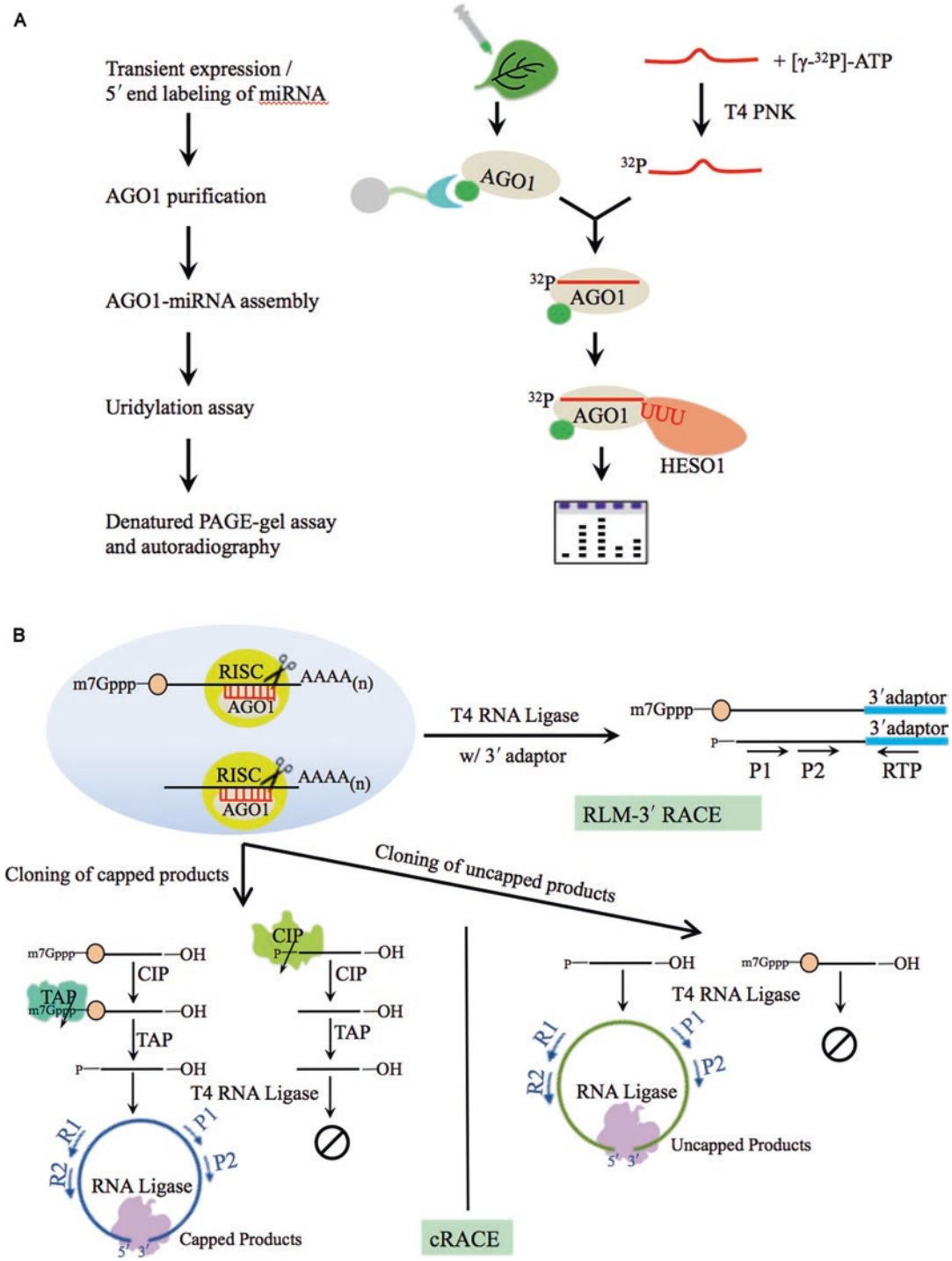


Fig. 1 Schematic flowcharts for uridylation assays on AGO1-bound miRNAs in vitro and 5′ cleavage products (CP) in vivo. (**a**) A schematic flowchart for HESO1-mediated uridylation assay on AGO1-bound miRNAs in vitro. AGO1 is obtained by transient expression of GFP–AGO1 in *N. benthamiana* followed by immunopurification using anti-GFP antibodies that were pre-coupled to protein A agarose beads. miRNA was 5′ end ^{32}P labeled with T4 PNK and was then loaded onto GFP–AGO1. The assembled AGO1–miRNA complex is subject to HESO1-mediated uridylation assay. After wash, AGO1-bound miRNAs were extracted and analyzed on a denatured

2.1.2 Purification of GFP–AGO1

1. Liquid nitrogen.
2. Mortars and pestles.
3. End-over rotator wheel.
4. Antibodies: Rabbit anti-GFP (Clontech) or GFP-Trap (ChromoTek).
5. Diethyl pyrocarbonate (DEPC)-treated H_2O. Add 1 mL DEPC to 1 L ultrapure H_2O (e.g., Milli-Q H_2O), stir overnight, and autoclave (*see* **Note 1**).
6. Protein extraction buffer (prepared with DEPC-treated H_2O): 50 mM Tris–HCl pH 7.5, 150 mM sodium chloride (NaCl), 5 mM $MgCl_2$, 5% glycerol, 2 mM dithiothreitol (DTT), 0.1 mM phenylmethanesulfonyl fluoride (PMSF), and 1/100 protease inhibitor (Thermo Fisher Scientific) (*see* **Note 2**).
7. Protein A agarose beads (Sigma-Aldrich).

2.1.3 Preparation of Radio-Labeled miRNA Probes

1. [γ-^{32}P]ATP (10 mCi/mL, 3000 Ci/mmol, PerkinElmer).
2. miR166a RNA oligonucleotide: 5′ UCGGACCAGGCUU CAUUCCCC 3′.
3. T4 polynucleotide kinase (T4 PNK) (New England Biolabs).
4. MicroSpin G-25 columns (GE Healthcare).
5. Phenol pH 4.5.
6. Chloroform.
7. Ethanol.
8. 3 M sodium acetate (NaOAc) pH 4.3.
9. Glycogen.

2.1.4 AGO1 Assembly and Uridylation Assay

1. RiboLock RNase inhibitor (Thermo Fisher Scientific).
2. Maltose-binding protein (MBP) and MBP-tagged HESO1 protein [20].
3. Reaction exchange buffer (prepared with DEPC-treated H_2O): 50 mM NaCl, 10 mM Tris–HCl, 10 mM $MgCl_2$, pH 7.9.
4. Reaction buffer (prepared with DEPC-treated H_2O): 50 mM NaCl, 10 mM Tris–HCl, 10 mM $MgCl_2$, 100 μg/mL bovine serum albumin (BSA), 0.5 mM UTP, 1 U/μL RNase inhibitor, pH 7.9.
5. 5× Tris–borate–EDTA (TBE) buffer: 54 g/L Tris, 27.5 g/L boric acid, 10 mM EDTA, pH 8.3.
6. 16% polyacrylamide denaturing gel stock solution: 16% acrylamide/bis (29:1), 42% urea, 0.5× TBE. Dissolve the mixture in 42 °C water bath and pass through a 0.22 μm filter (*see* **Note 3**).

7. 10% ammonium persulfate (APS).
8. N,N,N′,N′-tetramethylethylenediamine (TEMED).
9. 2× formamide RNA loading dye: mix 8 mL of formamide with 2 mL of 5× TBE, and add trace amount of xylene cyanol and bromophenol blue.
10. Bio-Rad PROTEAN II xi Cell (Bio-Rad).
11. BAS Storage Phosphor Screen (GE Healthcare).
12. Typhoon FLA 9500 scanner (GE Healthcare).

2.2 Analysis of the 3′ Uridylation of 5′ Cleavage Products

2.2.1 Total RNA Extraction

1. TRI Reagent (Molecular Research Center).
2. β-Mercaptoethanol.
3. Chloroform.
4. Isopropanol.
5. 70% ethanol.
6. DEPC-treated H_2O.
7. NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).

2.2.2 RNA Ligase-Mediated 3′ Rapid Amplification of cDNA Ends (RLM-3′ RACE)

1. 3′ RNA adaptor: pUUUdCdTdGdTdAdGdGdCdAdCdCdAdTdCdAdAdTidT.
2. Primers: RT primer, 5′-ATTGATGGTGCCTACAG-3′; P1 primer (For MYB33), 5′-AAGCGACTTTGGGAATCTGA-3′; P2 primer (For MYB33), 5′-AAGAATTCTCGTCGCCTGAA-3′.
3. T4 RNA ligase1 (New England Biolabs).
4. SuperScript III Reverse Transcriptase (Thermo Fisher Scientific).
5. RiboLock RNase inhibitor (Thermo Fisher Scientific).
6. 10 mM ATP.
7. pGEM-T Easy Vector (Promega).
8. Phenol pH 4.5.
9. Chloroform.
10. Ethanol.
11. 3 M NaOAc pH 4.3.
12. Glycogen.
13. DEPC-treated H_2O.

2.2.3 Circulation-Mediated Rapid Amplification of cDNA Ends (cRACE)

1. Alkaline phosphatase, calf intestinal (CIP) (New England Biolabs).
2. Tobacco acid pyrophosphatase (TAP) (Epicentre) (*see* **Note 4**).

3. Primers: P1, P2 as in Subheading 2.2.2; R1 primer (For MYB33), 5′-GCCATACGTGCCCATCTATT-3′; R2 primer (For MYB33), 5′-TTGGCCTCAGATGATTAGCC-3′.
4. T4 RNA ligase1 (New England Biolabs).
5. Phenol pH 4.5.
6. Chloroform.
7. Ethanol.
8. 3 M NaOAc pH 4.3.
9. Glycogen.
10. DEPC-treated H_2O.

3 Methods

3.1 Analysis of the 3′ Uridylation of AGO1-Bound miRNAs

3.1.1 Transient Expression of GFP–AGO1 in Nicotiana Benthamiana

A straightforward way to obtain AGO1 is to immunoprecipitate it from *Arabidopsis* tissues using an anti-AGO1 antibody (Agrisera or produced according to [23]). Alternatively, one may express tagged AGO1 protein (e.g., GFP–AGO1) either stably in *Arabidopsis* (i.e., stable transgenic line) or transiently in *N. benthamiana* and purify it by using a commercial antibody recognizing the tag [16, 18]. In this protocol, we describe methods for the transient expression and purification of GFP–AGO1 from *N. benthamiana* leaves.

1. Grow *N. benthamiana* at 23 °C under long day condition (16 h light) for 4–5 weeks.
2. Inoculate three to four fresh colonies of *A. tumefaciens* containing either GFP–AGO1 or P19 into 10 mL of YEB medium supplemented with appropriate antibiotics.
3. Grow the culture at 28 °C with vigorous shaking (~230 rpm) to an OD600 of 1.2–2.0, which usually takes 12–20 h when using fresh prepared colonies in the *A. tumefaciens* strain GV3101 (*see* **Note 5**).
4. Centrifuge the culture at 1500 × *g* for 20 min and discard the supernatant.
5. Resuspend the agrobacteria in 10 mL infiltration medium, and incubate at room temperature for 3 h with gentle shaking (~75 rpm).
6. Centrifuge the culture at 1500 × *g* for 20 min, and resuspend the agrobacteria pellet in fresh infiltration buffer, and adjust the final concentration to an OD600 of 0.4–0.5 (for P19, adjust the final concentration to an OD600 of 0.8–0.9).
7. Mix the GFP–AGO1 agrobacteria suspension with P19 at a ratio of 1:1 (v/v).

8. Infiltrate the third to sixth leaf (from top) using a 3 mL syringe without needle (*see* **Note 6**).
9. Collect leaf tissues 60–72 h after infiltration and quickly freeze in liquid nitrogen. Tissues can be stored at −80 °C for several months.

3.1.2 Purification of GFP–AGO1 with Anti-GFP Antibodies

Unless otherwise indicated, all the following procedures in this section should be conducted in cold conditions (e.g., in a 4 °C cold room or on ice), and all the buffers should be precooled. Use a microfuge to centrifuge samples.

1. Grind 0.3 g of leaf tissue using a mortar and a pestle in liquid nitrogen into fine powder. Add 1 mL protein extraction buffer, and incubate the homogenate on an end-over rotator at 4 °C for 1 h.
2. Centrifuge at 16,000 × *g* for 10 min and carefully transfer the supernatant to a new tube.
3. Repeat the centrifugation step and save a 40 μL aliquot as input sample.
4. Antibody–beads conjugation.

   Once incubation of the homogenate on an end-over rotator starts (**step 1**), conjugate the antibody to the protein A agarose beads. Take 40 μL of protein A agarose slurry per sample (50% slurry, 20 μL beads) using a 200 μL tip with wide orifice. Wash the beads with 1 mL of protein extraction buffer three times. Centrifuge at 100 × *g* for 1 min and discard the supernatant after each wash. Resuspend the beads with 600 μL of protein extraction buffer, and add 5–10 μL anti-GFP antibodies (check the product datasheet for recommended antibody dilution). Incubate for 2–3 h at 4 °C on an end-over rotating wheel (5–10 rpm), and collect the beads by centrifugation at 100 × *g* for 1 min. Wash the antibody/beads mixture with 1 mL of protein extraction buffer three times (*see* **Note 7**).
5. Preclear the protein lysate from **step 3** with 20 μL of prewashed protein A agarose beads.
6. Incubate 1 mL of total protein lysate with coupled antibody–beads for 2 h to overnight on the rotating wheel (5–10 rpm).
7. Sediment the beads by centrifuging at 100 × *g* for 1 min, and discard the supernatant. Wash the beads four to six times with 1 mL of protein extraction buffer.

3.1.3 Preparation of Radio-Labeled miR166a

1. Assemble the following reaction in a 1.5 mL tube, and incubate at 37 °C water bath for 1 h.

DEPC-treated H_2O	39 μL
T4 PNK buffer	5 μL
miR166a (10 μM)	1 μL
T4 PNK	2.5 μL
[γ-^{32}P]ATP (10 mCi/mL)	2.5 μL
Total volume	50 μL

2. To remove unincorporated [γ-^{32}P]ATP, pass the reaction through the G25 column, and centrifuge at 850 × *g* for 1 min in a microcentrifuge.
3. Phenol/chloroform extraction to remove the enzyme.

   Adjust the volume of elute to 120 μL with DEPC-treated H_2O. Add equal volume of acidic phenol pH 4.5/chloroform (1:1 premixed) and vortex vigorously for 30 s. Centrifuge at 16,000 × *g* for 5 min, and carefully transfer the aqueous phase (i.e., the top phase) to a new tube. Add 1 volume of chloroform and extract again to remove residual phenol. Carefully transfer the aqueous phase to a new tube (~100 μL).
4. Ethanol precipitation.

   Add 10 μL of 3 M NaOAc pH 4.3, 1 μL glycogen (20 μg), and 300 μL of ethanol. Mix thoroughly and incubate at −20 °C overnight or at −80 °C for at least 2 h. Centrifuge at 16,000 × *g* for 20 min and rinse the pellet with 70% ethanol. Note the position of the pellet and decant supernatant or remove the supernatant by pipetting. Be careful not to disturb the pellet. Remove the ethanol by centrifuging at 7000 × *g* for 5 min and air-dry for 3–5 min. Resolve the radio-labeled miR166a with 100 μL of RNase-free H_2O, and use an aliquot for AGO1 assembly.

3.1.4 AGO1 Assembly and Uridylation Assay

1. AGO1–miR166a assembly.

   Add 1–5 μL [^{32}P]-labeled miR166a to the purified GFP–AGO1 in 0.5 mL protein extraction buffer with 20 U of RNase inhibitor. Incubate at 4 °C for 1 h on an end-over rotating wheel (5–10 rpm), and collect the beads by centrifugation at 100 × *g* for 1 min in a microcentrifuge. Wash the beads with protein extraction buffer, and check the radioactivity of the supernatant with a Geiger counter after each wash. It usually requires three to five washes until the readout by the counter becomes stable.
2. HESO1-directed uridylation assay.

   Methods for construction and purification of MBP–HESO1 and MBP are according to [20]. Add 0.5 mL of reaction

exchange buffer to the assembled AGO1–miR166a beads, and split the beads into two aliquots; carefully remove the exchange buffer. Add 60 μL of reaction buffer to each aliquot. Add 1 μL of MBP–HESO1 or MBP (~30 ng) to each aliquot, and incubate at room temperature for 30 min. Collect the beads by centrifugation at 100 × *g* for 1 min in a microcentrifuge. Supernatant can be saved and assayed in parallel (optional). Wash the beads five times with protein extraction buffer. RNA was extracted with phenol/chloroform and precipitated with ethanol (*see* **steps 3** and **4** in Subheading 3.1.3). Resolve the washed and air-dried RNA in 10 μL of DEPC-treated H_2O. Add 10 μL of 2× formamide RNA loading dye, incubate at 70 °C for 5 min, and leave on ice.

3. RNA analysis by denaturing PAGE electrophoresis.

   The Bio-Rad PROTEAN II xi cell vertical electrophoresis system is used to analyze the pattern of AGO1-bound miRNA after HESO1 treatment. Assemble the glass plate and cast the gel according to the instruction manual. For 1.0 mm spacers, transfer 40 mL of 16% polyacrylamide denaturing gel stock solution into a 50 mL centrifuge tube, and if necessary, deaerate the solution under vacuum for 10–20 min. Add 420 μL of 10% APS and 24 μL of TEMED to the solution and mix. Pour the solution to the assembled glass plate and insert the comb in the gel sandwich. Let the gel polymerize for at least 1 h. Rinse the sample well thoroughly with 0.5× TBE running buffer. Load the samples on separate wells, and run until the bromophenol blue dye almost reaches the bottom of the gel. Remove the gel from the cassette, drain excess buffer, carefully put the gel on a Whatman paper support, and wrap it up with plastic wrap. Expose the gel using a storage phosphor screen, and the signal is monitored on a Typhoon FLA 9500 machine.

3.2 Analysis of the 3′ Uridylation of 5′ Cleavage Products of Target RNAs

Use a microfuge to centrifuge samples.

3.2.1 Extraction of Total RNA from Wild Type and the heso1–2 Mutant

1. Homogenize 0.1 g tissue samples from respective genotypes in liquid nitrogen, and transfer about 100 μL of fine powder to a 1.5 mL tube.
2. Add 1 mL of TRI Reagent (add 10 μL/mL β-mercaptoethanol just before use) and vortex thoroughly.
3. Place at room temperature for 5 min and centrifuge at 16,000 × *g* for 5 min at 4 °C.
4. Transfer the supernatant to a new 1.5 mL tube and add 200 μL of chloroform. Vortex thoroughly and incubate at room temperature for 5 min.
5. Centrifuge at 16,000 × *g* for 5 min at 4 °C, and carefully transfer the aqueous phase (~600 μL) to a new 1.5 mL tube.

6. Add 500 μL isopropanol, mix thoroughly, and incubate the samples at room temperature for 15–20 min.
7. Centrifuge at 16,000 × *g* for 10 min at 4 °C. A white RNA pellet appears at the bottom of the tube.
8. Wash the RNA pellet with 70% ethanol, and recover the pellet by centrifuging at 5,000 × *g* for 3 min.
9. Discard ethanol and air-dry the RNA for ~10 min. Dissolve the RNA in 20 μL DEPC-treated H_2O, and check the RNA concentration with a NanoDrop spectrophotometer.

3.2.2 Cloning of 5′ Cleavage Product by RNA Ligase-Mediated 3′ Rapid Amplification of cDNA Ends (RLM-3′ RACE)

1. 3′ adaptor ligation.

   Assemble the following mixture in a nuclease-free PCR tube:

DEPC-treated H_2O	Variable
3′ RNA adaptor (100 pmol/μL)	1 μL
Total RNA (500 ng–1 μg)	1–11 μL
Total volume	12 μL

   Incubate at 70 °C for 2 min and quickly chill on ice. Add the following components to the tube and mix by pipetting. Incubate at 22 °C for 8 h.

10× T4 RNA ligase buffer	2 μL
ATP (10 mM)	2 μL
RNase Inhibitor (40 U/μL)	1 μL
DMSO	2 μL
T4 RNA ligase 1 (10 U/μL)	1 μL
Total volume	8 μL

   After incubation, perform the phenol/chloroform extraction and ethanol precipitation (*see* **steps 3** and **4** in Subheading 3.1.3).

2. Reverse transcription, PCR, and detection.

   Resuspend RNA in 10 μL of DEPC-treated H_2O, and use 5 μL of ligated RNA for reverse transcription. Combine the following components in a nuclease-free PCR tube:

Ligated RNA	5 μL
RT primer (25 μM)	1 μL
dNTP (10 mM)	1 μL
DEPC-treated H_2O	6 μL
Total volume	13 μL

Incubate at 65 °C for 5 min and quickly chill on ice. Add the following components to each tube and mix by pipetting. Incubate at 50 °C for 1 h and place the tube on ice.

5× first-strand buffer	4 μL
RNase inhibitor (40 U/μL)	1 μL
DTT (100 mM)	1 μL
SuperScript III Reverse Transcriptase (200 U/μL)	1 μL
Total volume	7 μL

Perform first round PCR using P1 and RT primer. Assemble the following components in a PCR tube as follows:

cDNA	1 μL
H_2O	13.9 μL
10× PCR reaction buffer	2 μL
10 mM dNTP	1 μL
P1 primer (10 μM)	1 μL
RT primer (10 μM)	1 μL
Taq DNA polymerase	0.1 μL
Total volume	20 μL

Perform 25–30 cycles of PCR with default annealing temperature at 55 °C (*see* **Note 8**). Dilute the first round PCR product 50 times, and use 1 μL as template for second round PCR, which uses P2 and RT primer. Check the PCR product by running a 1.5% agarose gel. Perform gel purification and subclone the PCR product to a pGEM-T Easy Vector. Sequence the positive clones with the M13F sequencing primer. The 3′ end signature of the cloned 5′ cleavage product can be retrieved subsequently.

3.2.3 Cloning of 5′ Cleavage Product by Circulation-Mediated Rapid Amplification of cDNA Ends (cRACE)

1. 5′ end RNA dephosphorylation by CIP treatment (optional).

   Assemble the following reaction in a 1.5 mL nuclease-free tube:

Total RNA (~1 μg/μL)	6 μL
H_2O	19 μL
10× CIP buffer	3 μL
RNase inhibitor (40 U/μL)	1 μL
CIP (10 U/μL)	1 μL
Total volume	30 μL

Mix by pipetting and incubate at 37 °C for 45 min. Perform the phenol/chloroform extraction and ethanol precipitation (*see* **steps 3** and **4** in Subheading 3.1.3). Resuspend CIP-treated RNA in 9 μL of DEPC-treated H_2O. Use 1–2 μL to check the RNA concentration with a NanoDrop spectrophotometer.

2. 5′ end RNA decapping by TAP treatment.

   Assemble the following reaction in a 1.5 mL nuclease-free tube.

CIP-treated RNA or total RNA (~2–5 μg)	7 μL
H_2O	26 μL
10× TAP buffer	4 μL
RNase inhibitor (40 U/μL)	1 μL
TAP (10 U/μL)	2 μL
Total volume	40 μL

   Mix by pipetting and incubate at 37 °C for 2 h. Perform the phenol/chloroform extraction and ethanol precipitation (*see* **steps 3** and **4** in Subheading 3.1.3).

3. RNA self-circulation.

   Resuspend the TAP-treated RNA in 16 μL of DEPC-treated H_2O. Use 1–2 μL to check the RNA concentration with a NanoDrop spectrophotometer. Incubate the RNA at 65 °C for 5 min and chill on ice. Add the following reagent to the RNA (16 μL) and mix by pipetting. Incubate at 22 °C for 8 h.

10× T4 RNA ligase buffer	2 μL
ATP (10 mM)	2 μL
RNase inhibitor (40 U/μL)	1 μL
T4 RNA ligase 1 (10 U/μL)	1 μL
Total volume	6 μL

   After incubation, perform the phenol/chloroform extraction and ethanol precipitation (*see* **steps 3** and **4** in Subheading 3.1.3).

4. Reverse transcription, PCR, and detection.

   Resuspend ligated RNA in 15 μL of DEPC-treated H_2O. Split the RNA solution into three to four aliquots for different 5′ CPs. For each 5′ CP, use 2 pmol of R1 primer instead of RT primer, and perform reverse transcription (*see* **step 2** in Subheading 3.2.2). Nested PCR is performed as

described (*see* **step 2** in Subheading 3.2.2), except that the RT primer is replaced by R1 and R2 in the first and second round PCR, respectively. PCR product subcloning and Sanger sequencing are performed as described (*see* **step 2** in Subheading 3.2.2). Both 3′ end and 5′ end signatures are retrieved by cRACE (*see* **Note 9**).

4 Notes

1. DEPC is highly toxic. Handling of DEPC should be carried out carefully in a fume hood with good protection. The bottle cap should be tightly sealed with Parafilm after DEPC is added.
2. DTT, PMSF, and protease inhibitor should be added just before use.
3. The stock solution can be stored at 4 °C for several months without affecting performance. Old stocks of acrylamide/bis-acrylamide (29:1) may result in poor resolution (i.e., a smear rather than discrete bands appear after exposure).
4. TAP is no longer distributed. Tebu-bio's Decapping Pyrophosphohydrolase or CellScript's Cap-Clip™ Acid Pyrophosphatase could be possible replacement (http://being–bioreactive.com/2015/09/03/two–new–enzymes–available–to–replace–tap/).
5. When starting from an older plate or a frozen stock, recover the agrobacteria by streaking it on a YEB agar plate 3–5 days before inoculation.
6. We found that young and healthy leaves show much higher protein expression, whereas the first two top leaves are difficult to be infiltrated.
7. If using GFP-Trap or other pre-coupled antibody/agarose beads, simply wash the resin with protein extraction buffer three times before using.
8. Annealing temperature may be further optimized, and touch-down PCR may be performed if the PCR result is not satisfactory.
9. End nucleotides that match both 5′ end and 3′ end of the genome cannot be distinguished with this method.

Acknowledgments

We thank Yuan Wang for helping with the preparation of the figure. This work was supported by the Ministry of Science and Technology of China (Grant No. 2016YFA0503200) and National Natural

Science Foundation of China (Grant Nos. 31471221 and 31622009) to G.R. and Nebraska Soybean Board (Grant No. 16R–05–3/3 #1706) and National Science Foundation (Award No. OIA–1557417) to B.Y.

No conflict of interest was declared.

References

1. Chen X (2009) Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol 25:21–44. doi:10.1146/annurev.cellbio.042308.113417
2. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233. doi:10.1016/j.cell.2009.01.002
3. Ren G, Yu B (2012) Critical roles of RNA-binding proteins in miRNA biogenesis in Arabidopsis. RNA Biol 9(12):1424–1428. doi:10.4161/rna.22740
4. Ha M, Kim VN (2014) Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15(8):509–524. doi:10.1038/nrm3838
5. Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. Science 307(5711):932–935. doi:10.1126/science.1107130
6. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110(4):513–520. doi:10.1016/S0092-8674(02)00863-2
7. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798. doi:10.1016/S0092-8674(03)01018-3
8. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
9. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8(4):517–527. doi:10.1016/j.devcel.2005.01.018
10. Yekta S, Shih IH, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. Science 304(5670):594–596. doi:10.1126/science.1097434
11. Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP (2010) Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell 38(6):789–802. doi:10.1016/j.molcel.2010.06.005
12. Zhang X, Zhu Y, Liu X, Hong X, Xu Y, Zhu P, Shen Y, Wu H, Ji Y, Wen X, Zhang C, Zhao Q, Wang Y, Lu J, Guo H (2015) Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science 348(6230):120–123. doi:10.1126/science.aaa2618
13. Li J, Yang Z, Yu B, Liu J, Chen X (2005) Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. Curr Biol 15(16):1501–1507. doi:10.1016/j.cub.2005.07.029
14. Lee M, Kim B, Kim VN (2014) Emerging roles of RNA modification: m(6)A and U-tail. Cell 158(5):980–987. doi:10.1016/j.cell.2014.08.005
15. Ren G, Chen X, Yu B (2014) Small RNAs meet their targets: when methylation defends miRNAs from uridylation. RNA Biol 11(9):1099–1104. doi:10.4161/rna.36243
16. Ren G, Xie M, Zhang S, Vinovskis C, Chen X, Yu B (2014) Methylation protects microRNAs from an AGO1-associated activity that uridylates 5′ RNA fragments generated by AGO1 cleavage. Proc Natl Acad Sci U S A 111(17):6365–6370. doi:10.1073/pnas.1405083111
17. Shen B, Goodman HM (2004) Uridine addition after microRNA-directed cleavage. Science 306(5698):997. doi:10.1126/science.1103521
18. Tu B, Liu L, Xu C, Zhai J, Li S, Lopez MA, Zhao Y, Yu Y, Ramachandran V, Ren G, Yu B, Li S, Meyers BC, Mo B, Chen X (2015) Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in Arabidopsis. PLoS Genet 11(4):e1005119. doi:10.1371/journal.pgen.1005119
19. Wang X, Zhang S, Dou Y, Zhang C, Chen X, Yu B, Ren G (2015) Synergistic and independent actions of multiple terminal nucleotidyl transferases in the 3′ tailing of small RNAs in Arabidopsis. PLoS Genet 11(4):e1005091. doi:10.1371/journal.pgen.1005091
20. Ren G, Chen X, Yu B (2012) Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr Biol 22(8):695–700. doi:10.1016/j.cub.2012.02.052
21. Zhao Y, Yu Y, Zhai J, Ramachandran V, Dinh TT, Meyers BC, Mo B, Chen X (2012) The Arabidopsis nucleotidyl transferase HESO1

uridylates unmethylated small RNAs to trigger their degradation. Curr Biol 22(8):689–694. doi:10.1016/j.cub.2012.02.051

22. Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33(5):949–956. doi:10.1046/j.1365-313X.2003.01676.x
23. Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19(3):421–428. doi:10.1016/j.molcel.2005.06.014

Chapter 3

In Vitro Formation of Plant RNA-Induced Silencing Complexes Using an Extract of Evacuolated Tobacco Protoplasts

Taichiro Iki, Masayuki Ishikawa, and Manabu Yoshikawa

Abstract

Small RNA-mediated gene silencing is involved in a variety of biological processes among many eukaryotic organisms. The silencing effector, generally referred to as RNA-induced silencing complex (RISC), comprises an ARGONAUTE (AGO) protein and a small single-stranded guide RNA in its core. RISCs recognize target genes containing sequences complementary to the guide RNA and repress their expression transcriptionally or posttranscriptionally. In vitro systems that recapitulate RISC assembly are useful not only to decipher the molecular mechanisms underlying the assembly process itself but also to dissect the downstream silencing pathways mediated by RISCs. Here, we describe a method for in vitro plant RISC assembly, which relies on an extract of evacuolated protoplasts derived from *Nicotiana tabacum* BY-2 suspension-cultured cells. In this extract, synthetic duplexes of small RNAs are incorporated into AGO proteins that are synthesized by in vitro translation, and then duplex unwinding and selective strand elimination result in formation of mature RISCs.

Key words miRNA, siRNA, RISC, ARGONAUTE, Tobacco

1 Introduction

It is well known that <100-nucleotide (nt) small RNAs play important roles in gene regulation in both prokaryotes and eukaryotes. In general, these small RNAs carry out their functions in complexes with one or more proteins. A representative example includes the 19–28-nt small RNAs that engage in the eukaryotic gene silencing pathways. Among such small RNAs, small interfering RNAs (siRNAs) and most microRNAs (miRNAs) are excised as duplexes from inter- or intramolecularly double-stranded RNAs (dsRNAs) by RNase III-like enzymes such as Drosha and Dicer in animals or DICER-LIKE (DCL) in plants. The small RNA duplexes are then incorporated into ARGONAUTE (AGO) proteins, followed by the selective elimination of one strand of the duplex, referred to as a passenger strand or miRNA* [1]. Finally, the reaming guide

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_3,

strand and AGO combine to form the primary effector of gene silencing, designated RNA-induced silencing complex (RISC). Through the recognition of complementary sequences by the guide strands, RISCs transcriptionally or posttranscriptionally repress the target genes in a sequence-specific manner.

In the model plant *A. thaliana* whose genome encodes ten *AGO* genes, AGO1 plays key roles in developmental and physiological events via posttranscriptional gene silencing [2]. AGO1-containing RISCs induce cleavage of target RNA and/or repress the translation of target mRNAs by slicer-independent mechanisms.

In vitro RISC formation by exogenous addition of a specific small RNA is referred to as "programming." Previous work has demonstrated that AGO1 in crude plant extracts or immunopurified AGO1 can be programmed by single-stranded siRNAs [3]. The formed RISCs exhibit sequence-specific RNA cleavage activity directed by the guide siRNAs. However, the biogenesis of many small RNAs relies on DCL proteins that generate small RNAs as duplexes with diagnostic 2-nt 3′ overhangs. Thus, duplex-initiated programming has been considered the canonical RISC formation pathway in vivo. We have established an in vitro system where AGO1 programming is initiated using synthetic small RNA duplexes in an extract of evacuolated protoplasts from *Nicotiana tabacum* BY-2 cultured cells (BYL) [4]. In this system, AGO1 protein is synthesized by in vitro translation using BYL, and then small RNA duplexes are incorporated into the AGO1 proteins followed by duplex unwinding and formation of small RNA-programmed RISCs.

Plant AGO proteins exhibit preference among the 5′ nucleotides of guide RNAs, and AGO1 prefers those with 5′ uracil (U) [5–8]. This preference is based on the MID domain of AGO1, which contains the binding pocket for the guide RNA 5′ base and exhibits a higher affinity to U than for other bases [9]. A 22-nt siRNA duplex containing 5′ U, which is incorporated into AGO1 complexes in the leaves of *N. benthamiana* expressing an inverted repeat of the green fluorescent protein (GFP) mRNA sequence [10], has been tested for loading into AGO1 in BYL [4]. Synthetic miRNA duplexes containing 5′ U have been used to demonstrate miRNA duplex loading. Consistent with the in vivo observation, these small RNAs with 5′ U that serve as guide strands are selectively associated with AGO1, while the other strands in the duplexes are eliminated in BYL-based AGO1 programming. Similarly, other AGO proteins show differential preference among 5′ nucleotides, and these features are also recapitulated in BYL [11, 12].

The BYL-based AGO1 programming system has deepened our understanding of RISC loading [4]. In BYL, RISC loading requires ATP and a molecular chaperone, HSP90. A poorly hydrolyzable ATP analog stabilizes the interaction between HSP90 and AGO1, resulting in accumulation of complexes that contain HSP90, AGO1, and small RNA duplexes, most likely due to inhibition of

HSP90-mediated ATP hydrolysis. HSP90 dissociates from AGO1 after ATP hydrolysis, which presumably induces AGO1 conformational changes and unwinding of small RNA duplex. Subsequent BYL-based study has identified several HSP90 co-chaperones in HSP90-AGO1 complexes. Among them, cyclophilin 40 (CYP40) has been characterized as a unique co-chaperone that facilitates HSP90-mediated AGO1-RISC loading [13]. This finding is consistent with the genetic evidence that the *SQUINT* gene, encoding CYP40 in *A. thaliana*, is required for miRNA activity in vivo [14].

The in vitro RISC assembly system has been also used to examine the activities of RISCs, including target RNA cleavage and mRNA translational repression [15]. Recent advances using this system include studies on RISC-mediated antiviral defense, the molecular activity of tombusvirus P19, a well-established viral suppressor of RNA silencing [16], and the initiation of secondary siRNA biogenesis [17]. To further expand our understanding of the formation and activity of plant RISCs, the BYL-based RISC formation system can serve as a practical platform for in vitro biochemical analysis. This chapter describes in detail the methods of BYL preparation and formation of RISCs in BYL.

2 Materials

2.1 Equipment

1. Centrifuge with a swing rotor and 15 mL tubes that may be used up to 10,000 × *g* (e.g., Avanti HP-25, JS24.15, and 16 × 96 mm polyallomer tubes, Beckman Coulter).
2. Gradient forming instrument, used to make a Percoll linear gradient (e.g., Gradient Mate, BioComp).
3. Rotator (e.g., RT-50 and SC-0200, TAITEC).
4. Tabletop ultracentrifuge and rotor that can be operated up to 30,000 × *g* (e.g., Optima MAX-TL and TLA-100, Beckman Coulter).
5. 7 mL Dounce homogenizer (e.g., 7 mL Dounce tissue grinder 357524, Wheaton).

2.2 Medium, Solutions, and Buffers

1. Tobacco (*Nicotiana tabacum*) BY-2 suspension-cultured cells: 0.46% (w/v) Murashige and Skoog Plant Salt Mixture, 3% (w/v) sucrose, 200 μg/mL monopotassium phosphate (KH_2PO_4), 100 μg/mL myoinositol, 0.2 μg/mL 2,4-dichlorophenoxyacetic acid

   1 μg/mL thiamine-HCl, distilled water. Adjust the pH to 5.8 with 1 M potassium hydroxide (KOH). Prepare and pour 100 mL of the medium into a 300 mL conical flask. Cap with a silicon cap or aluminum foil, autoclave, and store at room temperature until use.

2. Protoplast wash solution: 12.5 mM with sodium acetate (CH_3COONa), 5 mM calcium chloride ($CaCl_2$), 0.37 M mannitol, distilled water. Adjust the pH to 5.8 with 1 M acetic acid (CH_3COOH). Prepare on the day of the experiment, or autoclave and store at room temperature until use.
3. Protoplast isolation enzyme solution (PIM): add 1.0 g of cellulase Onozuka RS (Yakult Pharmaceutical) and 0.1 g of pectolyase Y-23 (Kyowa Chemical Products) to 100 mL of protoplast wash solution. Prepare on the day of the experiment.
4. 0% Percoll solution: prepare a 60 mL solution by mixing 7.65 g of mannitol, 1.2 mL of 1 M $MgCl_2$, 3 mL 0.1 M PIPES buffer (pH 7.0), and distilled water to volume.
5. 70% Percoll solution: prepare a 60 mL solution by mixing 7.65 g of mannitol, 1.2 mL of 1 M $MgCl_2$, 3 mL 0.1 M PIPES buffer (pH 7.0), 42 mL Percoll (GE Healthcare), and distilled water to volume.
6. 30% Percoll solution: mix 16 mL of 0% Percoll solution with 12 mL 70% Percoll solution.
7. 40% Percoll solution: mix 12 mL of 0% Percoll solution with 16 mL 70% Percoll solution.
8. Translation (TR) buffer: 30 mM HEPES-KOH (pH 7.4), 80 mM CH_3COOK, 1.8 mM $(CH_3OO)_2Mg$, 2 mM dithiothreitol (DTT), 1 tablet cOmplete Mini (Roche) for 10 mL of buffer, RNase-free water to volume. Prepare, aliquot, and store at −20 °C until use.
9. 10× translation mix: 7.5 mM ATP, 1 mM GTP, 250 mM creatine phosphate (Roche), amino acid mix (500 μM each, Promega), 80 μM spermine, RNase-free water to volume. Prepare, aliquot, and store at −20 °C until use.
10. 5× annealing buffer: 50 mM Tris–HCl (pH 7.6), 100 mM KCl, 5 mM $MgCl_2$, RNase-free water to volume. Prepare, aliquot, and store at −20 °C until use.
11. 10× ATP mix: 7.5 mM ATP, 1 mM $MgCl_2$, 200 mM creatine phosphate, 2 mg/mL creatine phosphokinase, RNase-free water to volume. Prepare on the day of the experiment.
12. 2× native gel sample buffer: 0.2 mg/mL bromophenol blue, 0.2 mg/mL xylene cyanol, 1× TBE buffer, 10% (v/v) glycerol, RNase-free water to volume. Prepare, aliquot, and store at −20 °C until use.
13. 15% native acrylamide gel: prepare a 10 mL solution by mixing 5 mL of 30% polyacrylamide solution (acrylamide/bisacrylamide 37.5:1), 1 mL of 5× TBE, 4 mL of RNase-free water, 10 μL of 10% ammonium persulfate, and 10 μL of tetramethylethylenediamin (TEMED). Prepare on the day of the experiment.

14. 1.6× urea dye: 0.2 mg/mL bromophenol blue, 0.2 mg/mL xylene cyanol, 12.5 M urea, RNase-free water to volume. Prepare, aliquot, and store at −20 °C until use.
15. 7 M-urea-5% acrylamide gel: prepare a 10 mL solution by mixing 4.2 g urea, 1.7 mL 30% polyacrylamide solution (acrylamide/bis-acrylamide 37.5:1), 1 mL 5× TBE, 4.2 mL RNase-free water, 10 μL 10% ammonium persulfate, 10 μL TEMED. Prepare on the day of the experiment.
16. AmpliCap SP6 High Yield Message Maker Kit (CELLSCRIPT).
17. Mini Quick Spin RNA Column (Roche).
18. Phenol/chloroform.
19. 3 M CH_3COONa (pH 5.2).
20. Ethanol.
21. T4 polynucleotide kinase.
22. ATP.
23. MicroSpin G-25 Column (GE Healthcare).
24. Glycogen.
25. 111 TBq/mmol [γ-^{32}P]ATP.
26. 29.6 TBq/mmol [α-^{32}P]CTP.
27. TE buffer: 10 mM Tris–HCl, 1 mM EDTA-2Na pH 8.0.
28. TE-saturated phenol.
29. SP6-Scribe Standard RNA IVT Kit (CELLSCRIPT).
30. ScriptCap m^{7}G Capping System kit (CELLSCRIPT).

3 Methods

3.1 Preparation of Evacuolated Protoplast Extracts

1. Maintain tobacco BY-2 suspension-cultured cells by transferring 0.5–2.0 mL of cell suspension to 100 mL of fresh medium at 7-day intervals. Cells are cultured in the darkness at 26–27 °C in a rotary shaker at ~130 rpm. To prepare protoplasts, transfer 5–10 mL of 7-day-old cells to 100 mL of fresh medium, and culture the cell suspension under the same conditions for 3 days (*see* **Note 1**).
2. Transfer 2 × 100 mL of 3-day-old cell culture to four 50 mL conical centrifuge tubes, and collect the cells by centrifugation at 150 × *g* for 2–3 min at room temperature.
3. Discard the supernatant by decanting as much as possible.
4. Suspend the cell pellets in PIM solution, and divide the suspension equally between two bottles of 200 mL conical flasks.
5. Incubate the cell suspension at room temperature for 2–3 h, using a rotator for gentle swirling.

6. During the incubation for protoplast production, prepare a Percoll gradient using Percoll solutions. To prepare a 0–30% linear gradient, add 3.5 mL of 0% Percoll solution to each of six centrifuge tubes, and then add 3.5 mL of 30% Percoll solution to the bottom of the tube using a 230-mm-long Pasteur pipette. After, run the gradient protocol using a gradient forming instrument, and add 3.5 mL of 40% Percoll solution and then 2 mL of 70% Percoll solution to the bottom of the tube.
7. Transfer the protoplast suspension to two 50 mL conical centrifuge tubes, and pellet the protoplasts by centrifuging for 5–10 min at 150 × *g* at room temperature. Discard the supernatant by decantation.
8. Add protoplast wash solution to the tubes (≈50 mL), and mix gently until the protoplasts are fully suspended.
9. Centrifuge for 5–10 min at 150 × *g* and room temperature, and then discard the supernatant by decanting.
10. Repeat the wash (**steps 8** and **9**) two more times.
11. After discarding the supernatant, resuspend the protoplasts in the remaining wash solution, and combine the protoplast suspensions in one tube (~15 mL).
12. Gently overlay the protoplasts onto the top of the Percoll gradient prepared at **step 6** using a Pasteur pipette.
13. Centrifuge with a swing rotor at 10,000 × *g* and 20–25 °C for 1 h.
14. Using a Pasteur pipette, remove the layers above the 40–70% Percoll boundary, which contain vacuole-rich vesicles, undivided protoplasts, and incompletely evacuolated protoplasts. Leave 1–2 mL of the 40% Percoll layer to avoid contamination by vacuole-containing vesicles of the evacuolated protoplasts that accumulate at the 40–70% Percoll boundary (*see* **Note 2**).
15. Working on ice, collect the evacuolated protoplasts in a new 15 mL conical centrifuge tube using a clean Pasteur pipette.
16. Add precooled protoplast wash solution (~15 mL) to the tube, suspend protoplasts in the wash solution, and then recollect the evacuolated protoplasts by centrifugation at 150 × *g* at 4 °C for 10 min.
17. Decant and discard the supernatant. Add precooled wash solution, mix gently until the pellets are invisible, and centrifuge at 150 × *g* at 4 °C for 10 min. Repeat the washing step for a total of three times.
18. Remove as much as possible the supernatant using a micropipette.
19. Add four volumes of TR buffer to the evacuolated protoplasts, resuspend, and transfer to a 7 mL tight-fitting Dounce homogenizer that had been precooled on ice. Observe under

a microscope to ensure that protoplasts are free of vacuoles (*see* **Note 3**).

20. Homogenize the evacuolated protoplasts with ≈100 strokes, working on ice. Check a microscope for thorough homogenization.
21. Transfer the lysate to 1.5 or 2.0 mL microtubes on ice.
22. Centrifuge at 800 × *g* at 4 °C for 10 min to pellet nuclei and unbroken cells.
23. Transfer the supernatant (BYL) to a new tube.
24. Store aliquots (e.g., 200 μL) in 1.5 mL microtubes at −80 °C until use (*see* **Note 4**).

3.2 Preparation of FLAG-Tagged Tobacco AGO1 mRNA

1. The pSP64Poly(A) vector-derivative plasmid containing a gene that encodes FLAG-tagged tobacco AGO1 (*FLAG-NtAGO1*) under the SP6 promoter was linearized with *Sma*I. *FLAG-NtAGO1* mRNA was prepared using the AmpliCap SP6 High Yield Message Maker Kit. Set up the following reaction mixture in a microtube:

	Amount for 20 μL
Linearized template DNA (1 μg)	X μL
10× AmpliCap SP6 transcription buffer	2 μL
Cap/NTP PreMix	5 μL
100 mM DTT	2 μL
ScriptGuard RNase inhibitor	0.5 μL
AmpliCap SP6 enzyme solution	2 μL
RNase-free water	Y μL

2. Incubate at 37 °C for 30 min, add 1 μL of 10 mM GTP, and continue incubation at 37 °C for 1.5–2.5 h.
3. Add 1 μL of RNase-free DNase I and incubate at 37 °C for 15 min.
4. Add 29 μL of RNase-free water, and load the entire reaction onto a mini Quick Spin RNA Column.
5. Adjust the volume of the flow-through to ≈100 μL with RNase-free water, add 100 μL phenol/chloroform, and vortex vigorously. Centrifuge at maximum speed at room temperature for 5 min.
6. Transfer the aqueous phase to a new tube, add 10 μL of 3 M CH_3COONa (pH 5.2) and 250 μL of ethanol, and mix well. Centrifuge at maximum speed at 4 °C for 15 min. Discard the

supernatant, add 160 μL of 70% ethanol, centrifuge at maximum speed at 4 °C for 5 min, and then discard the supernatant.

7. Dissolve the pelleted RNA in RNase-free water after briefly drying up, measure the concentration, and adjust it to 0.5 μg/μL.
8. Store the mRNA solution at −80 °C until use.

3.3 Preparation of siRNA Duplexes for RISC Loading

1. Synthesize an siRNA and a complementary strand so that they form a duplex with 2-nt 3′ overhangs. The 3′ ends of both RNAs are 2′-*O*-methylated, and their 5′ ends are unmodified. The siRNAs are purified by HPLC.
2. To phosphorylate the 5′ end of siRNAs, set up the following reaction mixture separately for each siRNA:

	Amount for 20 μL
100 μM siRNA or complementary siRNA	2 μL
10× T4 polynucleotide kinase reaction buffer	2 μL
10 mM ATP	2 μL
10 U/μL T4 polynucleotide kinase	1 μL
RNase-free water	13 μL

3. Incubate for 1.5 h at 37 °C and then for a further 10 min at 65 °C.
4. Combine the two reactions and load onto a MicroSpin G-25 Column.
5. Adjust volume to ≈100 μL with RNase-free water, add 100 μL of phenol/chloroform, and vortex vigorously. Centrifuge at maximum speed at room temperature for 5 min.
6. Transfer the aqueous phase to a new tube. Add 1 μL of 20 mg/mL glycogen, 10 μL of 3 M sodium acetate (pH 5.2), and 250 μL of ethanol and then mix. Centrifuge at maximum speed at 4 °C for 15 min. Discard the supernatant, add 160 μL 70% ethanol, centrifuge at maximum speed at 4 °C for 5 min, and discard the supernatant again.
7. Dry the pellet at room temperature for a few minutes.
8. Suspend in 200 μL of 1× annealing buffer.
9. Incubate at 96 °C for 5 min, at 55 °C for 30 min, and at 25 °C for another 30 min with gradual cooling during the intervals.
10. Store at −80 °C until use. The final concentration of the siRNA duplex is 1 μM.

3.4 Preparation of ^{32}P-Labeled siRNA Duplexes

1. Follow **steps 1–9** from Subheading 3.3, omitting **step 2**. Set up the following reaction mixture for ^{32}P labeling:

	^{32}P-labeled strand (μL)	Unlabeled strand (μL)
5 μM siRNA or complementary siRNA	2	2.2
10× T4 polynucleotide kinase reaction buffer	2	2
50 μM ATP	2	2
111 TBq/mmol [γ-^{32}P]ATP	2	–
10 U/μL T4 polynucleotide kinase	1	1
RNase-free water	11	13

2. Store the ^{32}P-labeled siRNA duplex at −80 °C or −20 °C until use. The final concentration of ^{32}P-labeled siRNA duplex is 50 nM.

3.5 RISC Loading

1. Thaw a 200 μL aliquot of BYL, centrifuge at 30,000 × *g* at 4 °C for 15 min, and collect the supernatant (BYL-S30).
2. To synthesize FLAG-NtAGO1 in BYL, prepare the reaction mixture for in vitro translation as follows:

	Amount for 100 μL
BYL-S30	50 μL
10× translation mix	10 μL
40 U/μL RNase inhibitor	2 μL
10 mg/mL creatine phosphokinase	2 μL
TR	31 μL
0.5 μg/μL FLAG-NtAGO1 mRNA or RNase-free water (mock)	5 μL

3. Incubate for 1–1.5 h at 25 °C.
4. To check for synthesis of FLAG-NtAGO1, take 3 μL of the reaction mixture and mix with 12 μL of 1.25× SDS-PAGE sample buffer, and analyzed by SDS-PAGE.
5. Add 11 μL of 10 × ATP mix and 5.5 μL of 50 nM ^{32}P-labeled siRNA duplex (from Subheading 3.4) to the in vitro translation reaction (97 μL), and mix.
6. Incubate at 25 °C for 30–40 min.

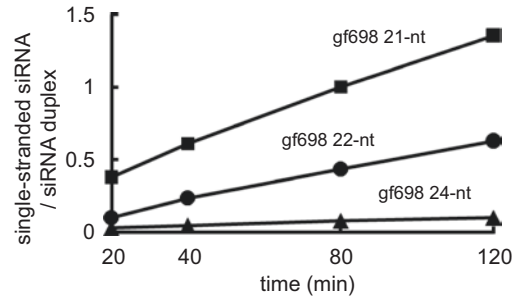


Fig. 1 Incorporation of 21-, 22-, or 24-nt siRNAs (gf698) complementary to a portion of the GFP mRNA (GF-s) in AGO1. ^{32}P-labeled gf698 duplexes (50 nM) were mixed with FLAG-NtAGO1 produced in the lysate and incubated for the indicated period. The ratio between single-stranded siRNA and siRNA duplex was calculated from the radioactivity of generated single-stranded siRNAs and siRNA duplexes

7. Mix 4 μL of the reaction with 16 μL TE and 20 μL of TE-saturated phenol, and vortex vigorously.
8. Centrifuge at maximum speed for 5 min at room temperature.
9. Mix 5 μL of the aqueous phase with 5 μL of 2× native gel sample buffer.
10. Load 2 μL of the sample onto a 15% native polyacrylamide gel, electrophorese in 0.5× TBE, and dry the gel using a gel dryer. Detect radioactivity using image analyzer. An example of the results is shown in Fig. 1.

3.6 Preparation of ^{32}P-Labeled Target RNA

1. For in vitro transcription of ^{32}P-labeled target RNAs, prepare a linear template DNA with a SP6 promoter by digesting plasmids with restriction enzymes or by PCR. Do not use restriction enzymes or DNA polymerases that produce 3′ overhangs.
2. Prepare the reaction mixture for in vitro transcription using the SP6-Scribe Standard RNA IVT Kit.

	Amount for 20 μL
Linearized DNA (0.1–0.3 μg)	X μL
10 × SP6-Scribe Transcription Buffer	2 μL
100 mM DTT	2 μL
100 mM ATP	1 μL
100 mM GTP	1 μL
100 mM UTP	1 μL
1 mM CTP (prepare from 100 mM CTP on the day of experiment)	1 μL

(continued)

	Amount for 20 µL
29.6 TBq/mmol [α-^{32}P]CTP	2 µL
40 U/µL ScriptGuard RNase inhibitor	0.5 µL
SP6-scribe enzyme solution	2 µL
RNase-free water	Y µL

3. Incubate for 2–3 h at 37 °C.
4. Add 1 µL of 1 U/µL RNase-free DNase I and incubate for 15 min at 37 °C.
5. Add 29 µL of RNase-free water and mix well.
6. To quantify the products, take 1 µL of the reaction into a tube containing 19 µL of RNase-free water.
7. Load the remaining reaction onto a mini Quick Spin RNA Column, following the manufacturer's instructions.
8. Adjust the volume of the flow-through to ≈100 µL with RNase-free water, then add 100 µL of phenol/chloroform, and vortex vigorously Centrifuge at maximum speed for 5 min at room temperature.
9. Transfer the upper aqueous phase to a new tube.
10. Add 1 µL of 20 mg/mL glycogen, 10 µL of 3 M sodium acetate (pH 5.2), and 250 µL of ethanol, and then mix.
11. Place at −20 °C for 15 min, and then centrifuge at maximum speed at 4 °C for 15 min.
12. Remove the supernatant, add 500 µL of 70% ethanol, and then centrifuge at maximum speed at 4 °C for 5 min.
13. Decant the supernatant, and allow the pellet to dry at room temperature for a few minutes.
14. Add 34.25 µL of RNase-free water and incubate at 65 °C for 10 min, and then quickly transfer the reaction tube onto ice.
15. To cap the target RNA, set up the following reaction mixture using the ScriptCap m^{7}G Capping System kit (*see* **Note 5**):

	Amount for 50 µL
Uncapped ^{32}P-labeled target RNA	34.25 µL
10 × ScriptCap capping buffer	5 µL
10 mM GTP	5 µL
2 mM *S*-adenosylmethionine (prepare from 20 mM *S*-adenosylmethionine on the day of experiment)	2.5 µL
40 U/µL ScriptGuard RNase inhibitor	1.25 µL
10 U/µL ScriptCap capping enzyme	2 µL

16. Incubate for 1 h at 37 °C.
17. Load the reaction onto a mini Quick Spin RNA Column by following the manufacturer's instructions.
18. Adjust the volume of the flow-through to ~100 μL with RNase-free water, then add 100 μL of phenol/chloroform, and vortex vigorously. Centrifuge at maximum speed for 5 min at room temperature.
19. Transfer the upper, aqueous phase to a new tube.
20. Add 1 μL of 20 mg/mL glycogen, 10 μL of 3 M sodium acetate (pH 5.2), and 250 μL of ethanol, and mix.
21. Place at −20 °C for 15 min, and then centrifuge at maximum speed at 4 °C for 15 min.
22. Remove the supernatant, add 500 μL of 70% ethanol, and centrifuge at maximum speed at 4 °C for 5 min.
23. Remove the supernatant, and dry the pellet at room temperature for a few minutes.
24. Add 50 μL of RNase-free water, dissolve well, and transfer 1 μL to a tube containing 19 μL of RNase-free water.
25. Measure the radioactivity of the samples from **steps 6** to **24** of this section using a scintillation counter.
26. Calculate the concentration of products based on radioactivity, the number of cytosine residues, and the initial concentration of cytidine triphosphate at **step 2** in Subheading 3.4.
27. Adjust the concentration of target RNA (usually to 50 nM) with RNase-free water.
28. Store at −20 °C or −80 °C until use.

3.7 Assay of Target Cleavage by Programmed RISCs

1. Follow **steps 1–4** of Subheading 3.5.
2. Add 11 μL of 10× ATP mix and 5.5 μL of 1 μM siRNA duplex to the in vitro translation reaction (113.5 μL), and mix (*see* **Note 6**).
3. Incubate at 25 °C for 30–40 min.
4. Add 2 μL of 50 nM target RNA into 18 μL of siRNA-programmed RISC reaction, and incubate for 15 min at 25 °C.
5. Mix 4 μL of the reaction with 16 μL of TE buffer and 20 μL of TE-saturated phenol, and vortex vigorously.
6. Centrifuge at maximum speed for 5 min at room temperature.
7. Mix 5 μL of the aqueous phase with 8 μL of 1.6× urea dye.
8. Load 2 μL of the sample onto a 7 M-urea, 5% polyacrylamide gel, then electrophorese with 0.5× TBE, and dry the gel using a gel dryer. Use an image analyzer to measure radioactivity. Two examples of the results are shown in Figs. 2 and 3 (*see* **Note 7**).

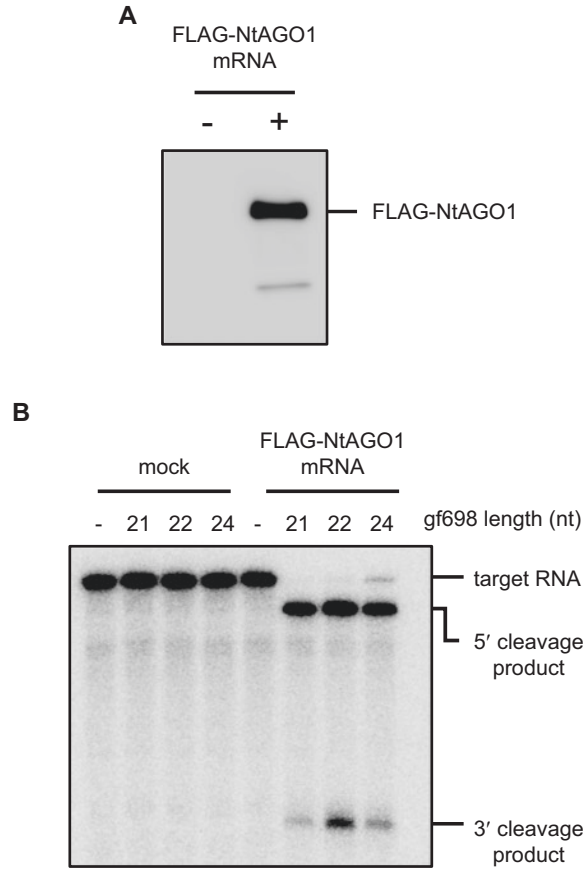


Fig. 2 Characterization of in vitro synthesized AGO1. (**a**) Synthesis of FLAG-NtAGO1 protein by in vitro translation in BYL. In vitro translation was performed as described in Subheading 3.5. Anti-FLAG antibody was used for detection of FLAG-NtAGO1 synthesized by in vitro translation. (**b**) Target cleavage by RISCs. The mock- or FLAG-NtAGO1-translated BYL were programmed by gf698 siRNA duplexes and incubated with GF-s RNA (5 nM) with the target sequence of gf698 for 15 min at 25 °C

4 Notes

1. Growth conditions for the tobacco BY-2 cells critically affect the efficiency of in vitro translation. In our experience, overgrowth decreases in vitro translation efficiency.
2. Evacuolated protoplasts should be collected carefully. Do not take the protoplasts locating on the layer between 30% Percoll and 40% Percoll.
3. The use of tight-fitting Dounce homogenizers is critical.
4. BYL can be stored at −80 °C at least 1 year. Repeated freezing and thawing of BYL should be avoided. Liquid nitrogen is not needed for freezing BYL.

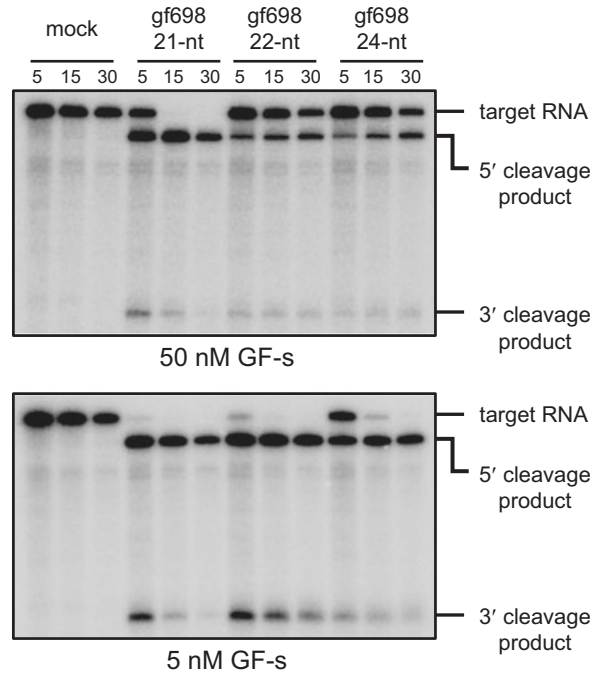


Fig. 3 Efficiency of target cleavage by RISCs programmed with 21-, 22-, or 24-nt gf698. The target cleavage assay was performed as described in Subheading 3.5. Under these conditions, 5 nM GF-s was completely cleaved in 5 min, and 50 nM GF-s was completely cleaved in 15 min. The target cleavage efficiency correlates with the efficiency of siRNA incorporation into AGO1 (*see* in Fig. 1)

5. In BYL, uncapped RNA is unstable.
6. FLAG-NtAGO1 synthesized in BYL is also programmed with miRNA/miRNA*.
7. This in vitro system is applied to analyses of translational inhibition by AGO1 or used for analyses of other plant AGO proteins.

References

1. Kawamata T, Tomari Y (2010) Making RISC. Trends Biochem Sci 35(7):368–376. doi:10.1016/j.tibs.2010.03.009
2. Fang X, Qi Y (2016) RNAi in plants: an Argonaute-centered view. Plant Cell 28(2): 272–285. doi:10.1105/tpc.15.00920
3. Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19(3):421–428. doi:10.1016/j.molcel.2005.06.014
4. 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. doi:10.1016/j.molcel.2010.05.014
5. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/j.cell.2008.02.034

6. Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature 443(7114):1008–1012. doi:10.1038/nature05198
7. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol 49(4):493–500. doi:10.1093/pcp/pcn043
8. Iki T (2017) Messages on small RNA duplexes in plants. J Plant Res 130(1):7–16
9. Poulsen C, Vaucheret H, Brodersen P (2013) Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. Plant Cell 25(1):22–37. doi:10.1105/tpc.112.105643
10. Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing-associated small RNAs in plants. Plant Cell 14(7):1605–1619
11. Endo Y, Iwakawa HO, Tomari Y (2013) Arabidopsis ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. EMBO Rep 14(7):652–658. doi:10.1038/embor.2013.73
12. Ye R, Wang W, Iki T, Liu C, Wu Y, Ishikawa M, Zhou X, Qi Y (2012) Cytoplasmic assembly and selective nuclear import of Arabidopsis Argonaute4/siRNA complexes. Mol Cell 46(6):859–870. doi:10.1016/j.molcel.2012.04.013
13. Iki T, Yoshikawa M, Meshi T, Ishikawa M (2012) Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. EMBO J 31(2):267–278. doi:10.1038/emboj.2011.395
14. Smith MR, Willmann MR, Wu G, Berardini TZ, Moller B, Weijers D, Poethig RS (2009) Cyclophilin 40 is required for microRNA activity in Arabidopsis. Proc Natl Acad Sci U S A 106(13):5424–5429. doi:10.1073/pnas.0812729106
15. Iwakawa HO, Tomari Y (2013) Molecular insights into microRNA-mediated translational repression in plants. Mol Cell 52(4):591–601. doi:10.1016/j.molcel.2013.10.033
16. Schuck J, Gursinsky T, Pantaleo V, Burgyan J, Behrens SE (2013) AGO/RISC-mediated antiviral RNA silencing in a plant in vitro system. Nucleic Acids Res 41(9):5090–5103. doi:10.1093/nar/gkt193
17. Yoshikawa M, Iki T, Tsutsui Y, Miyashita K, Poethig RS, Habu Y, Ishikawa M (2013) 3′ fragment of miR173-programmed RISC-cleaved RNA is protected from degradation in a complex with RISC and SGS3. Proc Natl Acad Sci U S A 110(10):4117–4122. doi:10.1073/pnas.1217050110

Chapter 4

In Vitro Analysis of ARGONAUTE-Mediated Target Cleavage and Translational Repression in Plants

Yukihide Tomari and Hiro-oki Iwakawa

Abstract

MicroRNAs (miRNAs) are endogenous small RNAs, which negatively regulate expression of complementary target genes at the post-transcriptional level. In plants, miRNAs are mainly loaded onto ARGONAUTE1 to form RNA-induced silencing complexes (RISCs), which mediate target mRNA cleavage as well as translational repression. The cell-free system derived from tobacco BY-2 protoplasts has become a powerful tool not only for the analysis of RISC assembly mechanism but also for mechanistic dissection of plant RISC functions. Here we describe the detailed protocols for the preparation of BY-2 cell lysate and the procedure to analyze the dual function of plant RISC—target cleavage and translational repression—in vitro.

Key words RNA silencing, ARGONAUTE, RISC, MicroRNA, Translational repression, Target cleavage, Cell-free system, BY-2 protoplasts

1 Introduction

MicroRNAs(miRNAs) are endogenous small RNAs, which regulate expression of complementary target genes at the posttranscriptional level. In *Arabidopsis thaliana* (*Arabidopsis*), more than 300 miRNAs have been registered in the miRNA database (miRBase) so far [1]. The cumulative results indicate that miRNAs act as a master regulator of plant growth, development, and stress responses [2].

miRNAs are produced as their duplex form, consisting of the miRNA and miRNA* strands, by the excision of the double-stranded region of their long precursor hairpin RNAs by an RNase III enzyme DICER-LIKE PROTEIN 1 (DCL1) [3]. The miRNA/miRNA* duplex is methylated by HUA ENHANCER 1 (HEN1) at the 3′ end of each strand [4] and then generally loaded into ARGONAUTE 1 (AGO1), assisted by chaperon machinery [5–8]. Subsequently, the miRNA* strand is generally ejected from the complex as the passenger strand, while the miRNA strand stays in AGO1 as the guide strand, forming the RNA-induced silencing

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_4,

complex (RISC). The RISC containing AGO1 and the single-stranded miRNA interact with complementary target mRNAs, and if the complementarity between the miRNA and the binding site is extensively high, RISC cleaves the targets by the RNase H-like activity of the AGO protein [9, 10]. Because conserved miRNAs in plants often possess a few target mRNAs with extensively complementary sequence, target cleavage was thought as the main mode of miRNA action [11]. In contrast, a number of reports have shown that plant miRNAs reduce protein levels much more than mRNA levels of target mRNAs, suggesting that plant miRNAs can also induce translational repression without target cleavage [12–19]. However, the molecular mechanism of miRNA-mediated translational repression has been poorly understood [20].

In vitro assay systems have served as powerful tools for elucidating molecular mechanisms of RNA silencing [21]. Indeed, assays in cell lysates derived from *Arabidopsis* tissues and cultured cells as well as those in wheat germ extract revealed detailed mechanisms for many important aspects of RNA silencing (e.g., target cleavage, RNA-dependent RNA polymerization, and Dicer activities) [9, 22, 23]. Plant RISC assembly has also been recapitulated in a cell-free system derived from tobacco BY-2 protoplasts [6, 24–26], which was initially developed for in vitro viral replication [27]. In this assay system, synthetic small RNA duplexes can be loaded into in vitro translated AGOs to generate RISCs with a desired combination of an AGO protein and a small RNA guide sequence. Because the BY-2 cell lysate can also recapitulate canonical translation dependent on the 5′ cap and 3′ poly(A) tail [28], we applied this system to investigate the mechanisms of miRNA-mediated translational repression in plants. In order to monitor translational repression without triggering target cleavage, we used a catalytic mutant AtAGO1 and found that *Arabidopsis*AGO1 (AtAGO1)–RISC can inhibit translation of target mRNAs at the elongation and/or initiation steps without promoting deadenylation and mRNA destabilization [10].

In this chapter, we describe the detailed protocols for the preparation of BY-2 cell lysate and the procedure to analyze the dual function of RISC—target cleavage and translational repression—in vitro. Besides understanding the mechanism of translational repression, this system will provide a versatile platform to investigate if and how a miRNA silences a target mRNA of interest, simply by changing sequences of the target mRNA and its cognate miRNA.

2 Materials

2.1 Preparation of BY-2 Cell Lysate

1. BY-2 cell suspension (RIKEN BRC).
2. Mannitol buffer: 0.4 M mannitol, 20 mM magnesium chloride hexahydrate ($MgCl_2$ $6H_2O$). Prepare fresh.

3. Enzyme solution: 2.0% (w/v) cellulase Onozuka RS (Yakult), 0.2% (w/v) pectolyase Y-23 (Wako), 0.4 M mannitol, 20 mM $MgCl_2$ $6H_2O$. Prepare fresh.
4. 10% Percoll solution: 10% (w/v) Percoll (GE Healthcare), 5 mM HEPES–potassium hydroxide (KOH) pH 7.4, 20 mM $MgCl_2$, 0.7 M mannitol. Prepare fresh.
5. 40% Percoll solution: 40% (w/v) Percoll, 5 mM HEPES–KOH (pH 7.4), 20 mM $MgCl_2$, 0.7 M mannitol. Prepare fresh.
6. 70% Percoll solution: 70% (w/v) Percoll, 5 mM HEPES–potassium hydroxide (KOH) (pH 7.4), 20 mM $MgCl_2$, 0.7 M mannitol. Prepare fresh.
7. Large-orifice pipette tips: enlarge the tip of P1000 pipette tips by cutting with a sterile razor blade. Diameter, ~4 mm.
8. Thick-wall centrifuge tube (Beckman; 25 × 89 mm).
9. Syringe 2.5 mL (Terumo), needle 18G × 11/2 (Terumo).
10. Gradient maker (Biocomp; Gradient Master).
11. TR buffer: 30 mM HEPES–KOH (pH 7.4), 100 mM potassium acetate (KOAc), 2 mM magnesium acetate [$Mg(OAc)_2$], 1 × protease inhibitor cocktail (Roche; cOmplete Mini EDTA-free). Prepare fresh.
12. Dounce homogenizer (Wheaton; 7 mL, "tight" pestle).

2.2 Preparation of Target mRNAs and AtAGO1 Expression mRNAs

1. T7-Scribe™ Standard RNA IVT Kit (CELLSCRIPT). Store at −20 °C.
2. Phenol/chloroform/isoamyl alcohol 25:24:1 mixed, pH 5.2 (Nacalai). Store at 4 °C.
3. Chloroform/isoamyl alcohol 24:1: mix 24 vol of chloroform and 1 vol of isoamyl alcohol. Store at 4 °C.
4. Spectrophotometer (e.g., NanoDrop, Thermo Fisher Scientific).
5. ScriptCap™ m^7G Capping System (CELLSCRIPT). Store at −20 °C.
6. A-Plus™ Poly(A) Polymerase Tailing Kit (CELLSCRIPT). Store at −20 °C.

2.3 Preparation of MicroRNA Duplex for RISC Assembly

1. Synthetic miR156 strand (guide strand): 5′-UGACAGAAGAGAGUGAGCAC(M)-3′. "M" indicates the 2′-OMe modification.
2. Synthetic miR156* strand (passenger strand): 5′-GUGCUCUCUCUCUUCUGUCA(M)-3′.
3. 10× T4 polynucleotide kinase buffer: 500 mM Tris–HCl (pH 8.0), 100 mM $MgCl_2$, 50 mM dithiothreitol (DTT).
4. T4 polynucleotide kinase (Takara).

5. 100 mM ATP: dissolve in water. Adjust to pH 7.0–8.0 with KOH. Store in aliquots at −20 °C.
6. 20 mg/mL glycogen (Nacalai): store in aliquots at −20 °C.
7. MicroSpin G-25 column (GE Healthcare).
8. 2× lysis buffer: 60 mM HEPES–KOH (pH 7.4), 200 mM KOAc, and 4 mM $Mg(OAc)_2$. Store at 4 °C.

2.4 In Vitro RISC Function Analysis

1. Substrate mixture (3 mM ATP, 0.4 mM GTP, 100 mM creatine phosphate, 160 mM KOAc, 200 μM each of 20 amino acids, 320 μM spermine, 0.4 U/μL creatine phosphokinase (Calbiochem).
2. TRIzol Reagent (Thermo Fisher Scientific).
3. 4× Laemmli sample buffer: 250 mM Tris–Cl (pH 6.8), 400 mM DTT, 8% sodium dodecyl sulfate (SDS), 0.05% bromophenol blue, 40% glycerol. Store in aliquots at −20 °C.

2.5 Northern Blotting for Detection of Target mRNAs

1. TRIzol Reagent (Thermo Fisher Scientific).
2. Chloroform.
3. Isopropanol.
4. 70% ethanol.
5. 2× formamide dye: 10 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 98% (w/v) deionized formamide, 0.025% (w/v) xylene cyanol, 0.025% bromophenol blue. Store in aliquots at −20 °C.
6. 10× DIG RNA labeling mix (Roche).
7. 10× transcription buffer: 400 mM Tris–HCl (pH 8.0), 80 mM $MgCl_2$, 20 mM spermidine, 50 mM DTT.
8. T7 RNA polymerase (Takara).
9. 3 M sodium acetate (NaOAc) pH 5.2: dissolve in water. Adjust to pH 5.2 with acetic acid. Store in aliquots at room temperature.
10. 20× 3-(N-morpholino)propanesulfonic acid (MOPS) buffer: 400 mM MOPS, 100 mM NaOAc, 25 mM EDTA. Adjust pH 7.0 with sodium hydroxide (NaOH). Store in aliquots at room temperature.
11. Formaldehyde (Nacalai).
12. Gel tray (Gel Tray ex-L, Mupid).
13. Comb [Comb-ex 6.0 mm (W) × 1.0 mm (D) × 13 well, Mupid].
14. Submarine electrophoresis system (Mupid-2 plus, Mupid).
15. Pre-stained RNA marker (DynaMarker® Prestain Marker for RNA High, BioDynamics Laboratory Inc).
16. Downward blotting device (TurboBlotter system, GE Healthcare).

17. Nylon membrane (Hybond-N+, GE Healthcare).
18. 20× SSC: 3 M sodium chloride (NaCl), 0.3 M sodium citrate. Adjust to pH 7.2 with HCl. Store at room temperature.
19. UV crosslinker (FUNA UV crosslinker, Funakoshi).
20. 0.1% methylene blue staining solution: 0.1% (w/v) methylene blue in 0.5 M NaOAc (pH 5.2).
21. Hybridization buffer: dissolve DIG Easy Hyb (GE Healthcare) in 100 mL pre-warmed RNase-free water.
22. Hybridization bottle: diameter 35 mm × length 150 mm.
23. Hybridization oven (HB-80RWR with an exchangeable bottle rotation kit, TAITEC).
24. 2× SSC + 0.1% SDS: 2× SSC containing 0.1% SDS. Store at room temperature.
25. 0.1× SSC + 0.1% SDS: pre-warm up to ~68 °C.
26. 10× maleic acid buffer: 1 M maleic acid, 1.5 M NaCl. Adjust to pH 7.5 with NaOH. Store at room temperature.
27. 1× maleic acid buffer. Store at room temperature.
28. 10× blocking solution: dissolve blocking reagent (Roche) in 1× maleic acid buffer to a final concentration 10% (w/v) with shaking and heating, and autoclave. Store at 4 °C.
29. 1× blocking solution: dilute 10× blocking regent 1:10 in 1× maleic acid buffer by tenfold.
30. 1× wash buffer: 1× maleic acid buffer containing 0.3% (v/v) Tween 20. Store at room temperature.
31. Anti-DIG-AP (Anti-Digoxigenin-AP Fab fragments, Sigma-Aldrich).
32. 1× detection buffer: 0.1 M Tris–HCl, 0.1 M NaCl. Adjust to pH 9.5 with NaOH. Store at room temperature.
33. CDP-Star Detection Reagent (Roche).
34. Thin plastic sheet.
35. LAS-3000 system (Fujifilm).

2.6 Western Blotting for Detection of Target Proteins

1. 5–20% gradient gel (SuperSep 5–20%, Wako).
2. Pre-stained protein marker (Precision Plus Protein™ All Blue Prestained Protein Standards, Bio-Rad).
3. PVDF membrane (Immobilon-P, Millipore).
4. 1× transfer buffer: dilute EzFastBlot (ATTO) by tenfold in protease-free water.
5. Semidry transfer system (Trans-Blot SD Semi-Dry Transfer Cell, Bio-Rad).
6. 10× TBS: 250 mM Tris–HCl (pH 7.4), 1.5 M NaCl.

7. 1× TBST: 25 mM Tris–HCl (pH 7.4), 150 mM NaCl, 0.1% (v/v) Tween.
8. Blocking buffer: 1 × TBST containing 1% (w/v) low-fat milk powder.
9. Anti-FLAG IgG (Anti-FLAG M2 monoclonal antibody, Sigma-Aldrich).
10. HRP-conjugated goat anti-mouse IgG antibody (Thermo Fisher Scientific).
11. Chemiluminescent HRP detection reagent (Luminata Forte Western HRP Substrate, Millipore).

3 Methods

3.1 Preparation of BY-2 Cell Lysate

Preparation of BY-2 cell lysate consists of three steps: (1) preparation of BY-2 protoplasts, (2) separation of vacuoles by centrifugation, and (3) preparation of cell lysate from evacuolated protoplasts. Although the above three steps are important, the healthiness of BY-2 cells is critical for the preparation of translationally active BY-2 cell lysate. If possible, it is highly recommended to obtain several different lines of BY-2 cell suspension for preparation of best cell-free system.

3.1.1 Preparation of BY-2 Protoplasts

1. Transfer 200 mL of 3-day-old BY-2 cell suspension into four 50 mL conical centrifuge tubes.
2. Centrifuge in a swinging bucket rotor at 114 × *g* for 3 min at 25 °C. Remove the supernatant with an aspirator. Resuspend the cell pellet in 40 mL of mannitol buffer for each tube.
3. Centrifuge in a swinging bucket rotor at 114 × *g* for 3 min at 25 °C. Remove the supernatant with an aspirator. Resuspend the cell pellet in 25 mL of enzyme solution for each tube.
4. Combine and transfer the cell suspension (~100 mL in total) into a conical flask (300 mL), and incubate at 25 °C for 3 h with mild shaking in the dark.
5. Confirm the complete digestion of the cell wall by visual inspection of 10 μL aliquots on a glass slide using a dissection microscope.
6. Transfer the protoplast suspension into two 50 mL conical centrifuge tubes.
7. Centrifuge in a swinging bucket rotor at 114 × *g* for 3 min at 4 °C. Remove the supernatant with an aspirator.
8. Resuspend the pellet of protoplast gently in 15 mL of ice-cold mannitol buffer. Combine the protoplast suspension into a 50 mL conical centrifuge tube.

9. Centrifuge in a swinging bucket rotor at 114 × *g* for 3 min at 4 °C. Remove supernatant with an aspirator.
10. Add 40 mL of ice-cold mannitol buffer to the pellet of protoplast, and repeat **steps 9** and **10** until the supernatant is completely clear.
11. Resuspend the pellet of protoplast gently in 24 mL of ice-cold mannitol buffer. Keep it on ice.

3.1.2 Separation of Vacuoles by Centrifugation

To make a cell extract with high translation activity, vacuoles must be removed completely from protoplasts because they store abundant proteases and nucleases inside the membrane. A linear Percoll density gradient allows an efficient separation of vacuoles from BY-2 protoplasts.

1. Put 11 mL of 40% Percoll buffer on the bottom of the thick-wall centrifuge tube (*see* **Note 1**).
2. Overlay 3 mL of 10% Percoll buffer on the top of 40% Percoll buffer with large-orifice pipette tips.
3. Create 10–40% linear gradient with a gradient maker (Gradient Master program: Percoll, 15–40%, v/v).
4. Inject 2 mL of 70% Percoll buffer at the bottom of the tube with syringe and needle without introducing air bubbles or disturbing the gradient.
5. Overlay 4 mL of protoplast suspension on the top of the gradient with large-orifice pipette tips for each tube.
6. Adjust the weight of each tube with mannitol buffer. Centrifuge in a Beckman SW28 swing-out rotor at 11,077 × *g* for 1.5 h at 25 °C.
7. Collect evacuolated protoplasts concentrated at the boundary between 40% and 70% Percoll layer with a pipette, and resuspend them in 20 mL of ice-cold mannitol buffer in a new 50 mL conical centrifuge tube (*see* **Note 2**).
8. Confirm the separation of vacuoles by visual inspection of 10 μL aliquots on a glass slide using a dissection microscope.

3.1.3 Preparation of Cell Lysate from Evacuolated Protoplasts

1. Centrifuge the recovered evacuolated protoplasts in a swinging bucket rotor at 233 × *g* at 4 °C for 3 min. Remove supernatant with an aspirator.
2. Resuspend the pellet of protoplast gently in 40 mL of ice-cold mannitol buffer, and repeat **steps 1** and **2** until the supernatant is completely clear.
3. Remove supernatant completely with an aspirator and a pipette.
4. Weigh the protoplast and add ice-cold TR buffer of four times the amount of the protoplast.

5. Transfer the protoplast suspension into a clean, prechilled Dounce homogenizer. Break the protoplasts with 50 strokes.
6. Transfer the slurry to new microcentrifuge tubes. Centrifuge the slurry at 17,000 × *g* for 10 min at 4 °C.
7. Recover the supernatant into clean, prechilled centrifuge tubes. Aliquot 50 μL of the lysates per microcentrifuge tube, freeze in liquid nitrogen, and store at −80 °C (*see* **Note 3**).

3.2 Preparation of Target mRNAs and AtAGO1 Expression mRNAs

The 5′ and 3′ ends of the in vitro transcripts should be modified with m^7G cap and poly(A) tail, respectively, because both modifications are required for efficient translation in BY-2 cell lysate.

3.2.1 In Vitro Transcription

1. Transcribe reporter RNAs (e.g., FLAG-tagged SPL13 mRNA as a target for miR156) as well as mRNAs carrying AGO1 wild-type and AGO1 catalytic mutant from their template DNAs at 37 °C for 2 h using T7-Scribe Standard RNA IVT Kit (*see* **Note 4**).
2. Add 1 μL of DNase I, and digest the template DNAs for 20 min at 37 °C.
3. Adjust the volume of the reaction to 100 μL with RNase-free water.
4. Purify RNAs with an equal volume of phenol/chloroform/isoamyl alcohol.
5. Recover aqueous phase, and further purify RNAs with an equal volume of chloroform/isoamyl alcohol.
6. Transfer the aqueous phase into a new microcentrifuge tube, add an equal volume of 5 M ammonium acetate (*see***Note 5**), vortex for 10 s, and keep on ice for 15 min.
7. Centrifuge at 15,000 × *g* at 4 °C for 20 min.
8. Remove the supernatant and add 1 mL of 70% ethanol.
9. Centrifuge at 15,000 × *g* at 4 °C for 5 min and remove the supernatant completely.
10. Air-dry for 10 min and dissolve the pellet in 50 μL of RNase-free water.
11. Quantify the concentration of the transcripts with a spectrophotometer.

3.2.2 In Vitro Capping and Polyadenylation

1. Modify the 5′ end of transcript with the m^7G cap structure using ScriptCap m^7G Capping System.
2. Add the capping reaction directly to A-Plus™ Poly(A) Polymerase Reaction for poly(A) tailing of the 3′ end of the capped transcript (*see***Note 6**).

3. Purify RNAs with an equal volume of phenol/chloroform/isoamyl alcohol.
4. Recover aqueous phase, and further purify RNAs with an equal volume of chloroform/isoamyl alcohol.
5. Transfer the aqueous phase into a new microcentrifuge tube, add an equal volume of 5 M ammonium acetate, vortex for 10 s, and keep on ice for 15 min.
6. Centrifuge at 15,000 × *g* at 4 °C for 20 min.
7. Remove the supernatant and add 1 mL of 70% ethanol.
8. Centrifuge at 15,000 × *g* at 4 °C for 5 min and remove the supernatant completely.
9. Air-dry for 10 min, and dissolve the pellet in 50 μL of RNase-free water.
10. Quantify the concentration of the transcripts with a spectrophotometer. If not used immediately, store at −80 °C.

3.3 Preparation of MicroRNA Duplex for RISC Assembly

Because the 5′-monophosphate of the miRNA guide strand is critical for its binding to AGO proteins, synthetic small RNAs should be 5′ phosphorylated by T4 polynucleotide kinase. If you purchase 5′-phosphorylated synthetic small RNAs, you can skip the phosphorylation step.

1. Mix 5 μL of 100 μM guide-strand or passenger-strand RNA, 5 μL of 10× T4 polynucleotide kinase buffer, 5 μL of 10 mM ATP, 5 μL of T4 polynucleotide kinase, and 30 μL of RNase-free water.
2. Incubate the reaction mixture at 37 °C for 2 h.
3. Pass through a MicroSpin G-25 column to remove unincorporated ATP.
4. Mix the flow through with 5 μL of 3 M NaOAc (pH 5.2), 1 μL of 20 mL/mL glycogen, and 150 μL of 100% ethanol. Centrifuge at 20,000 × *g* at 4 °C for 30 min.
5. Rinse the pellet with 1 mL of 70% ethanol. Air-dry the pellet for ~15 min.
6. Dissolve the pellet in 50 μL of RNase-free water.
7. Quantify the concentration of the transcripts and adjust the concentration to 7.5 μM.
8. To make a 1.5 μM stock of microRNA duplex, mix 4 μL of 7.5 μM guide-strand RNA, 6 μL of 7.5 μM passenger-strand RNA, and 10 μL of 2× lysis buffer (*see* **Note 7**).
9. Heat the mixture at 95 °C for 2 min and gradually cool down to room temperature for 30 min. If not used immediately, store at −80 °C.

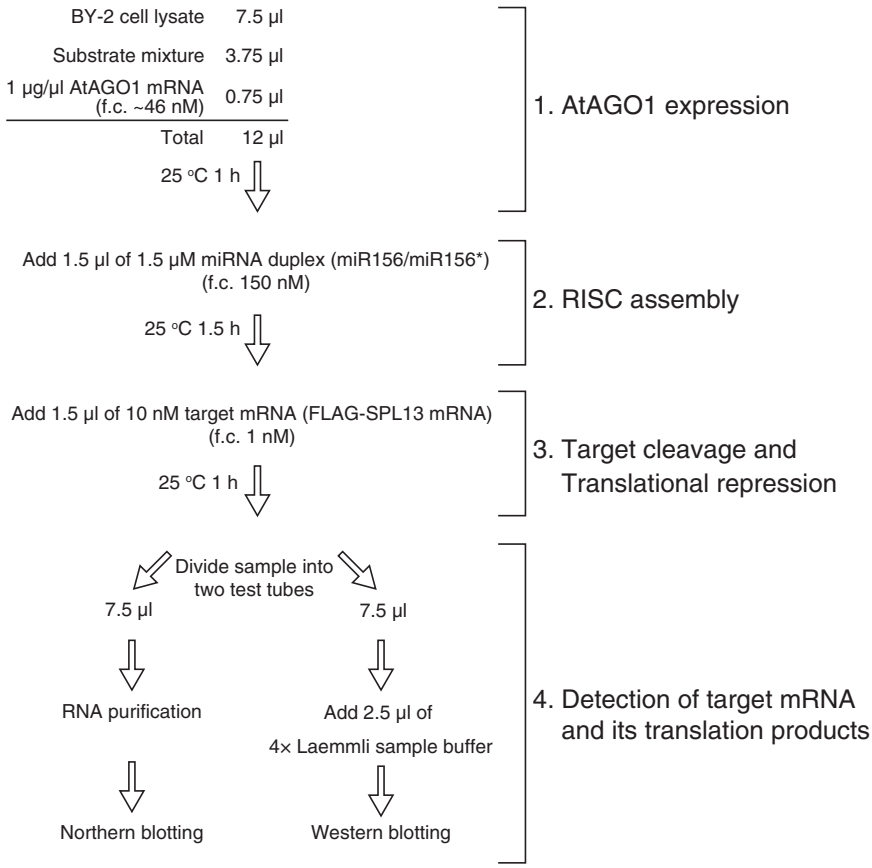


Fig. 1 Flowchart of the procedure for the in vitro analysis of miRNA-mediated target cleavage and translational repression

3.4 In Vitro RISC Function Analysis

In vitro RISC function analysis can be divided into four steps: (1) expression of the AGO protein, (2) RISC assembly, (3) target cleavage and translational repression, and (4) detection of the target mRNA and its translation products. The flowchart of this assay is shown in Fig. 1. To evaluate miRNA-mediated translational repression and overall silencing (cleavage + translational repression), we utilize catalytic mutant AtAGO1–RISC in parallel with wild-type AtAGO1–RISC. As a model pair of a target and its cognate miRNA, we selected the SPL13 mRNA and miR156. miR156 interacts with the SPL13 mRNA in the ORF with one mismatch at the guide position 14 shown in the Fig. 2a. For convenience in detection of SPL13 by Western blotting, we fused a 3× FLAG-tag to the N-terminus of full-length SPL13 (Fig. 2a). This assay clearly showed that AtAGO1–RISC has translational repression activity in addition to target cleavage activity (Fig. 2b, c). Accumulation of translation arrest products suggests that miRNA-mediated translational repression is partly due to blockage of translation elongation by binding of AtAGO1–RISC on the ORF of the target mRNA (Fig. 2b, c).

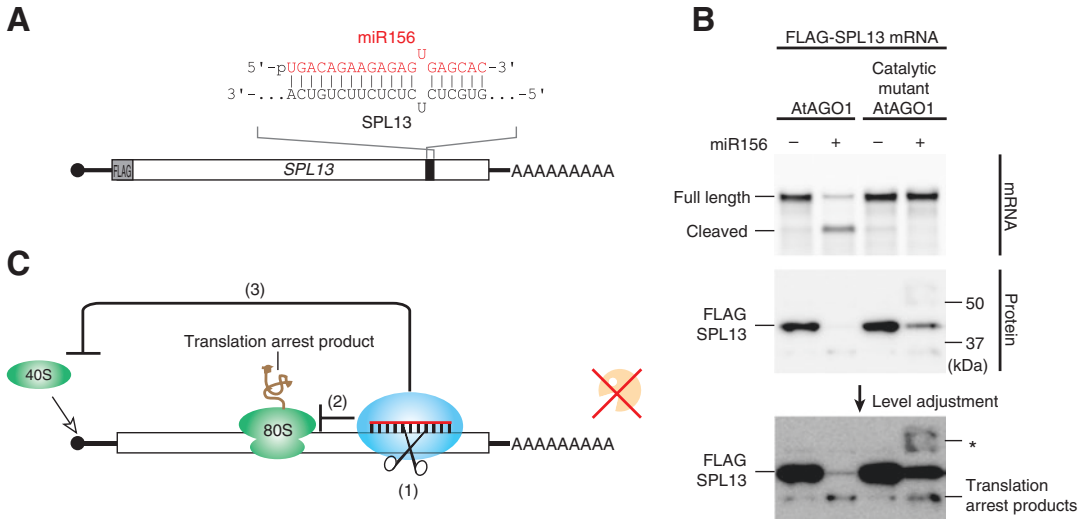


Fig. 2 In vitro miRNA-mediated target cleavage and translational repression. (**a**) Schematic representation of the N-terminal FLAG-tagged SPL13 mRNA and the base-pairing configuration between the target sequence and miR156. (**b**) Northern blotting of the FLAG–SPL13 mRNA (*top*) and anti-FLAG Western blotting of its product (*middle* and *bottom*), in the presence or absence of wild-type or catalytic mutant AtAGO1 programmed with miR156. In the presence of miR156-programmed wild-type AtAGO1–RISC, the full-length FLAG–SPL13 mRNA was efficiently cleaved, and the full-length SPL13 protein was also drastically decreased. In the presence of miR156-programmed catalytic mutant AtAGO1, the mRNA level was not changed, but the protein level was still decreased, indicating translational repression. Translation arrest products were detected in the presence of wild-type and catalytic mutant AtAGO1 programmed with miR156 (*bottom*). The *asterisk* indicates the position of the presumed elongation-arrested peptidyl–tRNAs from FLAG–SPL13 in the presence of catalytic mutant AtAGO1 programmed with miR156 (*bottom*). Even in the absence of miR156, faint bands corresponding to the cleaved FLAG–SPL13 mRNAs (*top*) and the translation arrest products (*middle* and *bottom*) were detected. This is presumably due to the activity of the endogenous tobacco miR156–RISC present in the BY-2 cell lysate. (**c**) A model for the activities of AtAGO1–RISC. AtAGO1–RISC not only (*1*) cleaves the target mRNA but also blocks its translation at (*2*) elongation and/or (*3*) initiation step. In contrast to animal RISCs, AtAGO1–RISC does not induce deadenylation or subsequent mRNA decay

1. For a 15 μL reaction, mix 7.5 μL of BY-2 cell lysate, 3.75 μL of substrate mixture, and 0.75 μL of 1 μg/μL wild-type AtAGO1 mRNA or catalytic mutant AtAGO1 mRNA [final concentration (f.c.), ~46 nM] in a microcentrifuge tube. Incubate at 25 °C for 1 h.
2. Add 1.5 μL of 1.5 μM miRNA duplex (f.c. 150 nM) for programing RISC or 1.5 μL of 1 × lysis buffer as a negative control. Incubate at 25 °C for 90 min.
3. Add 1.5 μL of 10 nM target mRNA (FLAG–SPL13 mRNA) (f.c. 1 nM). Incubate at 25 °C for 1 h (*see* **Note 8**).
4. Transfer 7.5 μL of the reaction mixture into a new microcentrifuge tube, and add 300 μL of TRIzol to purify the total RNAs for the detection of target mRNAs by Northern blotting.

Add 2.5 μL of 4× Laemmli sample buffer to the remaining 7.5 μL of reaction mixture for the detection of target products by Western blotting.

3.5 Northern Blotting for the Detection of Target mRNAs

3.5.1 Extraction of Total RNA

1. Shake the tube containing the mixture of 7.5 μL sample and 300 μL TRIzol for 10 s, and incubate for 5 min at 25 °C.
2. Flash spin and add 60 μL of chloroform.
3. Shake tubes and incubate at 25 °C for 2 min.
4. Centrifuge the samples at 10,000 × *g* for 10 min at 4 °C.
5. Recover 180 μL of the upper aqueous phase into a new microcentrifuge tube.
6. Precipitate the total RNA in the aqueous phase by mixing with 150 μL of isopropanol. Incubate at 25 °C for 10 min.
7. Centrifuge at 10,000 × *g* for 10 min at 4 °C. Remove the supernatant.
8. Rinse the RNA pellet with 1 mL of 70% ethanol. Centrifuge at 7000 × *g* for 5 min at 4 °C. Remove the supernatant completely and air-dry for 10 min.
9. Dissolve the pellet in 10 μL of 2× formamide dye. If not used immediately, store at −80 °C.

3.5.2 Preparation of Digoxigenin-Labeled RNA Probe

We normally use a digoxigenin (DIG)-labeled Riboprobe containing the complementary sequence of the ORF of the target RNA for the detection by Northern blotting.

1. Mix 1 μL of 1 μg/μL PCR product containing T7 promoter and the complementary sequence of the full-length target RNA, 2 μL of 10× DIG RNA labeling mix, 2 μL of 10× transcription buffer, 13 μL of RNase-free water, and 2 μL of T7 RNA polymerase. Incubate at 37 °C for 2 h.
2. Add 80 μL of RNase-free water, 1 μL of 20 mg/mL glycogen, 10 μL of 3 M NaOAc (pH 5.5), and 300 μL of 100% ethanol. Vortex, and centrifuge at 20,000 × *g* at 4 °C for 15 min.
3. Remove the supernatant. Rinse the pellet with 1 mL of 70% ethanol. Centrifuge at 20,000 × *g* at 4 °C for 5 min and air-dry for ~10 min.
4. Dissolve the pellet in 100 μL of RNase-free water.
5. Aliquot 10 μL into new microcentrifuge tubes, and store at −80 °C.

3.5.3 Northern Blotting

1. Mix 67.5 mL of RNase-free water and 0.71 g agarose in a 300 mL conical flask, heat the mixture to boiling in a microwave oven, and cool to ~60 °C.
2. In a fume hood, gently mix the molten agarose with 3.75 mL of 20× MOPS buffer and 3.75 mL of 37% formaldehyde.

3. Pour the gel solution into a gel tray, immediately insert a comb, and allow to solidify in the fume hood.
4. Set the gel in the submarine electrophoresis system, and fill the buffer tank with 1× MOPS buffer.
5. Heat the samples at 95 °C for 3 min, and then load 5 μL of aliquots into the well.
6. After loading samples, load 2 μL of pre-stained RNA marker at the both sides of the sample lanes. Perform electrophoresis at 100 V until the 200 nt of pre-stained RNA marker reaches ~10 mm away from the bottom of the gel.
7. Transfer RNAs from the agarose gel to a nylon membrane with a downward blotting device overnight with 20× SSC transfer buffer.
8. After transfer, UV cross-link the membrane with 1200 × 100 μJ/cm^2, immediately soak the membrane in RNase-free water for 30 s, remove the water, and then add 0.1% methylene blue staining solution to check if the amounts of loaded total RNAs are equal among the samples.
9. After the detection of total RNA, wash the membrane four times with 40 mL of RNase-free water, and then dry the membrane completely.
10. Insert the membrane into a hybridization bottle, and pour 10 mL of pre-warmed (~68 °C) hybridization buffer. After closing the cap, incubate at 68 °C for 10 min at a constant rotation speed (25 rpm) in a hybridization oven.
11. Dilute 5 μL of DIG-labeled probe with 500 μL of hybridization buffer. Add the total amount of the diluted probe solution into the bottom of the hybridization bottle.
12. After closing cap, incubate at 68 °C for ~16 h at a constant rotation speed (25 rpm) in a hybridization oven.
13. Remove hybridization buffer. Wash the membrane twice with 50 mL of 2× SSC + 0.1% SDS for 5 min at 25 °C.
14. Further wash the membrane three times with 50 mL of 0.1× SSC + 0.1% SDS at 68 °C for 15 min at a constant rotation speed (25 rpm) in a hybridization oven.
15. Take out the membrane from the bottle with a flat head tweezers, and rinse the membrane with 1× maleic acid buffer for 1 min.
16. Incubate the membrane in 1× blocking solution for 30 min at 25 °C with rocking.
17. After blocking, incubate the membrane with anti-DIG-AP (1:10,000) in 1× blocking solution for 30 min at 25 °C with rocking.

18. Wash the membrane twice, 15 min each, in 1× wash buffer at 25 °C with rocking.
19. Incubate the membrane in 1× detection buffer for 5 min at 25 °C.
20. Place the membrane on a thin plastic sheet, and pour the CDP-Star Detection Reagent on the membrane. Place a thin plastic sheet at the top of the membrane.
21. Visualize the signals with LAS-3000 system or equivalent.

3.6 Western Blotting for the Detection of Target Proteins

1. Load 5 μL of samples on the well of the 5–20% gradient gel. After loading the samples, load 2 μL of pre-stained marker at the both sides of the sample lanes.
2. Perform electrophoresis at 200 V until the bromophenol blue dye reaches ~5 mm away from the bottom of the gel.
3. Transfer the proteins on the PVDF membrane with a semidry blotting apparatus with 1× transfer buffer at 25 V for 30 min.
4. Incubate the membrane with blocking buffer for 10 min at 25 °C.
5. Incubate the membrane with mouse anti-FLAG IgG antibody (1:3000) in blocking buffer overnight at 25 °C with rocking.
6. Wash the membrane twice, 5 min each, in 1× TBST at 25 °C with rocking.
7. Incubate the membrane with HRP-conjugated goat anti-mouse IgG antibody (1:1000) in blocking buffer at 25 °C for 1 h with rocking.
8. Wash the membrane three times, 15 min each, in 1× TBST at 25 °C with rocking.
9. Place the membrane on a thin plastic sheet, and pour a chemiluminescent HRP detection reagent on the membrane. Place a thin plastic sheet on the top of the membrane.
10. Visualize the signals with LAS-3000 system or equivalent.

4 Notes

1. We usually prepare six density gradient tubes per 200 mL of BY-2 cell suspension.
2. Weigh the empty conical tube for accurate measurement of the weight of evacuolated protoplasts (*see* **step 4** in Subheading 3.1.3).
3. The lysate retains enough activity for translation and RISC assembly at least for a year at −80 °C.
4. Both PCR products and linearized plasmids containing T7, T3, or SP6 promoter can be used for the preparation of mRNAs

for in vitro translation. Indeed, we transcribe the FLAG–SPL13 mRNA, which is used as a model target mRNA, from a PCR product but AtAGO1 expression mRNAs, carrying wild-type or catalytic mutant AtAGO1, from *Not*I-linearized plasmids.

5. No alcohol is required.
6. The length of the poly(A) tail added to the 3′ end of transcripts can be varied by changing the quantity of the substrate RNA, incubation time, units of poly(A) polymerase, and reaction volume. We usually mix 100 μL of ScriptCap m7G Capping reaction with 60 μg of substrate transcript, 13.2 μL of 10× A-Plus Poly(A) Tailing Buffer, 13.2 μL of 10 mM ATP, and 5 μL of A-Plus Poly(A) Polymerase (20 U), and incubate for 30 min at 37 °C. This step can be skipped when poly(A) sequence has been already included in the transcription template. Indeed, the plasmids for in vitro transcription of AtAGO1 expression mRNAs contain ~60 nt of poly(A) sequence just upstream of the *Not*I site [10].
7. To ensure that all the guide-strand RNAs form duplexes, we usually use a 1.5-fold excess of the passenger-strand RNAs.
8. miRNA-mediated translational repression is sensitive to the concentration of target mRNAs. Assay should be performed using less than 2.5 nM target mRNA at final concentration.

Acknowledgments

We are grateful to the members of Tomari laboratory for comments on this manuscript. This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas ("Nascent-chain Biology") 26116003 (to H.I.) and ("Non-coding RNA neo-taxonomy") 26113007 (to Y.T.), Grant-in-Aid for Young Scientists (A) 16H06159 (to H.I.), and Grant-in-Aid for Challenging Exploratory Research 15K14444 (to H.I.).

References

1. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42(Database issue):D68–D73. doi:10.1093/nar/gkt1181
2. Sunkar R, Li YF, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. Trends Plant Sci 17(4):196–203. doi:10.1016/j.tplants.2012.01.010
3. Kurihara Y, Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through dicer-like 1 protein functions. Proc Natl Acad Sci U S A 101(34):12753–12758. doi:10.1073/pnas.0403115101
4. Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. Science 307(5711):932–935. doi:10.1126/science.1107130
5. Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short

interfering RNAs. Proc Natl Acad Sci U S A 102(33):11928–11933. doi:10.1073/pnas.0505461102

6. 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. doi:10.1016/j.molcel.2010.05.014
7. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/j.cell.2008.02.034
8. Vaucheret H, Vazquez F, Crete P, Bartel DP (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev 18(10):1187–1197. doi:10.1101/gad.1201404
9. Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. Genes Dev 17(1):49–63. doi:10.1101/gad.1048103
10. Iwakawa HO, Tomari Y (2013) Molecular insights into microRNA-mediated translational repression in plants. Mol Cell 52(4):591–601. doi:10.1016/j.molcel.2013.10.033
11. Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53. doi:10.1146/annurev.arplant.57.032905.105218
12. Todesco M, Rubio-Somoza I, Paz-Ares J, Weigel D (2010) A collection of target mimics for comprehensive analysis of microRNA function in *Arabidopsis thaliana*. PLoS Genet 6(7):e1001031. doi:10.1371/journal.pgen.1001031
13. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8(4):517–527. doi:10.1016/j.devcel.2005.01.018
14. Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science 303(5666):2022–2025. doi:10.1126/science.1088060
15. Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 15(11):2730–2741. doi:10.1105/tpc.016238
16. Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J 49(4):683–693. doi:10.1111/j.1365-313X.2006.02983.x
17. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190. doi:10.1126/science.1159151
18. Yang L, Wu G, Poethig RS (2012) Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. Proc Natl Acad Sci U S A 109(1):315–320. doi:10.1073/pnas.1114673109
19. Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, Zhang F, Raikhel N, Jiang L, Chen X (2013) MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. Cell 153(3):562–574. doi:10.1016/j.cell.2013.04.005
20. Iwakawa HO, Tomari Y (2015) The functions of microRNAs: mRNA decay and translational repression. Trends Cell Biol 25(11):651–665. doi:10.1016/j.tcb.2015.07.011
21. Liu Q, Paroo Z (2010) Biochemical principles of small RNA pathways. Annu Rev Biochem 79:295–319. doi:10.1146/annurev.biochem.052208.151733
22. Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19(3):421–428. doi:10.1016/j.molcel.2005.06.014
23. Fukudome A, Kanaya A, Egami M, Nakazawa Y, Hiraguri A, Moriyama H, Fukuhara T (2011) Specific requirement of DRB4, a dsRNA-binding protein, for the in vitro dsRNA-cleaving activity of Arabidopsis dicer-like 4. RNA 17(4):750–760. doi:10.1261/rna.2455411
24. Iki T, Yoshikawa M, Meshi T, Ishikawa M (2012) Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. EMBO J 31(2):267–278. doi:10.1038/emboj.2011.395
25. Ye R, Wang W, Iki T, Liu C, Wu Y, Ishikawa M, Zhou X, Qi Y (2012) Cytoplasmic assembly and selective nuclear import of Arabidopsis Argonaute4/siRNA complexes. Mol Cell 46(6):859–870. doi:10.1016/j.molcel.2012.04.013
26. Endo Y, Iwakawa HO, Tomari Y (2013) Arabidopsis ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. EMBO Rep 14(7):652–658. doi:10.1038/embor.2013.73

27. Komoda K, Naito S, Ishikawa M (2004) Replication of plant RNA virus genomes in a cell-free extract of evacuolated plant protoplasts. Proc Natl Acad Sci U S A 101(7):1863–1867. doi:10.1073/pnas.0307131101

28. Iwakawa HO, Mizumoto H, Nagano H, Imoto Y, Takigawa K, Sarawaneeyaruk S, Kaido M, Mise K, Okuno T (2008) A viral noncoding RNA generated by cis-element-mediated protection against 5′->3′ RNA decay represses both cap-independent and cap-dependent translation. J Virol 82(20):10162–10174. doi:10.1128/JVI.01027-08

Chapter 5

Establishment of an In Vivo ARGONAUTE Reporter System in Plants

Károly Fátyol and József Burgyán

Abstract

RNA silencing is not only an evolutionarily conserved gene regulatory mechanism, but in plants also serves as the basis for robust adaptive antiviral immune responses. ARGONAUTE (AGO) proteins form the catalytic cores of the RNA-guided ribonuclease complexes, which play a central role in RNA silencing. Here we describe an in vivo assay system for analyzing the activities of AGO proteins in the virological model plant *Nicotiana benthamiana*.

Key words ARGONAUTE, RNA silencing, Agroinfiltration, Transient reporter system, *Nicotiana benthamiana*

1 Introduction

Plants respond to viral infection by a defensive mechanism, which relies on recognition and subsequent inactivation of the invading foreign nucleic acids. Years of research have uncovered striking similarities between this process and RNA silencing, an important cellular phenomenon responsible for regulating gene expression at the transcriptional and posttranscriptional levels [1–5]. Besides the mechanistic resemblance, many of the molecular players employed by the two processes are also shared. For example, the protein complexes (RNA-induced silencing complexes, RISCs) that are executing the final steps of both RNA silencing and viral restriction contain members of the highly conserved ARGONAUTE (AGO) protein family. One of the most important biochemical properties of AGO proteins is their ability to bind small RNAs, which ensures sequence specific recognition of the substrate nucleic acids by RISCs. Once recognized, the RISCs can trigger inactivation of the targeted RNA or DNA molecules through various means.

Nicotiana benthamiana is a widely used experimental host in plant biology, due mainly to the large number of viruses that can

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_5,

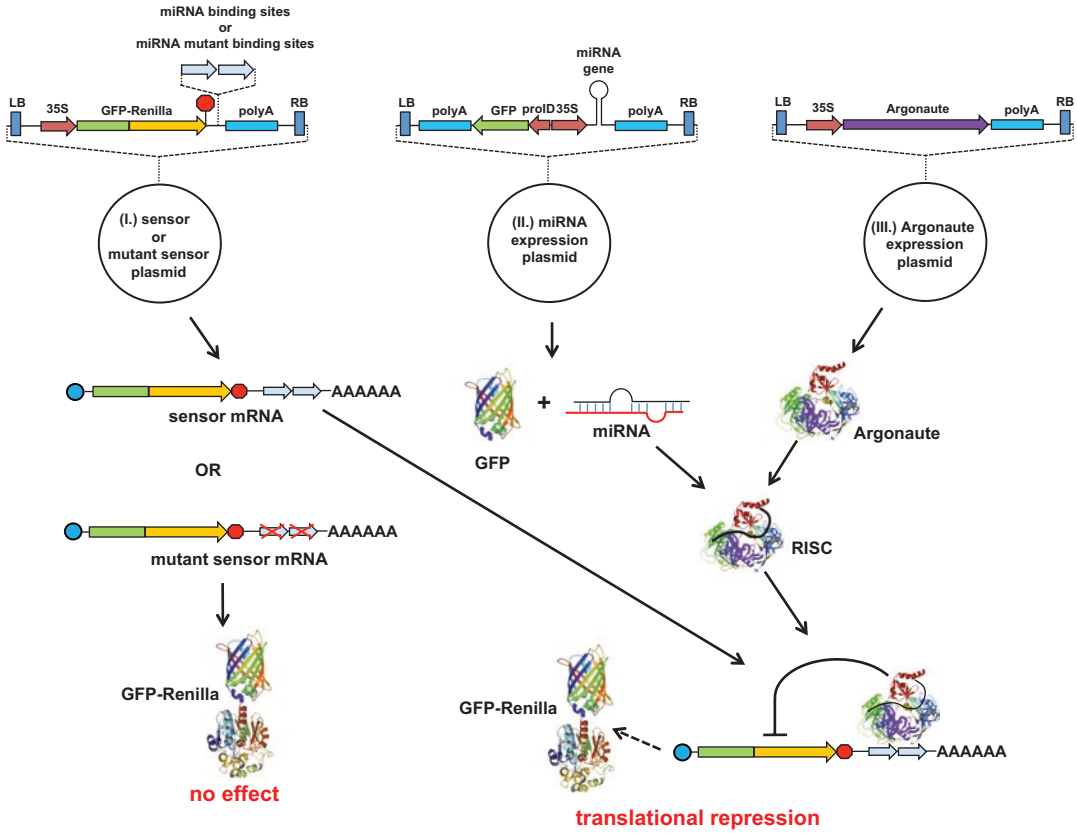


Fig. 1 Agroinfiltration-based transient gene silencing system. Schematic structures of components of the transient gene silencing system are shown. See details in the text

successfully infect it. Additionally, it is also susceptible to a wide variety of other plant pathogens (such as bacteria, oomycetes, fungi), making this species a cornerstone of host–pathogen research [6]. Recently, the draft genome sequence and various transcriptome data sets of *N. benthamiana* have become available from several sources [7–9]. Using this information, we have cloned the full-length cDNAs of the nine *N. benthamiana* AGO proteins [10]. We have developed an in vivo reporter assay to monitor the activity of the cloned proteins.

Our system is based on the *Agrobacterium tumefaciens*-mediated transient expression assay developed previously to analyze microRNA (miRNA)-dependent gene silencing [11–13]. The three components of the system are the following: (1) sensor transcript, (2) miRNA expression cassette, and (3) AGO expression cassette (Fig. 1). These elements are incorporated into binary plasmid vectors to allow their efficient co-delivery into tobacco leaves by agroinfiltration. The sensor plasmid encodes a GFP–*Renilla* luciferase fusion protein under the control of a constitutively active CaMV 35S promoter. MiRNA target sites are inserted into various

positions of the sensor mRNA (3′UTR or ORF). The second component of the reporter system encodes a miRNA. This plasmid also constitutively produces GFP protein, which is used as an internal control to normalize for the varying efficiencies of agroinfiltration. The AGO protein under study is expressed from the third binary plasmid. The effector complexes, assembled from the co-expressed AGO proteins and miRNAs, can interact with the sensor mRNAs via the miRNA target sites. Monitoring the expression of the sensor mRNA allows the assessment of gene silencing. By using this system, we have previously characterized the activities of several *N. benthamiana* AGO proteins and also successfully identified functionally important regions within the antiviral AGO2 molecule [10]. This chapter provides the reader with an in-depth protocol for the use of the above reporter system. First, we provide guidelines on the construction of the three components of the system (AGO expression vector, miRNA expression vector, sensor). Next, we give a protocol for the delivery on the above components into plants. Finally, the quantitative Western blot analysis of protein lysates is described.

2 Materials

Solutions should be prepared in autoclaved deionized water unless otherwise indicated. Producers of only specialty reagents, kits, and enzymes are indicated. General reagents and chemicals can be purchased from various suppliers and should be of molecular biology or analytical grade.

2.1 Vector Construction

1. Oligonucleotides (IDT, Sigma–Aldrich, other suppliers).
2. RNeasy Plant Mini Kit (Qiagen).
3. DNeasy Plant Mini Kit (Qiagen).
4. Titan One Tube RT-PCR kit (Roche).
5. FDE: 10 mL deionized formamide, 200 μL ethylenediaminetetraacetic acid (EDTA) pH 8, 10 mg/mL xylene cyanol, 10 mg/mL bromophenol blue.
6. pGEM-T easy vector system (Promega).
7. Restriction and modifying enzymes (Fermentas, NEB).
8. Agarose.
9. Ethidium bromide.
10. DNA molecular weight markers, 6× DNA loading dye (Fermentas).
11. TBE buffer: 89 mM Tris–HCl, 89 mM boric acid, 2 mM EDTA pH 8.
12. Luria–Bertani (LB) medium: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride (NaCl).

13. LB agar plates [1.5% (w/v) agar] with appropriate antibiotics (100 μg/mL ampicillin, 100 μg/mL spectinomycin, 50 μg/mL kanamycin).
14. Plasmid purification and fragment isolation kits (Qiagen, Macherey–Nagel, GE Healthcare, Thermo Fisher Scientific, or other suppliers).
15. Low- and high-fidelity PCR enzymes: Taq polymerase (Fermentas), Phusion polymerase (Thermo Fisher Scientific).
16. 100 mM dATP/dGTP/dCTP/dTTP (Fermentas). 10 mM dNTP mix is prepared form the 100 mM deoxynucleotide triphosphate stocks.
17. LR Clonase II Plus enzyme (Thermo Fisher Scientific).
18. Gateway entry vectors: pENTR11, pENTR2B, pENTR4 (Thermo Fisher Scientific).
19. Binary expression vectors: pBIN61, pK7WG2D, pK7WGF2 (*see* **Note 1**).
20. *E. coli* strains: TOP10, DB3.1 (*see* **Note 2**).
21. Equipment: gel electrophoresis apparatus, water bath, shaker, incubator, centrifuge, Nanodrop, thermocycler.

2.2 Agroinfiltration

1. *A. tumefaciens* strains: C58C1 or GV3101.

2.2.1 Agrobacterium Transformation/ Growth

2. 20 mM calcium chloride ($CaCl_2$).
3. YEB medium: 5 g/L tryptone, 1 g/L yeast extract, 5 g/L nutrient broth, 5 g/L sucrose 0.49 g/L magnesium sulfate heptahydrate ($MgSO_4 \cdot 7H_2O$).
4. RT plates: YEB agar plates [1.5% (w/v) agar], 20 μg/mL rifampicin, 5 μg/mL tetracycline.
5. RTK plates: YEB agar plates [1.5% (w/v) agar], 20 μg/mL rifampicin, 5 μg/mL tetracycline, 50 μg/mL kanamycin.
6. RTS plates: YEB agar plates [1.5% (w/v) agar], 20 μg/mL rifampicin, 5 μg/mL tetracycline, 100 μg/mL spectinomycin.
7. Liquid nitrogen.
8. Equipment: water bath, shaker, incubator, centrifuge.

2.2.2 Agroinfiltration

1. YEB medium: 5 g/L beef extract, 1 g/L yeast extract, 5 g/L peptone, 5 g/L sucrose, 0.5 g/L magnesium chloride ($MgCl_2$).
2. 1 M acetosyringone dissolved in dimethyl sulfoxide (DMSO).
3. 1 M 2-(N-morpholino) ethanesulfonic acid (MES) (pH 5.7).
4. 1 M $MgCl_2$.
5. Infiltration medium: 10 mM $MgCl_2$, 250 μM acetosyringone (prepare freshly each time from stock solutions).
6. 1 mL hypodermic syringe.

2.3 Western Blotting

1. Liquid nitrogen.

2.3.1 Protein Lysate Preparation

2. 100 mM phenylmethylsulfonyl fluoride (PMSF) dissolved in isopropanol.
3. Complete protease inhibitor (Roche).
4. 200 mM sodium orthovanadate (Na_3VO_4) (*see* **Note 3**).
5. 1 M dithiothreitol (DTT).
6. Lysis buffer: 10 mM Tris–HCl pH 7.6, 1 mM EDTA pH 8, 150 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v) Nonidet P-40, 5 mM sodium fluoride (NaF), 1 mM DTT, 0.5 mM Na_3VO_4, 1 mM PMSF, complete protease inhibitor (the last four components are added to the lysis buffer just before use).
7. Equipment: porcelain mortar and pestle, centrifuge.

2.3.2 Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. Resolving gel buffer: 1.5 M Tris–HCl pH 8.8, 0.4% (w/v) SDS.
2. Stacking gel buffer: 0.5 M Tris–HCl pH 6.8, 0.4% (w/v) SDS.
3. 10% (w/v) ammonium persulfate (APS) (prepared fresh each time).
4. *N*,*N*,N,*N*′-tetramethylethylenediamine (TEMED).
5. 30% acrylamide/bis-acrylamide solution, 37.5:1 (Bio-Rad).
6. Resolving gel solution—8% acrylamide (15 mL): 4 mL 30% acrylamide/bis-acrylamide solution, 3.75 mL resolving gel buffer, 7.25 mL water, 50 µL 10% APS, 10 µL TEMED.
7. Stacking gel solution (5 mL): 650 µL 30% acrylamide/bis-acrylamide solution, 1.25 mL stacking gel buffer, 3.05 mL water, 25 µL 10% APS, 5 µL TEMED.
8. SDS-PAGE running buffer: 0.025 M Tris, 0.192 M glycine, 0.1% (w/v) SDS.
9. 6× SDS sample buffer: 0.35 M Tris–HCL pH 6.8, 30% (v/v) glycerol, 10% (w/v) SDS, 6 M DTT, 0.012% (w/v) bromophenol blue.
10. Pre-stained protein molecular weight marker (Lonza).
11. Equipment: Mini-PROTEAN electrophoresis chamber (Bio-Rad).

2.3.3 Electro-Transfer

1. Towbin buffer: 0.025 M Tris, 0.192 M glycine, 20% (v/v) methanol.
2. Equipment: Mini Trans-Blot Cell electroblotting apparatus (Bio-Rad) or equivalent transfer chamber.
3. Nitrocellulose membrane, 0.45 µm pore size (Bio-Rad).

2.3.4 Immunological Detection

1. Ponceau staining solution: 0.1% (w/v) Ponceau S, 5% (v/v) acetic acid.
2. Tris-buffered saline (TBS): 50 mM Tris–HCL pH 7.6, 150 mM NaCl.

3. Nonfat dry milk powder.
4. TBS-T: TBS supplemented with 0.1% (v/v) TWEEN 20.
5. Blocking solution: TBS-T supplemented with 5% (w/v) nonfat dry milk powder.
6. Polyclonal anti-GFP antibody (Sigma–Aldrich).
7. Horseradish peroxidase (HRP)-conjugated rat monoclonal anti-HA (3F10) antibody (Roche).
8. HRP-conjugated goat anti-rabbit IgG (whole molecule) (Sigma–Aldrich).
9. Clarity Western ECL Substrate (Bio-Rad).
10. Equipment: orbital shaker, ChemiDoc XRS Plus System (Bio-Rad).
11. Software: Image Lab software (Bio-Rad).

3 Methods

Use kits according to manufacturer's instructions unless otherwise indicated.

3.1 Generation of AGO Expression Constructs

High-quality transcriptome assemblies have recently become available from *N. benthamiana* [7, 8]. Using these data, appropriate forward and reverse primers can be designed to isolate full-length cDNAs of AGOs. Given that AGOs are large proteins (>100 kDa, the encoding ORFs are ~3 kb), it is more sensible to amplify overlapping subfragments of the genes, instead of attempting the amplification of the full-length ORFs in one step. The obtained subfragments can subsequently be assembled into full-length ORFs using conventional cloning techniques or overlapping PCR. Currently, there are no good quality antibodies available for the detection of *N. benthamiana* AGOs; thus, epitope tagging is necessary if one wishes to monitor the expression of the proteins. N terminal tagging of AGOs has been found to be compatible with the proteins' functions. FLAG–HA epitope tag can be attached to the N terminus of the assembled full-length ORFs using double-stranded oligonucleotides. Finally, the cDNAs can be inserted into binary plasmid vectors, which allows expression of the tagged AGOs in plants.

3.1.1 Isolation and Cloning of Full-Length N. benthamiana AGO cDNAs

1. Isolate total RNA from young *N. benthamiana* leaves using the RNeasy Plant Mini Kit.
2. Quantify RNA with Nanodrop spectrophotometer.
3. Check the quality of the isolated RNA by running an aliquot (0.5–1 μg) on a 1.2% (w/v) agarose TBE gel (*see* **Note 4**). Good quality RNA should appear as a ladder of sharp bands (ribosomal RNAs).

4. Set up a combined cDNA synthesis/PCR reaction using the Titan One Tube RT-PCR kit according to manufacturer instructions.
5. Check an aliquot (1/10th) of the RT–PCR on a 1.2% agarose TBE gel. Use the rest of the sample to purify the amplified PCR fragment by a fragment isolation kit.
6. Ligate the purified PCR fragment into pGEM-T easy plasmid vector according to manufacturer instructions.
7. Transform ligation reaction into chemically competent TOP10 *E. coli* (*see* **Note 5**).
8. Spread bacteria onto LB agar plates containing 100 μg/mL of ampicillin. Incubate the plates overnight at 37 °C.
9. Screen the appearing colonies by colony PCR (*see* **Note 6**).
10. Purify plasmids from colonies carrying inserts of the expected size and verify identity of the fragments by sequencing.
11. Assemble full-length AGO ORFs from cloned subfragments using conventional cloning techniques or overlapping PCR (*see* **Note 7**).

3.1.2 Epitope Tagging of N. benthamiana AGOs

1. Mix 1 nmol of each of the complementary FLAG–HA oligonucleotides in 100 μL of 1× Tango buffer.
2. Heat the mixture to 100 °C for 5 min, then turn off the heating block, and allow it to cool down slowly to room temperature (usually overnight).
3. Digest 1 μg of full-length AGO ORF containing pGEM-T easy plasmid vector with *Pme*I–*Sal*I (*see* **Note 8**). Gel purify the digested plasmid.
4. Mix 100 ng of *Pme*I–*Sal*I digested AGO vector with 0.1 pmol of double-stranded FLAG–HA oligonucleotide. Set up a 10 μL ligation reaction by adding 1 μL of 10 × T4 DNA ligase buffer, appropriate amount of water, and 1 μL of T4 DNA ligase to the above mixture. Incubate the reaction at room temperature for 2–4 h.
5. Transform ligation reaction into chemically competent TOP10 *E. coli* and grow bacteria as above.
6. Identify colonies carrying the epitope tag in the proper orientation by colony PCR.
7. Confirm correct insertion of the tag-encoding oligonucleotide by sequencing.

3.1.3 Insertion of N. benthamiana AGOs into Binary Expression Vector

1. Digest 2 μg of the pGEM–FLAG–HA–AGO plasmid with appropriate restriction enzymes (*see* **Note 9**).
2. Resolve the digested plasmid on a 1.2% agarose gel and isolate the AGO ORF containing restriction fragment.

3. Mix threefold molar excess of the isolated AGO containing restriction fragment with pBIN61 plasmid linearized with the appropriate restriction enzymes, and set up ligation in 10 μL volume as above. Incubate reaction at room temperature for 2–4 h.
4. Transform ligation into competent TOP10 *E. coli*. Spread bacteria onto LB agar plates containing 50 μg/mL of kanamycin, and incubate plates overnight at 37 °C.
5. Screen the appearing colonies by colony PCR.

3.2 Generation of miRNA Expression Constructs

The draft genome sequence of *N. benthamiana* is available from several sources. Based on these data, forward and reverse primers can be designed to isolate conserved miRNA genes of *N. benthamiana*. In order to ensure efficient processing of the cloned miRNA, in addition to the region encoding the miRNA hairpin, sufficiently long flanking regions should be included in the amplified DNA fragment (at least 200–300 nt in both 5′ and 3′ directions).

3.2.1 Isolation and Cloning of miRNA Genes of N. benthamiana

1. Isolate plant genomic DNA from young *N. benthamiana* leaves using DNeasy Plant Mini Kit.
2. Quantify genomic DNA with Nanodrop.
3. Set up PCR reaction in 50 μL volume by mixing 100 ng of genomic DNA, 10 μL of 5× Phusion buffer, 1 μL of 10 μM forward primer, 1 μL of 10 μM reverse primer, 1 μL of 10 mM dNTP mix, appropriate amount of water, and finally 0.5 μL of Phusion polymerase. Cycling conditions are the following: 98 °C for 2 min (initial denaturation); 35 cycles of 98 °C for 10 s (denaturation), T_m + 3 °C of the lower T_m primer for 20 s (annealing), 72 °C for 30 s/kb of fragment to be amplified (extension); 72 °C for 5 min (polishing).
4. Test an aliquot (1/10th) of the PCR on a 1.2% agarose TBE gel. Use the rest of the sample to purify the amplified product by a PCR purification kit.
5. Digest PCR fragment with restriction enzymes (primers should be designed to contain recognition sites for appropriate restriction enzymes, which can subsequently be used for cloning of the PCR product). Gel purify digested PCR fragment on a 1.2% TBE agarose gel.
6. Digest 1 μg of pENTR11 plasmid vector with appropriate restriction enzymes. Gel purify the digested plasmid.
7. Set up a 10 μL ligation reaction by mixing 100 ng of digested pENTR11 plasmid, threefold molar excess of the digested PCR product, 1 μL of 10× T4 DNA ligase buffer, appropriate amount of water, and 1 μL of T4 DNA ligase. Incubate the reaction at room temperature for 2–4 h.

8. Transform ligation reaction into chemically competent TOP10 *E. coli*. Spread bacteria onto an LB agar plate containing 50 μg/mL of kanamycin. Incubate the plates overnight at 37 °C.
9. Screen the appearing colonies by colony PCR.
10. Purify plasmids from colonies carrying inserts of the expected size and verify identity of the fragments by sequencing.

3.2.2 Transfer of miRNA Genes into a Gateway Compatible Binary Expression Vector

1. Set up LR recombination reaction by mixing 1 μL (10 ng) of entry plasmid (pENTR11 containing the cloned miRNA gene) with 1 μL (50 ng) of plant binary destination vector (e.g., pK7WG2D). Add 0.5 μL of LR clonase and incubate the reaction at room temperature for 1–12 h.
2. Transform ligation reaction into chemically competent TOP10 *E. coli*. Spread bacteria onto an LB agar plate containing 100 μg/mL of spectinomycin. Incubate the plates overnight at 37 °C.
3. Screen the appearing colonies by colony PCR.
4. Purify plasmids from colonies and verify correct structure of the plasmid by restriction digestion.

3.3 Generation of Sensor Constructs

A crucial component of the reporter system is the sensor plasmid. We have designed a GFP–*Renilla* luciferase fusion gene platform for testing the effectiveness of miRNA binding sites in various configurations (3′UTR or ORF). The use of the fusion protein provides several advantages: (1) GFP serves as a tag to detect and quantify the expression levels of the fusion protein by sensitive GFP antibodies in Western blotting; (2) it allows parallel visual monitoring (via observing GFP fluorescence) and quantitative assessment of silencing (by measuring enzymatic activity of the *Renilla* luciferase) (*see* **Note 10**); (3) the linker region connecting the two functional domains of the fusion protein provides a suitable position for inserting miRNA target sites without disrupting the activities of GFP or *Renilla* luciferase (for generating ORF sensors).

To properly assess gene silencing, the activity measured on a wild-type sensor construct has to be compared to the activity observed on a corresponding mutant sensor. The two sensors differ from each other only in three point mutations. These mutations are introduced into the miRNA binding site at positions 10, 11 (flanking the AGO cleavage site), and 16 (the 3′ supplementary region).

3.3.1 Generation of miRNA Target Site Entry Vectors

1. pENTR11–*Renilla*–luciferase vector can be obtained by inserting the 982 nt long *Nco*I–*Not*I fragment of psiCHECK (Promega)—containing the *Renilla* luciferase ORF—into pENTR11 by conventional asymmetric sticky end ligation.

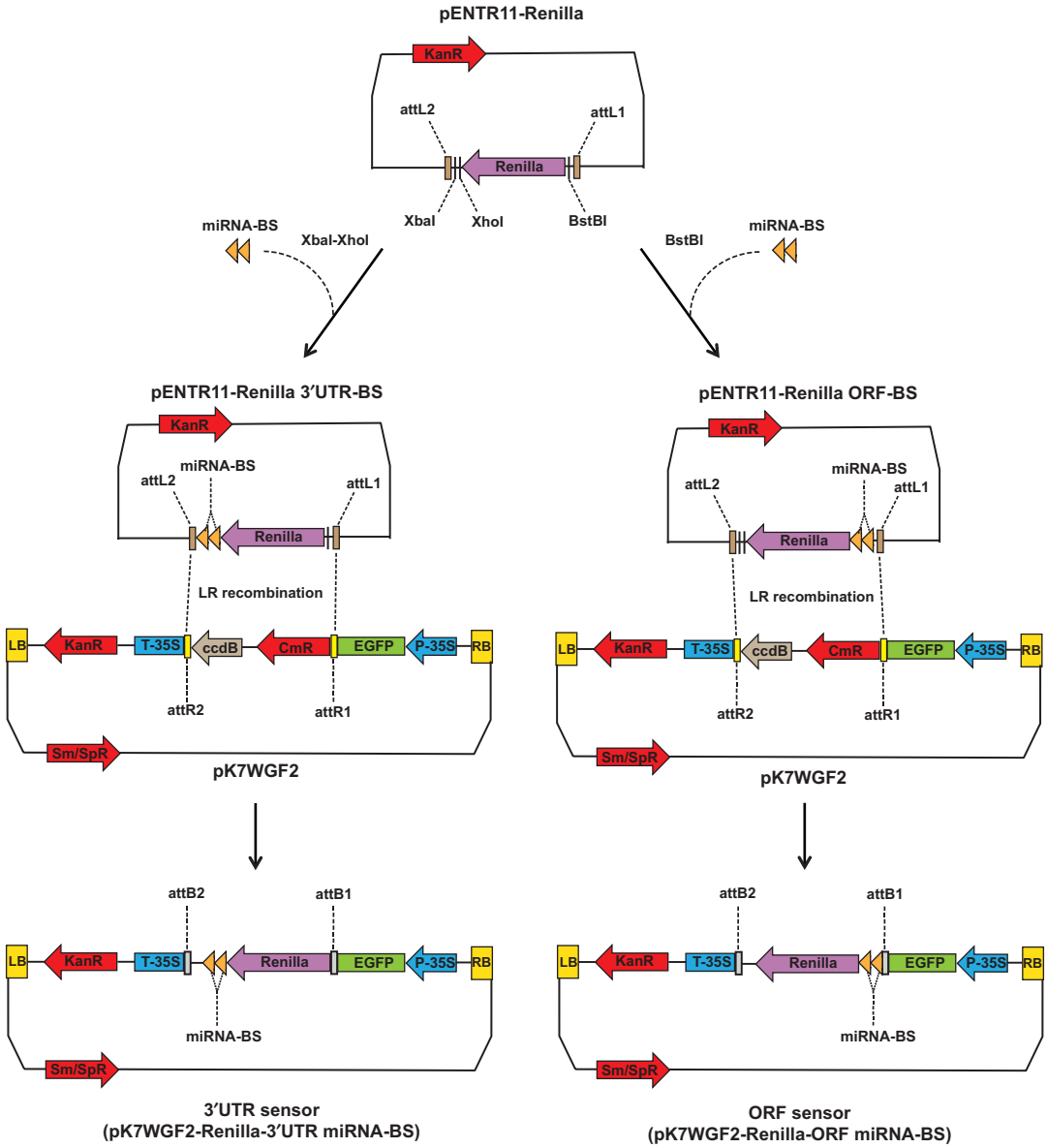


Fig. 2 Construction of binary sensor plasmids. Double-stranded DNA oligonucleotides containing tandem miRNA binding sites are inserted either at the 5′ or 3′ end of the *Renilla* luciferase coding region (in pENTR11 plasmid vector). In the subsequent LR recombination step, the *Renilla* luciferase coding region, along with the miRNA binding sites, is transferred into the pK7WGF2 binary plasmid. LR recombination ensures the precise in-frame fusion between the GFP and *Renilla* luciferase coding regions

2. Anneal tandem wild-type or mutant miRNA binding site containing complementary oligonucleotides as described above. Insert the double-stranded oligonucleotides either into *Xho*I–*Xba*I sites of pENTR11–*Renilla*–luciferase vector to obtain 3′ UTR entry vector or into the *Bst*BI site to get the ORF entry vector (*see* **Note 11**) (Fig. 2).

3.3.2 Generation of Binary 3′UTR and ORF Sensors

1. Use the 3′UTR or ORF entry vectors to set up LR recombination reactions with the Gateway compatible pK7WGF2 binary vector as described above.
2. Transform ligation reaction into chemically competent TOP10 *E. coli*. Spread bacteria onto an LB agar plate containing 100 μg/mL of spectinomycin. Incubate the plates overnight at 37 °C.
3. Screen the appearing colonies by colony PCR. Correct recombination leads to in-frame fusions between the *Renilla* luciferase and GFP coding regions. In the resulting plasmids the miRNA binding sites are located either in the 3′UTR (3′UTR sensors) or in the linker region connecting the GFP and *Renilla* luciferase domains of the fusion gene (ORF sensors) (Fig. 2).

3.4 Agroinfiltration

To ensure efficient co-delivery of components of the reporter system into plants, agroinfiltration is used. First, the constructed binary vectors are transformed into a suitable *A. tumefaciens* strain (e.g., C58C1). Next, the generated bacterium strains are introduced into *N. benthamiana* leaves in various combinations by syringe infiltration.

3.4.1 Agrobacterium Transformation and Growth

1. Streak C58C1 *A. tumefaciens* strain from a frozen glycerol stock onto a YEB–RT plate. Incubate the plate for 3–5 days at 28 °C.
2. Inoculate a single colony of C58C1 into 5 mL of YEB medium containing 20 μg/mL of rifampicin and 5 μg/mL of tetracycline. Shake the culture vigorously (200–250 rpm) overnight at 28 °C.
3. Add 2 mL of the overnight culture to 50 mL of YEB medium (w/o antibiotics) in a 250 mL flask. Shake the culture vigorously (250 rpm) at 28 °C until it grows to OD_{600} of 0.5–1.0 (this generally takes 4–6 h).
4. Chill the culture on ice for 10 min.
5. Centrifuge the cell suspension at 3000 × *g* for 10 min at 4 °C.
6. Discard supernatant and resuspend the bacterium pellet in 1 mL of ice-cold 20 mM $CaCl_2$ (*see* **Note 12**).
7. Add 1–2 μg of binary plasmid DNA to 100 μL of bacterium suspension.
8. Freeze mixture in liquid nitrogen.
9. Thaw the bacteria at 37 °C for 5 min.
10. Add 1 mL of YEB medium to the bacteria and incubate the mixture at 28 °C for 2–4 h. This period allows the bacteria to express the antibiotic resistance genes.
11. Centrifuge bacterium suspension in a centrifuge for 1 min at maximum speed.

12. Remove majority of supernatant leaving ~100 μL of YEB medium on the bacterium pellet.
13. Resuspend bacteria and spread suspension onto a YEB plate containing appropriate antibiotics (*see* **Note 13**).
14. Incubate the plate for 3–5 days at 28 °C.

3.4.2 Preparation of Agrobacteria for Infiltration

1. Inoculate a single colony of each C58C1 strain of the reporter system into 5 mL of YEB medium containing the appropriate antibiotics (AGO expression strain, miRNA expression strain, wild-type sensor/mutant sensor strains. Empty expression vector containing bacterium strains should be grown as well, as negative controls). In addition to the antibiotics, supplement the growth media with 10 mM of MES and 20 μM of acetosyringone.
2. Shake the cultures (250 rpm) at 28 °C overnight to reach saturation.
3. Centrifuge cultures at 3000 × *g* for 10 min at 25 °C.
4. Decant supernatants and resuspend bacterium pellets in 1 mL of freshly prepared infiltration media. Incubate suspensions at room temperature for at least 3 h.
5. Measure optical density of bacterium suspensions. Prepare $OD_{600} = 1$ dilutions from suspensions using infiltration medium.
6. Prepare infiltration mixtures from the strains of the reporter system (AGO expression strain, miRNA expression strain, sensor/mutant sensor strains). The mixing ratios of the strains have to be optimized for each AGO/miRNA/sensor combination. In addition, mixtures containing the appropriate empty vector strains, instead of the AGO or miRNA expression strains, should also be prepared as negative controls. In our hands, mixing the three strains at 1:1:1 ratio works well for most of the time. Setup of a typical experiment is given in Fig. 3.

3.4.3 Agroinfiltration

1. *N. benthamiana* plants are grown at 25 °C under long-day conditions (16 h light, 8 h dark). Four- to five-week-old plants are most suitable for infiltration. The day before the experiment, irrigate plants well in order to ensure that their stomata are fully open at the time of the infiltration.
2. Fill a 1 mL hypodermic syringe with the infiltration mixture. Infiltrate leaves from the abaxial side. For each infiltration mixture, use three plants. Infiltrate three leaves per plant. Into the left side of the leaf, infiltrate the wild-type sensor containing mixture, while into the right side of the same leaf, introduce the mutant sensor containing mixture (Fig. 3). Choose leaves that are fully expanded and avoid the wrinkled ones, which are highly resistant to infiltration.

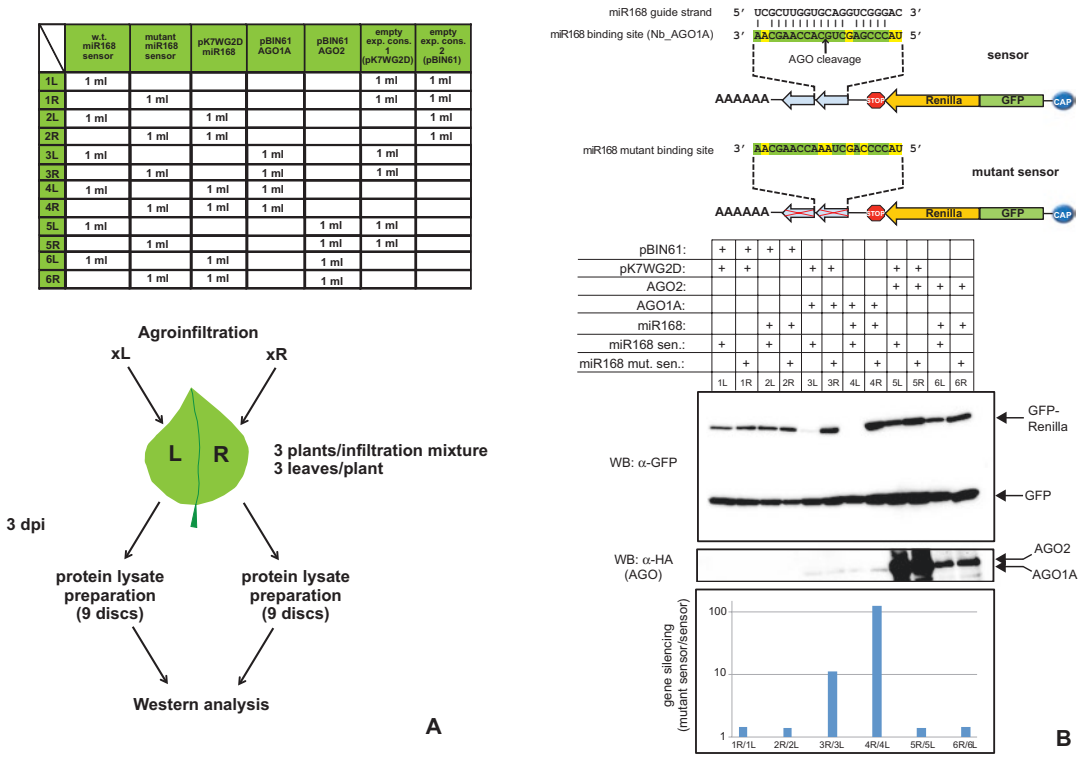


Fig. 3 Setup and result of a typical gene silencing assay. (**a**) Infiltration mixtures are assembled from diluted *Agrobacterium* suspensions (OD_{600} = 1) according to the table. For each infiltration use three plants. Infiltrate three leaves per plant. Into the left half of the leaf, the wild-type sensor containing mixture (xL) is infiltrated, while into the right half of the same leaf, the mutant sensor containing mixture (xR) is introduced. Samples are collected at 3 days post infiltration (dpi) and processed for protein lysate preparation. (**b**) Sensor activities are analyzed by quantitative Western blotting. The sensor-encoded GFP–*Renilla* luciferase fusion protein signals are normalized for the GFP signals. Gene silencing is plotted as the ratio of the amounts of GFP–*Renilla* fusion protein produced by the mutant and the wild-type sensors. Due to the complexity of the system, rigorous statistical analysis of the data is cumbersome and sometimes could even be misleading. Repeat of the experiments (at least three times) with sufficient number of biological parallels (at least three) is more effective. Conclusions should only be drawn from experiments, which are highly reproducible. Note that the differential effects of AGO1 and AGO2 on the miR168 sensors can clearly be observed. AGO1 preferentially incorporates miRNAs starting with a U, while AGO2 favors miRNAs with A in the same position (miR168 starts with a U). As a result, the miR168-specific sensor can only be silenced by AGO1-containing RISCs

3. Grow the infiltrated plants at 25 °C under long-day conditions for 72 h before analyzing protein expression.

3.5 Analysis of Protein Expression

As a measure of gene silencing, expression of the wild-type sensor/mutant sensor-encoded GFP–*Renilla* luciferase fusion protein is analyzed by quantitative Western blotting. Western blot analysis of proteins is composed of three consecutive steps. First the proteins are separated according to their apparent molecular masses using SDS-PAGE. Next, the gel is electro-transferred to nitrocellulose or

polyvinylidene fluoride (PVDF) membranes. Generally, nitrocellulose membranes give lower background than PVDF membranes; however, the latter is more durable. Therefore, if multiple rounds of stripping/immunological detection are planned, PVDF membranes should be used instead of nitrocellulose. The final step is the immunological detection of the protein to be studied. Since, Western blot analysis is a routine technique in most molecular biology laboratory, no detailed protocols for SDS-PAGE and electrotransfer of gels are provided here. Instead, the reader is referred to a number of excellent laboratory manuals [14, 15]. Only the protein lysate preparation and immunological detection/quantitation of signals are described here in more details.

3.5.1 Preparation of Protein Lysates from Infiltrated Leaves

1. Punch 1 cm diameter discs from both sides of infiltrated leaves. Pool discs from the same side of leaves infiltrated with the same suspensions (total of nine discs).
2. Collect the discs into a porcelain mortar. Freeze discs by pouring small volume of liquid nitrogen into the mortar and quickly grind them into a fine powder using a porcelain pestle.
3. Add 1800 μL of ice-cold lysis buffer to the mortar and continue homogenizing for 30 s.
4. Pour lysate into a 2 mL tube. Keep tubes on ice until all samples are collected.
5. Centrifuge lysates in a centrifuge at 20,000 × *g* for 20 min at 4 °C.
6. Transfer supernatants into clean tubes.
7. Store samples at −70 °C until further analysis.

3.5.2 SDS-PAGE and Electro-Transfer of Gel

1. Prepare resolving gel solution for 8% SDS-PAGE.
2. Cast 1.5 mm thick gels using the PROTEAN gel casting system from Bio-Rad.
3. Prepare stacking gel solution.
4. Cast stacking gel on top of the 8% resolving gel. Use 1.5 mm thick 10- or 15-well comb depending on the number of samples to be analyzed.
5. Prepare samples by mixing 30 μL of cleared protein lysate (from Subheading 3.5.1) with 6 μL of 6 × SDS sample buffer (*see* **Note 14**).
6. Heat samples at 95 °C for 5 min.
7. Load samples immediately onto the gel. Load pre-stained protein marker along the samples as well. Run gels in a Mini-PROTEAN electrophoresis chamber using SDS-PAGE running buffer.
8. Run gel at 100 V constant voltage until bromophenol blue reaches the bottom of the gel (usually ~2 h).

9. Blot the gel onto nitrocellulose membrane using Mini Trans-Blot Cell electroblotting apparatus (or equivalent equipment). Transfer is performed in Towbin buffer at 300 mA constant current in cold room (at ~4 °C) for 2 h.
10. Stain nitrocellulose membrane with Ponceau S solution to check protein transfer efficiency. If transfer is satisfactory, proceed with immunological detection.

3.5.3 Immunological Detection and Quantitation

1. Block membrane in blocking solution with gentle agitation on an orbital shaker for 1 h at room temperature.
2. To detect GFP and GFP–*Renilla* luciferase fusion protein, incubate membrane in diluted polyclonal anti-GFP antibody (1:2000 dilution in blocking solution) with constant agitation for 1 h at room temperature.
3. Wash membrane in TBS-T, three times for 10 min with agitation.
4. Incubate membrane with HRP-conjugated anti-rabbit IgG (1:10,000 dilution in blocking solution) with constant agitation for 1 h at room temperature.
5. Wash membrane in TBS-T, three times for 10 min with agitation.
6. Develop membrane using Clarity Western ECL Substrate (or similar) according to manufacturer's instructions.
7. Use Bio-Rad ChemiDoc XRS Plus System (or equivalent imaging system) to image membrane. The sensor-encoded GFP–*Renilla* luciferase fusion protein can be detected on the membrane at ~66 kDa, while the miRNA expression vector (or the pK7WG2D empty vector)-encoded GFP runs at ~33 kDa (Fig. 3). Take multiple pictures of the membrane with varying exposure times. Be sure to have images where signal intensities are not at saturating levels. For accurate quantitation of protein bands, use these images.
8. Use Image Lab software for identification and accurate quantitation of protein bands. Identification of bands can be done automatically or manually. Manual labeling of the bands becomes necessary if they are too faint for automatic detection. Image Lab generates an analysis table, which contains absolute values of band intensities (with background subtracted). The analysis table can be exported to Excel for further analysis. The GFP signal serves as an internal infiltration control and is used to normalize the GFP–*Renilla* luciferase fusion protein signal. Gene silencing can be defined as the ratio of the normalized fusion protein levels:

$$\text{gene silencing} = \text{GFP–}\textit{Renilla}\text{ protein}_{\text{(mutant sensor)}} \div \text{GFP–}\textit{Renilla}\text{ protein}_{\text{(wild–type sensor)}}$$

9. After exposure, rinse membrane twice with TBS-T and proceed with the detection of AGO protein expression. Repeated blocking of the membrane is not necessary.
10. Incubate membrane in diluted HRP-conjugated rat monoclonal 3F10 anti-HA antibody (1:2000 dilution in blocking solution) with constant agitation for 1 h at room temperature (*see* **Note 15**).
11. Wash membrane in TBS-T, three times for 10 min with agitation.
12. Develop membrane using Clarity Western ECL Substrate (or similar) as described above.

4 Notes

1. The Gateway compatible pK7WG2D and pK7WGF2 binary destination vectors along with a number of similar plasmids can be obtained from University of Gent, VIB, Plant Systems Biology (https://gateway.psb.ugent.be).
2. For general cloning, subcloning purposes TOP10, DH5α, or similar *E. coli* strains can be used. For propagating plasmid components of the Gateway system (entry vectors, destination vectors), the DB3.1 *E. coli* strain must be used. This strain tolerates the presence of the *ccd*B gene (a component of the destination cassette), which is highly toxic to routinely used *E. coli* strains.
3. Na_3VO_4 should be activated for maximal inhibition of protein phosphotyrosyl phosphatases as follows:
   (a) Prepare a 200 mM solution of Na_3VO_4.
   (b) Adjust the pH to 10 using either 1 N NaOH or 1 N HCl. At pH 10 the solution will be yellow.
   (c) Boil the solution until it turns colorless.
   (d) Cool to room temperature.
   (e) Readjust the pH to 10 and repeat **steps b** and **c** until the solution remains colorless and the pH stabilizes at 10.
   (f) Store the activated Na_3VO_4 in aliquots at −20 °C.
4. TBE agarose gel electrophoresis of RNA: Mix 1 μL of RNA (0.5–1 μg) and 4 μL of water. Add 5 μL of FDE to diluted RNA. Heat sample to 65 °C for 10 min (heating RNA to 65 °C in the absence of chelating agent results in hydrolysis of the molecule). Cool sample on ice and then load immediately onto a 1.2% TBE agarose gel.
5. Good quality chemically competent *E. coli* can be prepared following the Inoue protocol [16]. High-efficiency competent *E. coli* are also available from a number of commercial sources.

6. Colony PCR can be used to screen for bacterial colonies carrying plasmids with the desired insert. The protocol is carried out as follows:
   (a) Touch isolated bacterial colonies with a pipet tip and transfer the attached cells into 20 μL of water (in a PCR tube).
   (b) Lyse bacteria by heating the suspension to 95 °C for 5 min in a thermocycler.
   (c) Transfer 0.5 μL of the bacterial lysate into a PCR tube containing 12 μL of PCR mix. The PCR mix is assembled as follows: 0.25 μL of 10 μM forward primer, 0.25 μL of 10 μM reverse primer, 0.25 μL of 10 mM dNTP mix, 0.75 μL of 25 mM $MgCl_2$, 1.25 μL of 10× Taq buffer, 9 μL of water, and 0.25 μL of Taq polymerase. By upscaling the above volumes, a PCR master-mix can be prepared for screening of the desired number of colonies. At least one of the applied primers should be specific for the plasmid vector. This helps to eliminate amplification of background products, which derive from the unused insert. Additionally, it allows determining the orientation of the inserted DNA fragment within the vector.
   (d) Cycling conditions are the following: 95 °C for 2 min (initial denaturation); 40 cycles of 95 °C for 30 s (denaturation), T_m °C of the lower T_m primer for 30 s (annealing), 72 °C for 30 s/kb of the fragment to be amplified (extension); 72 °C for 7 min (polishing).
   (e) Check PCR reactions on a TBE agarose gel (percentage of the gel depends on the size of the expected fragment).
7. With careful choice of primers, subfragments of AGO ORFs can be amplified, which contain unique restriction sites in their terminal overlapping segments. These sites can be used to assemble the full-length ORF using conventional cloning techniques. Alternatively, the subfragments can be compiled by overlapping PCR. The two fragments to be fused have to be added to the PCR reaction as templates along with primers that are specific for the distal ends of the fragments. Since overlapping PCR is an inefficient process, highly processive enzymes with proofreading activity have to be used (e.g., Phusion).
8. Restriction sites should be chosen which are not present in the coding regions of *N. benthamiana* AGO ORFs. Design tag-encoding oligonucleotides, which after annealing produce ends that are compatible with different restriction enzymes (e.g., *Pme*I–*Sal*I). This allows cloning of the tag in front of the AGO ORFs in the proper orientation, thereby reducing the efforts of screening for properly tagged plasmid clones.
9. The polylinker region of pGEM-T easy plasmid vector contains recognition sites for a relatively large number of rare-cutting

restriction enzymes. Various combinations of these enzymes can be used to isolate the intact AGO ORFs from the vector. If necessary, additional recognition sites can also be added to the vector using oligonucleotides.

10. Using the specific substrate of *Renilla* luciferase (coelenterazine), we have demonstrated the functionality of the sensor plasmid-encoded GFP–*Renilla* fusion protein. We also found that the protein's enzymatic activity correlated perfectly with the signal intensity measured in quantitative Western blots using anti-GFP antibody. In addition, we demonstrated that the protein is suitable for dual-luciferase assays (combined with firefly luciferase normalization control).
11. When generating ORF sensors, carefully check that no in-frame STOP codons are introduced into the linker region between GFP and *Renilla* luciferase.
12. At this point the competent bacteria can be aliquoted into Eppendorf tubes (100 μL/tube) and snap frozen in liquid nitrogen. The frozen aliquots can be stored at −70 °C.
13. Either RTK or RTS plates are used. Rifampicin and tetracycline select for the strain-specific markers of C58C1. Kanamycin selection allows for the selection of the pBIN61-based AGO vectors. Spectinomycin is used to select for the pK7WG2D- or pK7WGF2-derived plasmids (miRNA expression vectors, sensors).
14. We found that measuring the exact protein concentration of lysates is not necessary if the volume of lysis buffer per leaf disc is kept constant (200 μL of lysis buffer/1 cm diameter leaf disc). In addition, the slight differences in protein concentration are corrected during the normalization for the internal GFP control.
15. Expression of the AGO proteins can be detected by either HA or FLAG antibodies. In our hands the 3F10 anti-HA antibody proved to be more sensitive than the M2 FLAG antibody. For detection of AGO proteins, stripping of the membrane is not necessary. All *N. benthamiana* AGOs are >100 kDa in their size; thus, none of the bands recognized by the GFP antibody (GFP at ~33 kDa, GFP–*Renilla* luciferase at ~66 kDa) interfere with their detection.

Acknowledgments

This work was supported by the Hungarian Scientific Research Fund (NK105850).

References

1. Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4):669–687. doi:10.1016/j.cell.2009.01.046
2. Pumplin N, Voinnet O (2013) RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat Rev Microbiol 11(11):745–760. doi:10.1038/nrmicro3120
3. Martinez de Alba AE, Elvira-Matelot E, Vaucheret H (2013) Gene silencing in plants: a diversity of pathways. Biochim Biophys Acta 1829(12):1300–1308. doi:10.1016/j.bbagrm.2013.10.005
4. Carbonell A, Carrington JC (2015) Antiviral roles of plant ARGONAUTES. Curr Opin Plant Biol 27:111–117. doi:10.1016/j.pbi.2015.06.013
5. Bologna NG, Voinnet O (2014) The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu Rev Plant Biol 65:473–503. doi:10.1146/annurev-arplant-050213-035728
6. Goodin MM, Zaitlin D, Naidu RA, Lommel SA (2008) Nicotiana benthamiana: its history and future as a model for plant-pathogen interactions. Mol Plant-Microbe Interact 21(8):1015–1026. doi:10.1094/MPMI-21-8-1015
7. Nakasugi K, Crowhurst R, Bally J, Waterhouse P (2014) Combining transcriptome assemblies from multiple de novo assemblers in the allo-tetraploid plant Nicotiana benthamiana. PLoS One 9(3):e91776. doi:10.1371/journal.pone.0091776
8. Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA, Martin GB (2012) A draft genome sequence of Nicotiana benthamiana to enhance molecular plant-microbe biology research. Mol Plant-Microbe Interact 25(12):1523–1530. doi:10.1094/MPMI-06-12-0148-TA
9. Nakasugi K, Crowhurst RN, Bally J, Wood CC, Hellens RP, Waterhouse PM (2013) De novo transcriptome sequence assembly and analysis of RNA silencing genes of Nicotiana benthamiana. PLoS One 8(3):e59534. doi:10.1371/journal.pone.0059534
10. Fatyol K, Ludman M, Burgyan J (2016) Functional dissection of a plant Argonaute. Nucleic Acids Res 44(3):1384–1397. doi:10.1093/nar/gkv1371
11. Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science 297(5589):2053–2056. doi:10.1126/science.1076311
12. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, Garcia JA, Paz-Ares J (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet 39(8):1033–1037. doi:10.1038/ng2079
13. Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121(2):207–221. doi:10.1016/j.cell.2005.04.004
14. Maniatis T, Fritsch E, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
15. Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K (eds) (2003) Current protocols in molecular biology. John Wiley & Sons, Inc, Hoboken, NJ
16. Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96(1):23–28. doi:10.1016/0378-1119(90)90336-P

Chapter 6

Immunoprecipitation and High-Throughput Sequencing of ARGONAUTE-Bound Target RNAs from Plants

Alberto Carbonell

Abstract

ARGONAUTE (AGO) proteins function in small RNA (sRNA)-based RNA silencing pathways to regulate gene expression and control invading nucleic acids. In posttranscriptional RNA silencing pathways, plant AGOs associate with sRNAs to interact with highly sequence-complementary target RNAs. Once the AGO–sRNA-target RNA ternary complex is formed, target RNA is typically repressed through AGO-mediated cleavage or through other cleavage-independent mechanisms. The universe of sRNAs associating with diverse plant AGOs has been determined though AGO immunoprecipitation (IP) and high-throughput sequencing of co-immunoprecipitated sRNAs. To better understand the biological functions of AGO–sRNA complexes, it is crucial to identify the repertoire of target RNAs they regulate. Here I present a detailed AGO–RNA IP followed by high-throughput sequencing (AGO RIP-Seq) methodology for the isolation of AGO ternary complexes from plant tissues and the high-throughput sequencing of AGO-bound target RNAs. In particular, the protocol describes the IP of slicer-deficient hemagglutinin (HA)-tagged AGO proteins expressed in plant tissues, the isolation of AGO-bound RNAs, and the generation of target RNA libraries for high-throughput sequencing.

Key words ARGONAUTE, RNA immunoprecipitation, High-throughput sequencing, RIP-Seq, Small RNA, RNA-Seq, Library preparation

1 Introduction

In plants, ARGONAUTE (AGO) proteins are the effectors of small RNA (sRNA)-based RNA silencing pathways regulating key biological processes such as development, response to stress, chromosome structure and genome integrity, and antiviral defense [1–4]. In posttranscriptional RNA silencing pathways, AGOs associate with specific sRNAs based on the identity of the sRNA 5′ nucleotide and on the structure of the sRNA duplex among other factors [5–9]. AGO-loaded sRNAs serve as guides for AGO recognition and interaction with highly sequence-complementary RNAs. Once the ternary complex including the AGO, the sRNA, and the target RNA is formed, target RNAs are typically repressed either through

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_6,

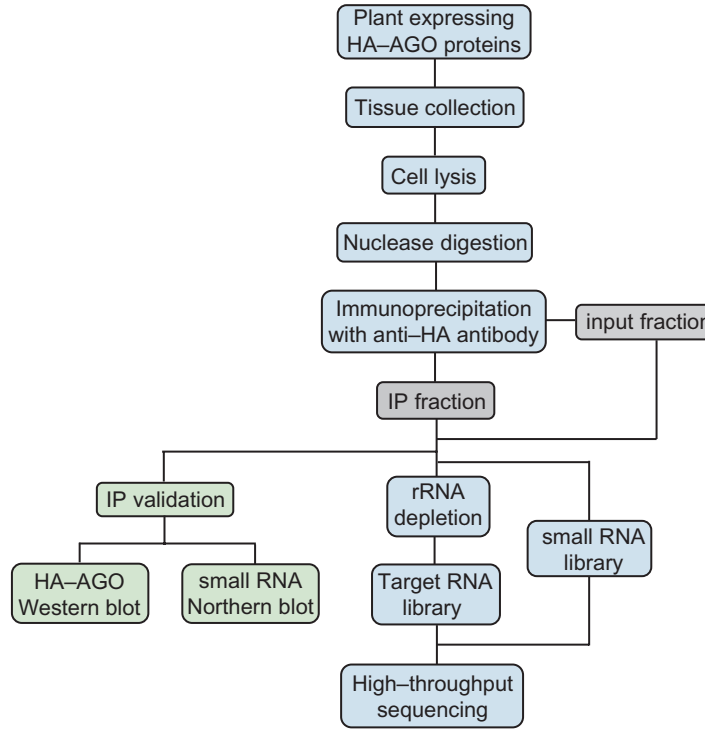


Fig. 1 Flowchart of the analytical steps of the AGO RIP-Seq methodology. Main steps are described in *light blue boxes*. Steps for the validation of the AGO immunoprecipitation (IP) are described in *light green boxes*

AGO-mediated endonucleolytic cleavage (or slicing) or through other cleavage-independent mechanisms such as target destabilization or translational inhibition [2].

AGO immunoprecipitation (IP) followed by high-throughput sequencing of bound sRNAs has been used to determine the universe of plant sRNAs associating with diverse AGOs in different species [5, 6, 8–11]. However, similar methods have not been described for the identification of AGO-bound target RNAs until recently [12]. *Arabidopsis thaliana* AGO1 target RNAs were efficiently co-immunoprecipitated and sequenced when using slicer-deficient hemagglutinin (HA)-tagged AGO1 forms but not with their wild-type catalytically active counterparts [12].

In this chapter, I present a detailed AGO–RNA IP followed by high-throughput sequencing (AGO RIP-Seq) protocol for the genome-wide analysis of AGO-bound target RNAs. I describe the IP of slicer-deficient HA-tagged AGO proteins from plant tissue, the isolation of associated RNAs, and the generation of target RNA libraries for high-throughput sequencing (Fig. 1).

2 Materials

2.1 Plant Material

This protocol is intended for processing plant tissue accumulating slicer-deficient HA-tagged AGO proteins (*see* **Note 1**).

2.2 Immunoprecipitation

1. IP buffer: 50 mM Tris–HCl at pH 7.4, 2.5 mM magnesium chloride ($MgCl_2$), 100 mM potassium chloride (KCl), 0.1% Nonidet P-40, 1 μg/mL leupeptin, 1 μg/mL aprotinin, 0.5 mM phenylmethanesulfonyl fluoride (PMSF), one tablet of cOmplete™ ethylenediaminetetraacetic acid (EDTA)-free Protease Inhibitor Cocktail tablets (Sigma–Aldrich) per 50 mL IP buffer, and 50 U/mL RNase inhibitor.
2. Proteinase K (PK) buffer: 0.1 M Tris–HCl pH 7.4, 10 mM EDTA, 300 mM sodium chloride (NaCl), 2% sodium dodecyl sulfate (SDS), 1 μg/μL proteinase K.
3. 2× Protein dissociation buffer (PDB): 0.0625 M Tris pH 6.8, 2% SDS, 10% glycerol, 10% ß-mercaptoethanol, and 0.02% bromophenol blue.
4. Anti-HA (12CA5) antibody (Sigma–Aldrich).
5. Protein A agarose (Sigma–Aldrich).
6. Micrococcal nuclease (MNase, Worthington).
7. Ethylene glycol-bis (ß-aminoethyl ether)-N, N, N′, N′-tetraacetic acid (EGTA) 0.5 M pH 8 sterile.
8. Sterile Miracloth paper (Merck Millipore).

2.3 RNA Extraction

1. IP buffer-equilibrated phenol: 75% Tris-saturated phenol pH 4.5 and 25% IP buffer.
2. Phenol/chloroform/isoamyl alcohol, 25:24:1.
3. Chloroform.
4. Ethanol.
5. Sodium acetate (NaOAc) 3 M and pH 5.2.
6. GlycoBlue (Thermo Fisher Scientific).
7. Diethyl pyrocarbonate (DEPC)-treated H_2O.
8. NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).

2.4 Western Blot

1. NuPAGE Novex 4–12% Bis–Tris Gel 1.0 mm (Thermo Fisher Scientific).
2. Protran nitrocellulose transfer membrane (VWR).
3. Ponceau S solution: 0.1% Ponceau S and 5% acetic acid.
4. Phosphate-buffered saline (PBS) pH 7.4 (10×): 1.06 mM potassium phosphate monobasic (KH_2P0_4), 155.17-mM NaCl, and 2.97 mM sodium phosphate dibasic (Na_2HPO_4–$7H_2O$).
5. 1× PBST: 1× PBS and 0.1% Tween 20.

6. Blocking solution: 1× PBST and 5% milk powder.
7. Anti-HA-peroxidase high-affinity clone 3F10 (Sigma–Aldrich).
8. Western Lightning Plus ECL kit (PerkinElmer).

2.5 Northern Blot

1. 0.5× Tris–borate–EDTA (TBE).
2. 17% Polyacrylamide gel containing 7 M urea in 0.5× TBE. Mix 17 mL of 30% polyacrylamide (acrylamide/bisacrylamide, 37.5:1), 12.6 g of urea, 1.5 mL of 10× TBE, and 2 mL of H_2O. Mix thoroughly by inversion; do not shake or vortex as this incorporates air bubbles to the solution which inhibit polymerization. Heat to 65 °C (bath) for 10 min to dissolve the urea. Let for additional 10 min on the bench to allow final resuspension of urea. Add 25 μL TEMED and mix by inversion. Add 150 μL of 10% ammonium persulfate and mix quickly by inversion. Pour the gel and allow it to polymerize for at least 30 min.
3. Whatman papers Protean XL size (Bio-Rad).
4. Nytran SuperCharge nylon membrane (Sigma–Aldrich).
5. USB OptiKinase (Affymetrix).
6. [^{32}P] γ-ATP, 6000 ci/mmol, and 10 mCi/mL (PerkinElmer).
7. P-6 spin columns (Bio–Rad).
8. PerfectHyb™ Plus buffer (Sigma–Aldrich).
9. Wash buffer 1: 2× SSC and 0.2% SDS.
10. Wash buffer 2: 1× SSC and 0.1% SDS.
11. 3× Saline sodium citrate (SSC).
12. DEPC-treated H_2O.

2.6 RT-PCR

1. Superscript III First-Strand Synthesis System (Thermo Fisher Scientific).
2. Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific).

2.7 RNA-Seq

1. DNA/RNA LoBind 1.5 mL tubes (Eppendorf).
2. Turbo DNA-free Kit (Thermo Fisher Scientific).
3. DEPC-treated H_2O.
4. Ribo-Zero rRNA Removal Kit (Plant Leaf) (Illumina).
5. Superscript II (Thermo Fisher Scientific).
6. Random primer.
7. SUPERase-In (Thermo Fisher Scientific).
8. Dithiothreitol (DTT) 0.1 M solution.
9. Agencourt RNAClean XP (Beckman Coulter).

10. 10× Blue buffer (Enzymatics).
11. Ribonuclease H (Thermo Fisher Scientific).
12. DNA polymerase I (Enzymatics).
13. Agencourt AMPure XP (Beckman Coulter).
14. EB (Qiagen).
15. 12P XP buffer: 12% PEG 8000 and 2.5 M NaCl.
16. End-Repair Mix LC (Enzymatics).
17. Klenow 3′–5′exo (Enzymatics).
18. 10× Hybridization buffer: 1 M NaCl, 0.1 M Tris–HCl pH 7.8, and 0.1 M EDTA pH 8.0.
19. TruSeq adapters (Table 1).
20. T4 DNA Ligase (Rapid) (Enzymatics).
21. Polyethylene glycol (Sigma–Aldrich).
22. Uracil DNA glycosylase (Enzymatics).
23. Phusion Hot Start II High-Fidelity DNA Polymerase (Finnzymes).
24. Paired-end (PE) primer and indexed adapters (Table 1).
25. Superscript II reverse transcriptase (Thermo Fisher Scientific).
26. Magnetic stands for PCR and 1.5 mL tubes.
27. NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).
28. Qubit DNA HS100 assay kit (Thermo Fisher Scientific).
29. Qubit fluorometer (Thermo Fisher Scientific).
30. RNA 6000 Nano kit (Agilent).
31. DNA High-Sensitivity (HS) kit (Agilent).
32. Bioanalyzer instrument (Agilent).

3 Methods

3.1 Immunoprecipitation

All the steps are performed in the cold room (at 4 °C) except the nuclease and PK treatments. It is recommended to prepare some of the materials the day before (*see* **Note 2)**.

1. Mix 9 mL of IP buffer and 2 mL of protein A agarose solution (previously well resuspended by pipetting) in a 15-mL conical tube. Rotate for 1 h at 4 °C. Stop rotation, allow beads to sediment at the bottom of the tube (*see* **Note 3**), and remove supernatant. Add 1 mL of IP buffer and resuspend beads by pipetting using a wide orifice tip. Rotate for 30 min at 4 °C.
2. Add 2 g of tissue to pre-cooled mortar (*see* **Note 4)**. Add a small amount of liquid nitrogen (*see* **Note 5)**, and grind

Table 1
Oligonucleotides used in transcript library preparation

Name	Nucleotide sequence
PE primer-F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACA CGACGCTCTTCCGATCT
TruSeq adapter 1	A*A*TGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAC GACGCTCTTCCGAT*C*T
TruSeq adapter 2	/5Phos/G*A*TCGGAAGAGCACACGTCTGAACTCCAGTC*A*C
Indexed adapter 1	CAAGCAGAAGACGGCATACGAGATtcgccttaGTGACTGGAGTTC AGACGTGT
Indexed adapter 2	CAAGCAGAAGACGGCATACGAGATctagtacgGTGACTGGAGTTC AGACGTGT
Indexed adapter 3	CAAGCAGAAGACGGCATACGAGATttctgcctGTGACTGGAGTTC AGACGTGT
Indexed adapter 4	CAAGCAGAAGACGGCATACGAGATgctcaggaGTGACTGGAGTTC AGACGTGT
Indexed adapter 5	CAAGCAGAAGACGGCATACGAGATggactcctGTGACTGGAGTTC AGACGTGT
Indexed adapter 6	CAAGCAGAAGACGGCATACGAGATtaggcatgGTGACTGGAGTTC AGACGTGT
Indexed adapter 7	CAAGCAGAAGACGGCATACGAGATctctctacGTGACTGGAGTTC AGACGTGT
Indexed adapter 8	CAAGCAGAAGACGGCATACGAGATcagagaggGTGACTGGAGTTC AGACGTGT
Indexed adapter 9	CAAGCAGAAGACGGCATACGAGATgctacgctGTGACTGGAGTTCA GACGTGT
Indexed adapter 10	CAAGCAGAAGACGGCATACGAGATcgaggctgGTGACTGGAGTTCA GACGTGT
Indexed adapter 11	CAAGCAGAAGACGGCATACGAGATaagaggcaGTGACTGGAGTTCA GACGTGT
Indexed adapter 12	CAAGCAGAAGACGGCATACGAGATgtagaggaGTGACTGGAGTTCA GACGTGT

*—Phosphorothioate bond
/5Phos/—5′ primer phosphorylation
Unique index sequences are in low case

vigorously with a pestle until obtaining a fine powder. Add 25 mL of IP buffer and homogenize. Transfer 12 mL of homogenate to each of two new 15 mL tubes (a total of 24 mL of homogenate are transferred). Centrifuge for 5 min at 12,000 × *g* at 4 °C to pellet cell debris. Filter the clarified homogenates with two layers of sterile Miracloth paper, and collect filtered homogenates in a single sterile flask.

3. Use 2 mL of filtered homogenate for input RNA (total RNA before IP) extraction (*see* Subheading 3.2). Mix 50 µL of filtered homogenate and 50 µL of 2× PDB in a 1.5 mL tube for input protein analysis, incubate at 100 °C for 3 min, and store at −80 °C until Western blot analysis (*see* Subheading 3.3).
4. Transfer 9 mL of filtered homogenate to each of two new 15-mL tubes for IP. Add 2250 U of MNase to each tube containing 9 mL of filtered homogenate (*see* **Note 6)**. Incubate for 5 min at 22 °C and transfer to ice. Add 90 µL EGTA 0.5 M pH 8 to each tube and mix.
5. Transfer 1 mL of digested homogenate from each tube to a new 2 mL tube for nuclease-digested (nd) input (nd-input) RNA extraction (*see* Subheading 3.2).
6. Add 128 µg anti-HA (12CA5) antibody to each 15 mL tube containing now 8 mL of nuclease-digested homogenate. Rotate for 30 min at 4 °C. Add 0.8 mL of protein A agarose solution to each tube and rotate for 30 min at 4 °C. Centrifuge for 30 s at 2500 × *g* at 4 °C to pellet beads. Remove the supernatant by careful pipetting.
7. Add 3.5 mL IP buffer to each tube and mix. Rotate for 2 min. Centrifuge for 30 s at 2500 × *g* at 4 °C to pellet beads. Remove the supernatant by careful pipetting. Do six washes of 10 min each: for each wash, add 7 mL IP buffer, rotate for 10 min, centrifuge for 30 s at 2500 × *g* at 4 °C, and discard supernatant.
8. Pipette 20 µL of beads from each tube, transfer to the same protein IP tube, add 40 µL of 2× PDB, mix with vortex for 10–30 s, incubate at 100 °C for 3 min, and store at −80 °C for Western blot analysis (*see* Subheading 3.3).
9. Add 800 µL of pre-warmed PK buffer containing RNase inhibitor to each tube containing the remaining beads (1600 µL approximately) (*see* **Note 7**). Add 2400 µg of proteinase K to each tube and mix with vortex. Incubate for 15 min at 65 °C with intermittent shaking (*see* **Note 8**).
10. Centrifuge for 1 min at 2500 × *g* at room temperature to pellet beads. Transfer each supernatant (1 mL approximately each) to a new 2 mL tube. Proceed to IP RNA extraction (*see* Subheading 3.2).

3.2 RNA Extraction

All centrifugations are done in a microfuge at 4 °C.

1. Mix the RNA-containing solution (nondigested filtered homogenate, digested filtered homogenate, or supernatant after PK treatment for input RNA, nd-input RNA, or IP RNA samples, respectively) with equal volume of IP buffer-saturated phenol.
2. Centrifuge at 14,000 rpm for 5 min.

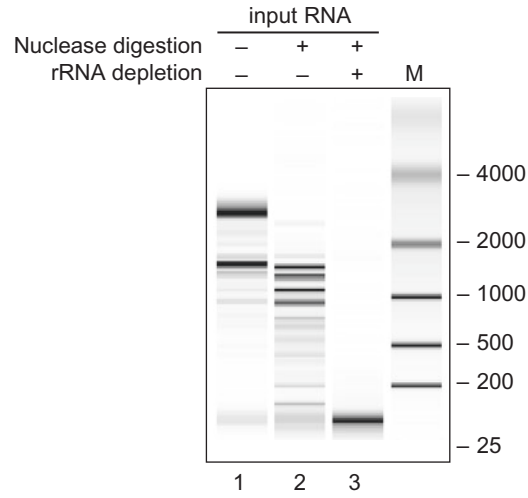


Fig. 2 Analysis of nuclease digestion and rRNA depletion. Bioanalyzer analysis of input RNA treated with nuclease and depleted for rRNAs (lane 3). Controls include input RNA untreated (lane 1) or solely digested with nuclease (lane 2). M, molecular weight marker (the size of the bands are indicated in nucleotides)

3. Transfer aqueous phase to fresh tube and repeat extraction with equal volume of phenol/chloroform/isoamyl alcohol.
4. Centrifuge at 14,000 rpm for 5 min.
5. Transfer aqueous phase to a new tube and repeat extraction with equal volume of chloroform.
6. Centrifuge at 14,000 rpm for 5 min at 4 °C.
7. Precipitate RNA from final aqueous phase with 0.1 volume of NaOAc, 20 μg of GlycoBlue, and 2.5 volume of 100% ethanol. Incubate overnight at −80 °C.
8. Centrifuge at 14,000 rpm for 30 min to pellet RNA.
9. Add cold 75% ethanol and centrifuge at 14,000 rpm for 10 min.
10. Remove ethanol and allow to air-dry for at least 10 min.
11. Resuspend input RNA and nd-input RNA in 20–30 μL DEPC-treated H_2O (*see* **Note 9**). Use 1 μL to measure the RNA concentration with a spectrophotometer. Verify nuclease digestion (*see* **Note 10**) (Fig. 2).
12. Resuspend IP RNA in 14 μL DEPC-treated H_2O (*see* **Note 9**). Use 12 μL for transcript library construction and 2 μL for sRNA library construction using standard protocols [13] if needed (*see* **Note 11**).

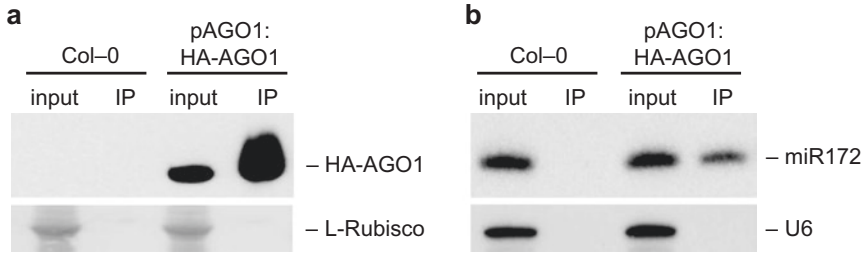


Fig. 3 Validation of the AGO immunoprecipitation. Samples from *Arabidopsis thaliana* wild-type Col-0 or transformed with a construct expressing HA-tagged AGO1 forms under endogenous regulatory sequences were analyzed. (**a**) Western blot analysis of HA–AGO1 accumulation in input and IP fractions. L-Rubisco (ribulose-1.5-bisphosphate carboxylase/oxygenase) stained membrane is included as input loading and immunoprecipitation control. (**b**) Northern blot analysis of *A. thaliana* miR172 accumulation in input and IP fractions. U6 blot is included as input loading and immunoprecipitation control

3.3 Western Blot Analysis of HA–AGO Protein Immunoprecipitation

Confirm HA–AGO protein IP by Western blot analysis using standard protocols. Next are some specifications:

1. Load 5–10 µL of input protein and 2.5–5 µL of IP protein samples into a 4–12% Bis–Tris mini gel. For IP protein samples, be sure to pipette from the top of the tube to avoid bead pipetting. Loading beads onto the gel will lead to poor protein separation. Load a protein marker, and run the gel at 150 V for 1 h and 30 min or long enough to resolve the size of the HA–AGO protein (e.g., ~116 kDa for *A. thaliana* HA–AGO1).
2. Transfer proteins to the nitrocellulose membrane. Stain membrane with Ponceau S solution to check protein transfer efficiency (*see* **Note 12**) (Fig. 3a).
3. Incubate membrane in blocking solution with shaking for 30 min at room temperature. Discard blocking solution and add 1× PBS with anti-HA-peroxidase 3F10 antibody (25 U/mL stock) at a 1:1000 dilution. Incubate with shaking for 2 h at room temperature, and wash four times in 1× PBST (10 min/wash). Follow the manufacturer's instructions of the Western Lightning Plus ECL kit or similar for electroluminescence-based detection of HA–AGO proteins (Fig. 3a).

3.4 Northern Blot Analysis of sRNA Co-Immunoprecipitation

Confirm the co-IP of at least one known AGO-interacting sRNA by Northern blot following standard protocols (*see* **Note 13**). Next are some specifications:

1. Prepare a 17% polyacrylamide gel containing 7 M urea in 0.5× TBE. Pre-run the gel at 180 V in 0.5× TBE for 1 h. Rinse well before loading samples as urea accumulates at the bottom of the gel. Heat samples for 10 min at 65 °C and immediately quench on ice briefly.

2. Load 2–10 μg of input RNA and 10–20% of IP RNA (*see* **Note 14**). Run at 180 V in 0.5× TBE until the bromophenol blue reaches the bottom of the gel (~4 h).
3. Assemble the blot sandwich with extra-thick Whatman papers (one on each side), gel, and positively charged nylon membrane in a semidry chamber. Transfer for 30 min at 500 mA. Auto-cross-link the membrane at 1200 μJ × 100. Store membranes between two sheets of filter paper until use.
4. Prepare the following reaction mix (*see* **Note 15**):

DNA or LNA oligonucleotide (10 μM)	1 μL
Polynucleotide kinase 10× buffer	1 μL
DEPC-treated H_2O	3 μL
[^{32}P] γ-ATP (6000 ci/mmol; 10 mCi/mL)	4 μL
USB OptiKinase	1 μL

5. Incubate for 60 min at 37 °C. Purify probe on P-6 spin columns according to manufacturer's instructions. Quantify counts per million (CPM)/μL.
6. Place the membrane (RNA side-up) in a hybridization tube and pre-hybridize with rotation for at least 5 min at 38–42 °C (*see* **Note 16**) in 5 mL of PerfectHyb™ Plus buffer.
7. Mix 1,000,000–2,000,000 CPM of probe with 200 μL of PerfectHyb™ Plus buffer, incubate 2 min at 95 °C, and transfer immediately to ice briefly.
8. Add to hybridization tube and incubate for 12–16 h at 38–42 °C (*see* **Note 16**).
9. Remove hybridization solution and wash the membrane five times with pre-heated wash solutions as follows: wash and rinse buffer 1 thoroughly, wash buffer 1 for 5 min with rotation at 38–42 °C, wash buffer 1 for 20 min with rotation at 55 °C, wash buffer 1 for 20 min with rotation at 55 °C, wash buffer 2 for 20 min with rotation at 55 °C, and wash buffer 2 for 30 min with rotation at 55 °C (*see* **Note 17**).
10. Rinse membrane briefly in 3× SSC, then air-dry briefly, and cover in transparent plastic wrap.
11. Autoradiograph (Fig. 3b).

3.5 RT-PCR Analysis of Target RNAs

AGO-bound target RNAs can be analyzed by RT-PCR (*see* **Note 6**). In this case, cDNA is obtained from 2 to 4 μg of input RNA or 10 to 25% of IP RNA using the Superscript III system according to manufacturer's instructions. Standard PCR reactions using Phusion High-Fidelity DNA Polymerase are designed to amplify fragments including the sRNA cleavage site present in the target transcript of analysis (*see* **Note 18**).

3.6 Preparation of Target RNA Libraries for High-Throughput Sequencing

In the target RNA library preparation protocol described below, nucleic acids are purified from all enzymatic reactions using SPRI magnetic beads (*see* **Note 19**). Transcript libraries are made solely from nd-input RNA and IP RNA samples as both samples have been nuclease treated, which favors later computational analyses.

3.6.1 Preparation of Transcript Fragments

The use of DNA/RNA LoBind 1.5 mL tubes is recommended to increase the RNA recovery in **steps 1** and **2**.

1. Prepare the DNase I digestion reaction by treating 20 μg of nd-input RNA or 80–100% of IP RNA with TURBO DNA-free kit following the manufacturer's instructions.
2. Measure the concentration of DNAse I-digested nd-input RNA with a NanoDrop spectrophotometer (*see* **Note 20**).
3. Use 5 μg or 100% of DNAse I-digested nd-input RNA or IP RNA for ribosomal RNA (rRNA) depletion treatment with Ribo-Zero kit following manufacturer's instructions (*see* **Note 21**).
4. Resuspend nd-input RNA and IP RNA in 12 μL and 8 μL of DEPC-treated H_2O, respectively.
5. Check rRNA depletion (Fig. 2) (*see* **Note 22**).
6. Use 1 μL of Ribo-Zero-treated nd-input RNA to measure its concentration with a spectrophotometer.

3.6.2 cDNA Synthesis

1. Prepare a 12 μL aliquot including 100 ng or 100% of Ribo-Zero-treated nd-input RNA or IP RNA, respectively, and 4 μL First-Strand buffer and DEPC-treated H_2O (*see* **Note 23**).
2. Prepare the reverse transcription reaction by mixing the following components:

RNA aliquot from the previous step	12 μL
Random primer (0.5 μg/μL)	0.5 μL
SUPERase-In (20 U/μL)	0.75 μL
0.1 M DTT	1 μL

3. Heat at 65 °C for 3 min, place on ice for 2 min, and add the following:

DEPC-treated H_2O	3.85 μL
0.1 M DTT	1 μL
10 mM dNTPs	0.25 μL
SUPERase-In (20 U/μL)	0.5 μL
Superscript II (200 U/μL)	0.5 μL

4. Transfer content to PCR strips, and proceed to reverse transcription reaction using the following PCR program: 10 min at 25 °C, 50 min at 42 °C, and 15 min at 70 °C and hold at 4 °C.
5. Purify the RNA/cDNA hybrid as follows: mix 21 μL of reverse transcription reaction, 38 μL of RNAClean XP beads, and 19 μL of 100% ethanol (*see* **Note 24**); pipette up and down ten times until the solution is well mixed; place the PCR strip in a magnetic stand for PCR tubes for 5–10 min to separate beads from solution; remove the supernatant without disturbing the ring of separated magnetic beads; wash beads twice with 100 μL 75% ethanol, incubating for 30 s at room temperature; remove PCR strip from the magnet; add 17 μL of DEPC-treated H_2O; pipette up and down 10 times until the solution is well mixed; centrifuge briefly at low speed; incubate for 2 min at room temperature; place the tubes back on the magnetic stand for 1 min; transfer the supernatant (16 μL eluate) to a new PCR strip.
6. Prepare the second strand synthesis reaction by mixing the following components (*see* **Note 25**):

RNA/cDNA hybrid	16 μL
10× Blue buffer	2 μL
dUTP mix (10 mM dA, dC, dG and 20 mM dU)	1 μL
RNAse H (2 U/μL stock)	0.5 μL
DNA polymerase I (10 U/μL)	1 μL
0.1 M DTT	0.5

7. Incubate at 16 °C for 2.5 h.
8. Purify double-stranded DNA (dsDNA) with AMPure XP beads (*see* **Note 19**) as follows: mix 21 μL of dsDNA, 38 μL of AMPure XP beads, and 19 μL of 100% ethanol; mix by pipetting ten times; let the mixed samples incubate for 5 min at room temperature for maximum recovery; place the PCR strip onto magnetic stand for 5 min to separate beads from solution; remove the supernatant without disturbing the ring of separated magnetic beads; wash beads twice with 100 μL 75% ethanol, incubating for 30 s at room temperature; air-dry for 5 min, lids opened; remove PCR strip from the magnetic stand; resuspend beads in 34 μL EB; incubate the elution suspension for 2 min at room temperature; place the tubes back on the magnetic stand for 1 min; transfer the supernatant (32 μL of eluate) to a new tube. Save 16 μL of eluate in −80 °C freezer.

3.6.3 Library Construction

To prepare transcript libraries from nd-input RNA or IP RNA, follow standard protocols [14]. Here are some specifications:

1. Prepare 12P XP beads by replacing the stock buffer of AMPure XP beads with 12P XP buffer as follows: add 1 mL of AMPure XP beads to a 1.5 mL tube; place onto a magnetic stand until the solution becomes clear; remove the supernatant; wash the beads twice with DEPC-treated H_2O; resuspend the beads in 1 mL 12P XP buffer (*see* **Note 26**).
2. Prepare the end-repair reaction on ice by mixing the following components:

dsDNA	16 μL
10× End-repair buffer	2 μL
10 mM dNTPs	1 μL
End-repair mix LC	1 μL

3. Incubate at 20 °C for 30 min in thermocycler.
4. Purify end-repaired DNA using 28 μL of AMPure XP beads with 14 μL of 100% ethanol. Elute with 18-μL DEPC-treated H_2O. Transfer 17 μL of eluate to a new tube.
5. Prepare the dA-tailing reaction by mixing the following components:

DNA	17 μL
10× Blue buffer	2 μL
10 mM dATP	1 μL
Klenow 3′–5′ exo	0.5 μL

6. Incubate at 37 °C for 30 min.
7. Purify A-tailed DNA using 28 μL of AMPure XP beads with 14 μL of 100% ethanol. Elute with 11 μL of DEPC-treated H_2O. Transfer 10 μL of eluate to a new tube.
8. Prepare the Y-shape adapter ligation by mixing the following components (*see* **Note 27**):

DNA	10 μL
Annealed TruSeq adapters (2 μM each)	1 μL
2× Ligation buffer	12 μL
T4 DNA ligase	1 μL

9. Incubate for 20 min at 20 °C in thermocycler.

Table 2
Multiplexing of transcript libraries for high-throughput sequencing

Plex	Indexed adapter selection
1 plex (no pooling)	Any indexed adapter
2 plex	Option A: 1 and 2 Option B: 2 and 4
3 plex	Option A: 1, 2, and 4 Option B: 3, 5, and 6
4 or 5 plex	Option A: 1, 2, 4, and any other indexed adapter Option B: 3, 5, 6, and any other indexed adapter
6 plex	1–6
7–12 plex, single index	1–6 and any other indexed adapter
7–12 plex, dual index	Option A: 1, 2, 4, and any other indexed adapter Option B: 3, 5, 6, and any other indexed adapter

10. Take 12 μL (half) and save in −80 °C freezer.
11. Mix the other half (12 μL) of the ligation product and mix with 12 μL of 12P XP beads. Incubate for 6 min at room temperature and 5 min in magnetic stand, keep the supernatant (24 μL), and discard the beads.
12. Mix the supernatant with 12 μL of AMPure XP beads and 5 μL 40% of PEG8000. Incubate for 6 min at room temperature, wash twice, and elute in 11 μL of DEPC-treated H_2O. Transfer 10 μL to a new tube. Mix with 12 μL of AMPure XP, incubate for 6 min at room temperature and 5 more min in the magnetic stand, and wash once. Elute in 32 μL of EB (*see* **Note 28**). Transfer 30 μL in a new tube.
13. Save 15 μL of elute in −80 °C freezer.

3.6.4 Library Amplification and Multiplexing

Up to 12 samples can be multiplexed for paired-end sequencing using the adaptors listed in Table 1. The diverse multiplexing options with these specific adaptors are described in Table 2.

1. Prepare the dUTP excision reaction by mixing the following components:

DNA	5 μL
Uracil DNA glycosylase	1 μL

2. Incubate for 30 min at 37 °C.

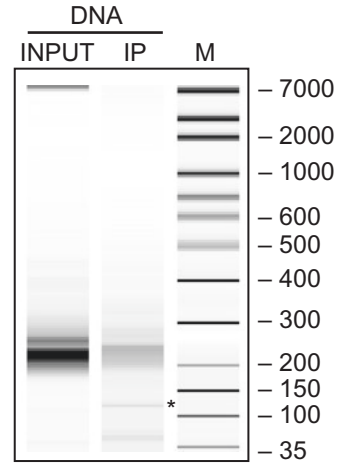


Fig. 4 Analysis of the transcript library amplicons. Bioanalyzer analysis of the library amplicons obtained with nuclease-digested input and IP fractions. The asterisk indicates the band corresponding to primer–dimer products. M, molecular weight marker (the size of the bands are indicated in base pairs)

3. Prepare the amplification reaction by mixing the following components:

Uracil DNA glycosylase-digested DNA	16 µL
5 mM PE primer F	1 µL
5× Phusion HF buffer	6 µL
10 mM dNTPs	1 µL
DEPC-treated H_20	4.5 µL
Phusion Hot Start II DNA polymerase	1 µL
5 mM indexed adapter	1 µL

4. Mix, centrifuge briefly at low speed, and transfer to thermocycler. Use the following PCR program: 30 s at 94 °C; 14–18 cycles (*see* **Note 29**) of 30 s at 98 °C, 30 s at 65 °C, and 30 s at 72 °C; 5 min at 72 °C; and hold at 4 °C.
5. Purify library using 43 µL of AMPure XP beads. Elute with 13 µL of EB. Transfer 12 µL in a new tube.
6. Measure the library concentration by loading 1 µL of purified library DNA in a Qubit fluorometer following the Qubit DNA HS Assay kit protocol (*see* **Note 30**). It is recommended to analyze DNA amplicons in a Bioanalyzer instrument following the DNA HS kit protocol (Fig. 4). If the amount of the band corresponding to primer dimers (~128 bp) is visibly high

(Fig. 4), then an additional purification with 1.2 volume of AMPure XP beads is highly recommended to remove primer dimers.

7. Prepare a 2–10 nM sample in a 20 μL volume for sequencing in a high-throughput sequencer (e.g., Illumina HiSeq 2000).

4 Notes

1. Inflorescence tissue from T4 transgenic *Arabidopsis thaliana* plants (4–6 weeks old, flower stages 1–12) expressing HA-tagged AGO proteins with authentic regulatory sequences or agroinfiltrated leaves from *Nicotiana benthamiana* plants transiently expressing HA-tagged AGOs with the *Cauliflower mosaic virus* (CaMV) 35S promoter have been analyzed [12]. Target RNA recovery was more efficient when using slicer-deficient AGO forms instead of catalytically active wild-type counterparts [12].
2. On the day before the IP, do the following: prepare PK buffer, store at room temperature; wash mortars and pestles, store in the cold (4 °C) room together with pipettes and trash containers; prepare a box with 0.2 and 1 mL tips with wide orifice by cutting the top of the tip to allow efficient bead pipetting, store in the cold room; put clean bench paper at the working area in the cold room; prepare IP buffer without leupeptin, aprotinin, PMSF, and proteinase inhibitor tablets (these will be added right before starting the IP protocol), store in the cold room; and label 15 mL conical and 1.5/2 mL tubes for bead washing, nuclease treatment, IP, and protein and RNA extractions, store in the cold room.
3. Because bead sedimentation usually takes 10–20 min, proceed to the following step while beads sediment.
4. Two grams of *A. thaliana* inflorescences or 5 g of seedlings have been typically processed for AGO RIP-Seq analysis. One gram of *N. benthamiana* agroinfiltrated leaves has been processed for AGO RIP–RT-PCR analysis [12].
5. The addition of an excessive amount of liquid nitrogen may lead to sample freezing. Sample thawing will take several minutes.
6. The nuclease treatment enriches for AGO-protected fragments, as regions of target RNAs not being protected by the AGO will be degraded. Because MNase is active in cell extracts, the treatment is done in filtered homogenates to nuclease treat both input and IP fractions. The nuclease treatment can be omitted if AGO-bound targets are detected solely by RT-PCR (and not by high-throughput sequencing).
7. Pre-warm an aliquot of PK buffer to 65 °C, and add 50 U/mL of RNase inhibitor.

8. If reactions are incubated in a ThermoMixer (Eppendorf), shake at 750 rpm every 3 min for 15 s. If reactions are incubated in a water bath, mix reactions every 3 min by inverting the tubes several times, and transfer the tubes back to the bath.
9. Each input RNA, nd-input RNA, and IP RNA samples will have two tubes and thus two pellets to resuspend. For each class of sample, resuspend one pellet and use this volume to resuspend the second pellet.
10. Nuclease digestion can be verified by polyacrylamide gel electrophoresis or by loading samples in a RNA 6000 nanochip in a Bioanalyzer instrument (Fig. 2).
11. It is recommended to do sRNA libraries in parallel to analyze AGO-bound sRNAs [13].
12. Ponceau S staining is optional but recommended to check the efficiency of the protein transfer to the membrane. For Ponceau S staining, add enough Ponceau S solution to cover the transferred membrane and shake for 5 min at room temperature. Wash with H_2O to remove the excess of Ponceau S solution. Bands in red corresponding to more abundant proteins become visible. Remove excess H_2O from the membrane by air-drying for 5 min maximum. Scan the membrane and keep the image for total protein-loading control panel. This staining step does not interfere with antibody probing. The red stain is removed in the following step.
13. As shown in Fig. 3, Northern blot detection of miR172 in the IP fraction can be used to confirm sRNA co-IP with *A. thaliana* AGO1.
14. If doing the complete AGO RIP-Seq protocol, do not use more than 10–20% of the IP RNA sample for Northern blot checking, as the amount of IP RNA left may be insufficient to obtain enough amount of final DNA amplicon for high-throughput sequencing.
15. Oligonucleotide sequence is antisense to the sRNA that is analyzed. Start by using a DNA probe for sRNA detection as DNA oligonucleotides are cheap and work well most of the times. However, if sRNA detection with a DNA probe fails, then use an LNA probe. Order the LNA oligonucleotide with every other three nucleotides locked, including the first one (e.g., an LNA probe to detect *A. thaliana* miR172 is A + TGC + AGC + ATC + ATC + AAG + ATT + CT, where the + indicates the locked nucleotide).
16. If using an LNA probe, pre-hybridize and hybridize at ~20 °C below the calculated dissociation temperature (T_d) [T_d (°C) = 4(G + C) + 2(A + T)] for the corresponding ^{32}P–labeled oligonucleotide.

17. This works in the majority of cases. If there is still a background problem, proceed to an additional incubation of the membrane with 0.1× SSC/0.1% SDS for 60 min at 50 °C or with 0.1× SSC/1% SDS for 60 min at 50 °C.
18. It may be necessary to increase the number of PCR cycles to 35–40 to detect target RNAs in IP fractions.
19. Every time when using SPRI beads, be sure of mixing very well the bead stock container with a vortex, even if the bead suspension seems uniform. Be careful to not over-dry SPRI beads as this will compromise rehydration and elution. Mix well with vortex and increase the elution time from 2 to 5 min if beads remain as black chunks in the elution H_2O. Do not try to elute SPRI beads more than once as this will dilute the eluate concentration. One elution is sufficient as the elution efficiency with H_2O or EB is very high. It is usually better not to transfer to a new tube all the eluate, as some beads could be co-pipetted. In this case, it is more convenient to leave a small amount of eluate in the tube and transfer a totally clean bead-free eluate.
20. Do not waste IP RNA sample trying to measure its concentration as it is too low.
21. An efficient rRNA depletion is necessary to obtain a high-quality transcript library with low number of rRNA-derived sequences. rRNA depletion can be verified by agarose gel electrophoresis or by loading samples in a RNA 6000 nanochip in a Bioanalyzer instrument (Fig. 2). Various rRNA depletion kits were tested, with the Ribo-Zero treatment being the most effective.
22. In contrast to other RNA-Seq protocols, no fragmentation step was included because the average size of RNAs after Ribo-Zero treatment according to Bioanalyzer analysis (Fig. 2) was appropriate for high-throughput sequencing (peak around 200 nt).
23. This protocol is suitable for RNA amounts ranging from 0.5 to 100 ng.
24. RNA XP-based recovery of RNA fragments smaller than 100 bp is increased adding half volume of 100% ethanol. The addition of PEG 8000 also helps.
25. Deoxyuridine triphosphate (dUTP) incorporation in the second strand synthesis instead of deoxythymidine triphosphate (dTTP) enforces strand specificity.
26. Note that when the 12P XP beads are used, the beads are discarded while the supernatant is retained for subsequent applications.
27. To prepare a 20 μM Universal PE adapter mix (20 μM each), mix 20 μL of 50 μM PE adaptor 1, 20 μL of 50 μM PE adaptor

2, 5 μL of 10× hybridization buffer, and 5 μL of DEPC-treated H_2O; heat for 5 min at 75 °C; ramp down to 25 °C by decreasing 1.5 °C/min; hold for 30 min at 25 °C; and put on ice. Freeze at −20 °C. To check annealing, load 2.5 μL of annealed oligos (50 μM each, 500 ng each approximately) on a 4% MetaPhor Agarose gel. Make a 2-μM working solution right before use.

28. Multiple SPRI bead cleanups are done after ligation to remove primer dimers.
29. The number of PCR cycles is variable. For input samples, 14 cycles are usually enough. For nuclease-treated IP samples, additional cycles (16–18) may be necessary to amplify enough amount of DNA for high-throughput sequencing (2–10 nM). Do not use excessive number of PCR samples as this will lead to low-complexity amplicons.
30. Do not quantify with spectrophotometer as it leads to unevenness among samples.

Acknowledgments

I would like to thank James C. Carrington for intellectual and funding support, as well as members of his laboratory. This work was supported by an Individual Fellowship from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska Curie grant agreement No. 655841 to A.C.

References

1. Carbonell A, Carrington JC (2015) Antiviral roles of plant ARGONAUTES. Curr Opin Plant Biol 27:111–117. doi:10.1016/j.pbi.2015.06.013
2. Fang X, Qi Y (2016) RNAi in plants: an Argonaute-centered view. Plant Cell 28(2): 272–285. doi:10.1105/tpc.15.00920
3. Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13(7):350–358. doi:10.1016/j.tplants.2008.04.007
4. Mallory A, Vaucheret H (2010) Form, function, and regulation of ARGONAUTE proteins. Plant Cell 22(12):3879–3889. doi:10.1105/tpc.110.080671
5. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133(1):128–141. doi:10.1016/j.cell.2008.02.033
6. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1): 116–127. doi:10.1016/j.cell.2008.02.034
7. Zhang X, Niu D, Carbonell A, Wang A, Lee A, Tun V, Wang Z, Carrington JC, Chang CE, Jin H (2014) ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. Nat Commun 5:5468. doi:10.1038/ncomms6468
8. Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X (2011) Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell 145(2):242–256. doi:10.1016/j.cell.2011.03.024
9. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol 49(4):493–500. doi:10.1093/pcp/pcn043

10. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC (2010) Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat Struct Mol Biol 17(8):997–1003. doi:10.1038/nsmb.1866
11. Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y (2009) Rice MicroRNA effector complexes and targets. Plant Cell 21(11):3421–3435. doi:10.1105/tpc.109.070938
12. Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC (2012) Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. Plant Cell 24(9):3613–3629. doi:10.1105/tpc.112.099945
13. Gilbert KB, Fahlgren N, Kasschau KD, Chapman EJ, Carrington JC, Carbonell A (2014) Preparation of multiplexed small RNA libraries from plants. Bio Protoc 4(21):e1275
14. Wang L, Si Y, Dedow LK, Shao Y, Liu P, Brutnell TP (2011) A low-cost library construction protocol and data analysis pipeline for Illumina-based strand-specific multiplex RNA-seq. PLoS One 6(10):e26426. doi:10.1371/journal.pone.0026426

Chapter 7

Identification of ARGONAUTE/Small RNA Cleavage Sites by Degradome Sequencing

Ivett Baksa and György Szittya

Abstract

The method described here enables the high-throughput identification of cleaved mRNA targets of ARGONAUTE/small RNA complexes. The protocol is based on a modified 5′-rapid amplification of cDNA ends combined with deep sequencing of 3′ cleavage products of mRNAs. Poly(A) RNA is purified from the tissue of interest which is followed by a 5′-RNA adapter ligation. The ligated products are then reverse transcribed, amplified, and digested with *Mme*I. After gel separation, a 3′ double-stranded DNA adapter is ligated to the fragments, which are then amplified and index labeled for the high-throughput sequencing of pooled degradome libraries. Sequencing datasets from pooled libraries can be analyzed with different bioinformatic approaches.

Key words ARGONAUTE, mRNA degradome, PARE, miRNA target, DNA library construction, Next-generation sequencing, Poly(A) mRNA, *Mme*I digestion, PAGE purification

1 Introduction

Small RNAs (sRNAs) are important gene expression regulators involved in multiple developmental processes and also in stress responses. These molecules are generally 21–24 nt long in their mature form and are produced by several overlapping but genetically separable biochemical pathways. In plants sRNAs are very diverse, and based on their biogenesis, we distinguish two main classes: microRNAs (miRNAs) and small interfering RNAs (siRNAs). MiRNAs function in a posttranscriptional manner by downregulating target mRNAs in a variety of cellular prosesses [1, 2]. Plant miRNAs have near-perfect complementarity to their target sites, and they guide the cleavage of target mRNAs [3]. However, there are examples where the translation of the mRNA is suppressed in the absence of cleavage [4–6]. MiRNAs are mainly 21 nt long and are important regulators of endogenous gene expression as shown by the pleiotropic developmental abnormalities seen in

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_7,

multiple miRNA biogenesis mutants. SiRNAs are produced from long double-stranded (ds) RNA (dsRNA) precursors and can be categorized into several different groups, based on the origin of the dsRNA and on the type of their target(s). One particular group of siRNAs regulates protein-coding genes through annealing to complementary sequence stretches on mRNAs and is called phased, secondary, small interfering RNAs (phasiRNAs) [7]. phasiRNAs are 21 nt long, and they belong to a special class of endogenous siRNAs. phasiRNAs can regulate their targets by mRNA cleavage in a similar way as miRNAs do [7].

Several bioinformatic approaches have been developed to predict miRNA targets in silico. However, these methods often lead to false predictions [8, 9]. For in vivo validation of miRNA-mediated cleavage, the commonly used 5′-rapid amplification of cDNA ends method can be an option [10]. However, this technique has several limitations. For example, it requires the knowledge of the sequence of the target mRNA, and multiple targets cannot be detected in the same experiment. With these disadvantages and the availability of high-throughput sequencing, a genome-wide experimental sRNA-cleaved target identification and validation technique has been developed and named PARE from **P**arallel **A**nalysis of **R**NA **E**nds [11]. Here we describe a modified version of the original PARE method used in our laboratory [12].

The protocol starts with a poly(A)-containing mRNA purification from a total RNA pool, which also contains the 3′ cleavage products of the sRNA-targeted mRNAs. As a result of the sRNA cleavage, the 3′ poly(A) mRNAs have a ligation competent 5′ monophosphate group instead of a 5′-cap. The following ligation step selects between the ligation competent 3′ cleavage products and full-length mRNAs that cannot be ligated to the 5′-RNA adapter. The used 5′-adapter is an RNA oligonucleotide, which contains the *Mme*I type IIS restriction enzyme recognition site on it. After the 5′-RNA adapter ligation, the transcripts are reverse transcribed with a 3′-adapter oligo (dT) primer. The resulting cDNA is then amplified through a few PCR cycles before the *Mme*I digestion. The *Mme*I enzyme cuts 20 nucleotides, 3′ of its recognition site, creating 20–21-nt long products with 2-nt 3′-overhang region. In the following steps, a 3′-dsDNA adapter is ligated to the template, which is followed by PCR amplification. If multiple PARE libraries (e.g., originated from different biological samples) are required to be sequenced in a parallel manner (multiplex-sequencing reaction), unique adaptor sequences are required to be incorporated into the oligos. So during this amplification step, the libraries can be labeled with a unique adaptor sequence for multiplex sequencing. The high-throughput sequencing datasets are analyzed by bioinformatic approaches. Most commonly used bioinformatic tools for degradome data processing are PAREsnip (available version 2.3) from the UEA small RNA workbench [13],

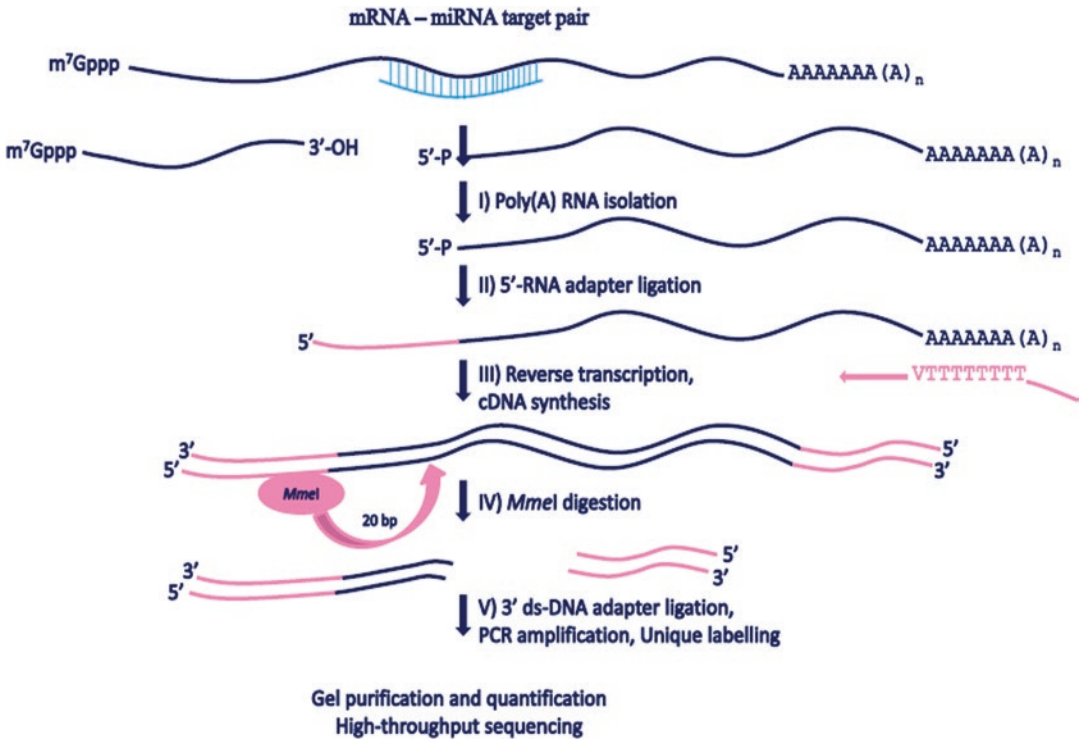


Fig. 1 A schematic representation of degradome library preparation. After the sequence-specific miRNA-induced RNA cleavage, the 3′ cleavage product of the mRNA contains a 5′-monophosphate and a poly(A) tail. The method starts with poly(A) selection and followed by a 5′-RNA adapter ligation. The next step is reverse transcription to make the first strand of cDNA with a 3′-adapter oligo (dT) primer. A few PCR cycle amplifications of the template and a control step are carried out before the *Mme*I digestion. This type IIS restriction enzyme cleaves 20 nucleotides, 3′ away from the recognition site. The cleaved fragment is ligated to a 3′-dsDNA adapter, and a PCR amplification with an index primer labeling is done. The PAGE-purified product is ready for deep sequencing

CleaveLand4 from the Axtell Lab [14], and the sPARTA from the Meyers Laboratory [15]. A schematic representation of the method is shown in Fig. 1.

2 Materials

2.1 Reagents

1. 10× EB (pH 9.5): 1 M glycine, 100 mM ethylenediaminetetraacetic acid (EDTA), and 1 M NaCl. Sterilize and protect from light.
2. Extraction buffer: prepare a 10 mL solution by mixing 1 mL 10× EB, 2 mL 10% SDS, and 7 mL sterile water.
3. 0.45 μm Spin-X CA centrifugal tube filters (Corning Costar).
4. 10% (w/v) Ammonium persulfate (APS) buffer.

5. 10× S-adenosylmethionine (SAM) (500 μM) (New England Biolabs).
6. 10× Tris–Borate–EDTA (TBE) buffer: 1 M Tris base, 1 M boric acid, and 0.02 M EDTA. Prepare with RNAse-free H_2O.
7. 20 bp DNA ladder and 100 bp Plus DNA ladder (Thermo Fisher Scientific).
8. 3 M Sodium acetate (NaAc).
9. STE buffer: 10 mM Tris pH 8.0, 50 mM sodium chloride NaCl, and 1 mM EDTA.
10. 3′-adapter (dT) primer 5′-CGAGCACAGAATTAATACGACTTTTTTTTTTTTTTTTTTV-3′.
11. 3′-adapter primer 5′-CGAGCACAGAATTAATACGACT-3′.
12. 3′-dsDNA adapter:

    top 5′–/5Phos/TGGAATTCTCGGGTGCCAAGG-3′.

    bottom 5′-CCTTGGCACCCGAGAATTCCANN-3′.
13. 5′-adapter primer 5′-GTTCAGAGTTCTACAGTCCGAC-3′.
14. 5′-RNA adapter:

    5′-rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrC-3′.
15. 6× DNA-loading buffer containing bromophenol blue and xylene cyanol.
16. Acrylamide/bis-acrylamide solution, 19:1, 40% (Bio-Rad).
17. 10% polyacrilamide gel: mix 0.6 mL of 10× TBE, 1.5 mL of 40% acrylamide/bis-acrylamide solution (19:1), 3.9 mL of nuclease-free water, 36 μL of 10% APS, and 2.4 μL of TEMED. Mix gently and pour.
18. 12% polyacrilamide gel: mix 0.6 mL of 10× TBE, 1.8 mL of 40% acrylamide/bis-acrylamide solution (19:1), 3.6 mL of nuclease-free water, 36 μL of 10% APS, and 2.4 μL of TEMED. Mix gently and pour.
19. Agarose.
20. Antarctic phosphatase (5 U/μL) (New England Biolabs).
21. Chloroform/isoamyl alcohol (24:1).
22. dNTPs.
23. Dynabeads mRNA DIRECT Purification Kit (Thermo Fisher Scientific).
24. Ethanol.
25. Ethidium bromide 10 mg/mL solution.
26. GlycoBlue (Thermo Fisher Scientific).
27. LigaFast™ Rapid DNA Ligation System (Promega).

28. Low-molecular weight (LMW) DNA Ladder (New England Biolabs).
29. Poly(A) Purist MAG Kit (Thermo Fisher Scientific).
30. Microcon Centrifugal Filters (30 k) (Merck Millipore).
31. *Mme*I (2 U/μL) (New England Biolabs).
32. Nuclease-free water.
33. PCR primer 5′-AATGATACGGCGACCACCGAGATCTAC ACGTTCAGAGTTCTACAGTCCGA-3.
34. Phenol/chloroform/isoamyl alcohol (25:24:1) pH 8.0.
35. Phusion High-Fidelity DNA Polymerase (2 U/μL) (Thermo Fisher Scientific).
36. RevertAid H Minus Reverse Transcriptase (200 U/μL) (Thermo Fisher Scientific).
37. RNaseOUT (40 U/μL) (Thermo Fisher Scientific).
38. Sequencing primer 5′-TACACGTTCAGAGTTCTACAGTC CGAC-3′.
39. T4 RNA ligase (5 U/μL) (Thermo Fisher Scientific).
40. Tetramethylethylene diamine (TEMED).

2.2 Equipment

1. 4 °C microfuge.
2. Agarose electrophoresis unit.
3. Benchtop microfuge.
4. Filtered tips.
5. Heat block and/or incubator.
6. Laboratory tube mixer.
7. Magnetic separation rack.
8. Nuclease-free tubes.
9. PCR machine.
10. Pestle and mortar.
11. Polyacrilamide electrophoresis unit.
12. Sterile blades.
13. Sterile needles.

3 Methods

3.1 RNA Extraction and Poly(A) Enrichment

1. Homogenize 2–5 g of tissue of interest in a pestle and mortar using liquid nitrogen and transfer to a 15-mL conical tube.
2. Extract RNA using TRIzol following the manufacturer's instructions (*see* **Note 1**).

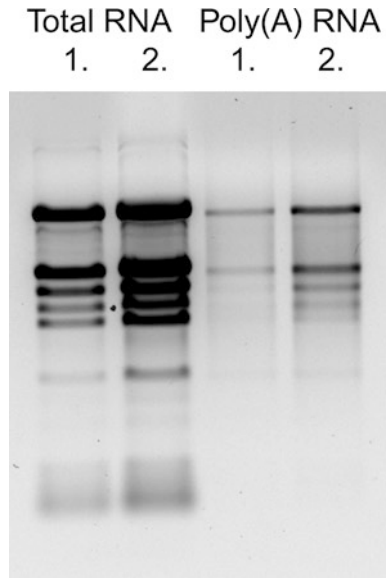


Fig. 2 Verification of the poly(A) mRNA purification by checking the integrity and quality of the RNA on a 1.2% agarose gel. For degradome library preparation, we used leaf tissues of *Nicotiana benthamiana* plants. The isolation of the poly(A) fraction reduces ribosomal RNA content, and mRNAs appear as a smear resulting from the heterogeneous population of differently sized mRNAs. Total RNA is used as control

3. Use 100–400 μg total RNA in 250 μL nuclease-free water to isolate poly(A) RNA using poly(A) Purist MAG Kit following the manufacturer's instructions (*see* **Note 2**). For the precipitation step, the use of GlycoBlue (stained glycogen) is recommended to easily detect your RNA pellet.
4. Dissolve the poly(A) RNA pellet in 17 μL nuclease-free water. Check the RNA integrity with a Bioanalyzer instrument or by 1.2% agarose gel electrophoresis (Fig. 2).

3.2 Ligation of the 5′-RNA Adapter

1. Add the following components into a nuclease-free tube on ice: 1 μg of poly(A) RNA (from Subheading 3.1, **step 4**), 1 μL of 200 μM 5′-RNA adapter, 2 μL of 10× RNA ligase buffer, 2 μL of T4 RNA ligase (5 U/μL), and nuclease-free water to a final volume of 20 μL.
2. Mix gently and incubate at 37 °C for 1 h.
3. Add 280 μL nuclease-free water to the reaction, extract once with 300 μL of phenol/chloroform/isoamyl alcohol (25:24:1), and extract once with 300 μL of chloroform/isoamyl alcohol (24:1).
4. For precipitation add 1 μL of GlycoBlue to the supernatant, 20 μL of 3 M NaAc, and 1 mL of ice-cold 100% ethanol. Centrifuge at 4 °C for 20 min at 18,400 × *g* in a microfuge.

5. Wash the pellet two times with 1 mL of 70% ethanol, and gently discard the supernatant.
6. Vacuum dry and resuspend the pellet in 20 μL of nuclease-free water.
7. Purify your poly(A) mRNA with Dynabeads mRNA DIRECT Purification Kit to remove unligated adapters following manufacturer's instructions. After preparing Dynabeads Oligo $(dT)_{25}$, follow the Direct mRNA Purification Kit protocol using 250 μL of beads. Elute the mRNA in 15 μL of 10 mM Tris–HCl.

3.3 Reverse Transcription

1. Prepare the following reaction in a nuclease-free tube on ice: 15 μL of 5′-RNA adapter-ligated mRNA (from Subheading 3.2, **step 7**), 1.5 μL of 50 μM 3′-adapter (dT) primer, 2.5 μL of dNTP's (10 μM each), and 13.5 μL of nuclease-free water. Incubate at 65 °C for 5 min followed by a 2-min incubation at room temperature. Gently spin down.
2. Add 10 μL of 5× RT reaction buffer, 2,5 μL of nuclease-free water, 2 μL of RNaseOUT (40 U/μL), and 3 μL of RevertAid H Minus Reverse Transcriptase (200 U/μL). Gently mix using pipette and incubate at 42 °C for 3 h and then at 70 °C for 15 min to heat and inactivate the enzymes. Place the tube on ice for 2 min and centrifuge gently for a few seconds.
3. Add 250 μL of nuclease-free water to the RT reaction, and repeat steps Subheading 3.2, **steps 3–5**.
4. Vacuum dry and resuspend the pellet in 7 μL of nuclease-free water (*see* **Note 3**).

3.4 PCR Amplification

1. Prepare a PCR master mix for two identical reactions by scaling the volumes listed below to the desired number of samples (*see* **Note 4**): 32 μL of nuclease-free water, 10 μL of 5× HF buffer, 0.5 μL of dNTPs (12.5 μM each), 2 μL of 10 μM 5′-adapter primer, 2 μL of 10 μM 3′-adapter primer, 0.5 μL of Phusion High-Fidelity DNA Polymerase (2 U/μL).
2. Transfer 47 μL of the PCR master mix to a PCR tube and add 3 μL of cDNA (from Subheading 3.3, **step 4**). Set up the following PCR program: 98 °C for 30 s, followed by 7 cycles of 98 °C for 20 s, 60 °C for 30 s, 72 °C for 3 min, and a final amplification at 72 °C for 5 min. Prepare a control reaction with the remaining 1 μL of cDNA template (*see* **Note 5**) (Fig. 3).
3. Combine the tubes from seven cycle reactions, add 200 μL of nuclease-free water and repeat steps from Subheading 3.2, **steps 3–5**.
4. Vacuum dry and resuspend the pellet in 200 μL of nuclease-free water. Clean the product by using Microcon Centrifugal Filters (30 k) according to the manufacturer's instructions (*see* **Note 6**).

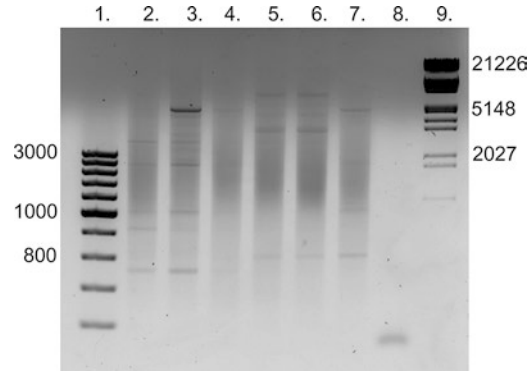


Fig. 3 Control step of the first PCR amplification (35 cycles). Detection of the control PCR products on 1.2% agarose gel. Lane 1: 100 bp plus DNA ladder. Lanes 2–7: different plant tissue samples from the control PCR reactions. Lane 8: PCR master mix control reaction without cDNA template. Lane 9: Lambda DNA/*EcoR*I + *Hind*III Marker

3.5 MmeI Digestion

1. Add the following components to a nuclease-free microfuge tube: 5 μg of PCR product (from Subheading 3.4, **step 4**), 4.5 μL of 10× CutSmart buffer, 10× of SAM (S-adenosylmethionine) (500 μM) (*see* **Note 7**), and 2.5 μL of *Mme*I (2 U/μL) and nuclease-free water to a final volume of 45 μL. Incubate the mixture at 37 °C for 2 h with occasional rotation (*see* **Note 8**).
2. Add 2 μL of Antarctic phosphatase (5 U/μL) and 5 μL of 10× Antarctic phosphatase reaction buffer and incubate at 37 °C for 1 h.

3.6 PAGE Purification I

1. Prepare a native 12% polyacrilamide gel (*see* **Note 9**) according to the recipe (6 mL is enough for one gel) (*see* **Note 10**).
2. Wash the wells with 1× TBE running buffer. Pre-run the gel at 100 V at 30 min and then wash the wells again with TBE running buffer before loading the samples. Add 6× DNA loading buffer to the dephosphorylated *Mme*I-digested samples from Subheading 3.5, **step 2**, and load them into two lanes of the polyacrilamide gel with equal volume. Load two lanes at the sides with LMW DNA Ladder and 20 bp DNA Ladder (*see* **Note 11**).
3. Run the gel at 50 V until all the samples have entered the gel, and then raise the voltage to 100 V. To get a good separation, let the bromophenol blue dye run out of the gel. The xylene cyanol dye needs to migrate to the middle of the PAGE gel. Prepare an ethidium bromide staining solution in a clean container using around 50 mL of 1× TBE and 2 μL of ethidium bromide solution (*see* **Note 12**). Take apart the glass plates and put the gel gently in the staining solution for 2–3 min.

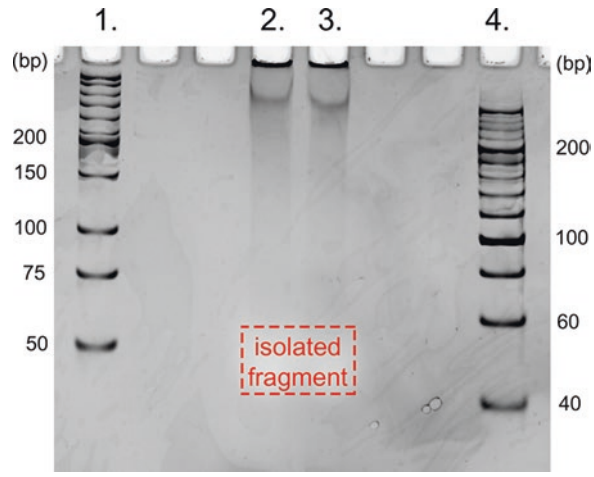


Fig. 4 PAGE purification of the *Mme*I-digested product. Lane 1: LMW DNA Ladder. Lanes 2–3: *Mme*I-digested product. Lane 4: 20 bp DNA ladder. The *red box* indicates the isolated gel piece, which overlaps with the xylene cyanol dye

4. Cut out the gel bands with a clean razor blade that (42–50 nt) contains the *Mme*I-cleaved fragments using the DNA Ladders as a marker (*see* **Note 13**) (Fig. 4).
5. For extracting the DNA from the gel, follow the method undermentioned. Prepare a sterile, nuclease-free, 0.5 mL microtube by puncturing the bottom of the tube three to four times with a sterile 21-gauge needle. Place the tube into a sterile, round-bottom, nuclease-free, 2 mL microtube. Transfer the gel slice into the 0.5-mL RNase-free microtube and centrifuge the stacked tubes at 16,000 × *g* for 1 min to shred the gel through the holes into the 2 mL tube. Discard the 0.5 mL microtube, add 0.5 mL of 0.3 M NaCl to the gel pieces, and elute the DNA overnight at 4 °C by rotating the tube (*see* **Note 14**).
6. Transfer the eluates and the gel debris to the top of a 0.45-μm Spin-X CA filter. Centrifuge the filter at 3,300 × *g* for 2 min in a microfuge, add 100 μL of 0.3 M NaCl to the top of the membrane, and repeat the centrifugal step.
7. Precipitate DNA with adding 1 μL of glycoblue and 600 μL of isopropanol (equal volume) to the sample, and incubate at −70 °C for at least 3 h (overnight is the best).
8. Centrifuge the samples at 13,000 rpm for 20 min at 4 °C in a microfuge. Carefully remove supernatant and wash the pellet two times with 70% ethanol. Vacuum dry the pellet, and resuspend the DNA in total of 7-μL nuclease-free water.

3.7 Ligation of the 3′-dsDNA Adapter

Before the ligation step, it is necessary to make dsDNA from the single-stranded top and bottom oligonucleotides (*see* **Note 15**).

1. Resuspend the oligonucleotides in STE buffer at a high concentration (1–10 OD260 U/100 uL) as the presence of some salt is necessary for the oligos to hybridize.
2. Mix equal volumes of both complementary oligos (at equimolar concentration) in a 1.5 mL microfuge tube. If you do not use equal molar amounts, there will be single-stranded material leftover.
3. Place tube in a standard heat block or water bath at 90–95 °C for 3–5 min. To avoid hairpin formation, gradually cool down the oligos (1 °C/min). The easiest way to achieve this is to remove the heat block from the apparatus or unplug the water bath and allow it to cool down to room temperature on the workbench.
4. Slow cooling to room temperature should take 45–60 min.
5. For short term store the dsDNA oligo on ice or at 4 °C until ready to use. For long-term storage, store at −20 °C.
6. Add the following reagents to the gel-purified digested product (from Subheading 3.6, **step 8**): 10 μL of 2× rapid ligation buffer, 1 μL of 3′-dsDNA adapter (6 μM), and 2 μL of T4 DNA ligase (2 U/μL).
7. Incubate the ligation reaction at room temperature for 2 h.

3.8 PAGE Purification II

1. Prepare a native 12% polyacrilamide gel.
2. Wash the wells with 1xTBE running buffer before pre-running the gel at 100 V to 30 min, and then wash it again before loading the sample to the gel. Add 6× DNA loading buffer to the ligation reaction from Subheading 3.8, **step 1**, and load two lanes of the polyacrilamide gel with ~12 μL of the ligated samples. Also load two lanes at the sides with both LMW DNA Ladder and 20 bp DNA Ladder.
3. Run the gel at 50 V until all of the samples have entered the gel, and then raise the voltage to 100 V. Run the gel until the bromophenol blue runs out of the gel, and the xylene cyanol is just above the bottom of the gel. Stain the gel in ethidium bromide solution.
4. Slice the appropriate gel bands (~63 nt) (*see* **Note 16**), and crush them (*see* Subheading 3.6, **steps 5–7**) (Fig. 5).
5. Resuspend the pellet in 8 μL of nuclease-free water.

3.9 PCR Amplification of DNA Library

Set up the following 50 μL PCR reaction: 3 μL of ligated product from Subheading 3.8, **step 5**, 10 μL of 5× Phusion HF buffer, 1 μL of 25 μM PCR primer, 1 μL of Index primer (Illumina) (*see* **Note 17**), 1 μL of dNTP mix (12.5 μM each), 0.5 μL of Phusion

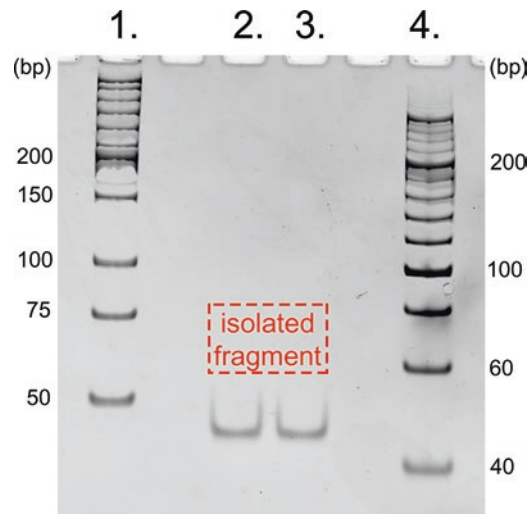


Fig. 5 PAGE purification of the 3′-ds DNA adapter-ligated product. Lane 1, LMW DNA Ladder, 2–3, 3′-ds DNA adapter-ligated product; 4, 20 bp DNA ladder. The *red box* indicates the isolated gel piece

DNA polymerase, and 33.5 µL of nuclease-free water in a final volume of 50 µL. Use the following PCR conditions: 98 °C for 30 s, followed by 22 cycles of 98 °C for 20 s, 60 °C for 30 s, 72 °C for 15 s, and a final amplification at 72 °C for 3 min (*see* **Note 18**).

3.10 PAGE Purification III

1. Prepare a 10% native polyacrylamide gel to visualize the PCR products.
2. Add 6× DNA loading buffer to the PCR products from Subheading 3.9, **step 1**, and load three lanes of the polyacrilamide gel with equal volumes of the PCR products. Also load two lanes at the sides with LMW DNA Ladder and 20 bp DNA Ladder in wells with at least one space in between ladder and samples.
3. Repeat the running and staining procedure (*see* Subheading 3.8, **step 3**).
4. Cut out the gel bands corresponding to aprox. 135 nt that contain the degradome library, and proceed (*see* Subheading 3.6, **steps 5**–7) (Fig. 6).
5. Resuspend the pellet in 12 µL of nuclease-free water.
6. Check the quality and the quantity of the degradome library. Several degradome libraries can be visualized on one gel before multiplexing. Run a 12% native polyacrylamide gel (*see* Subheading 3.6, **step 1**). Load 2 µl sample of each library on the gel (from Subheading 3.10, **step 5**) with 6× DNA loading buffer and use LMW DNA Ladder and 20 bp DNA Ladder (*see* **Note 19**). For quantification of the products, use a

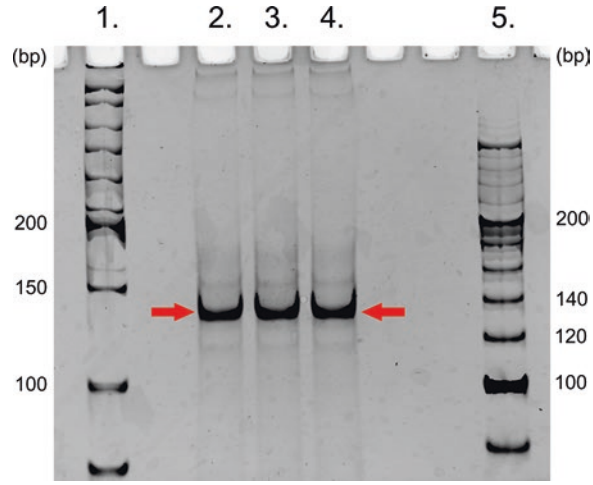


Fig. 6 PAGE purification of the PCR-amplified and index-labeled degradome library. Lane 1: LMW DNA Ladder. Lanes 2–4: final PCR product. Lane 5: 20 bp DNA ladder. The *red arrows* indicate the isolated bands

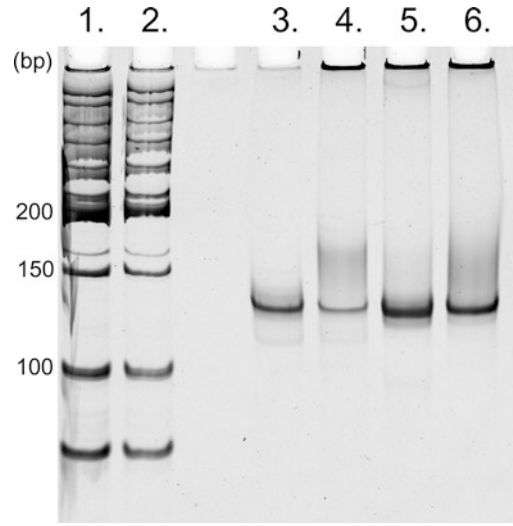


Fig. 7 Quality and quantity control of the final degradome libraries. Lanes 1–2: LMW DNA Ladder dilution series (150 bp: 16.5 ng in lane 1, 8.25 ng in lane 2). Lanes 3–6: different index-labeled degradome libraries

Bioanalyzer instrument or compare with a dilution series of a DNA Ladder (Fig. 7).

7. After quantifying each degradome library, prepare a sequencing sample by measuring the same amount from each together (*see* **Note 20**). Check this mixture before sequencing (*see* Subheading 3.10, **step 7**) on a 12% polyacrylamide gel (Fig. 8).

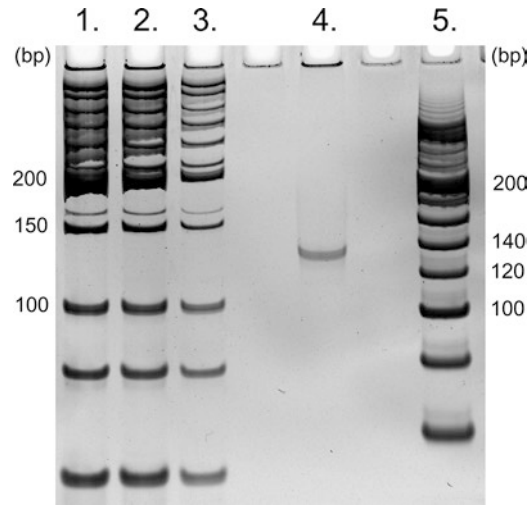


Fig. 8 Verification of multiplexed degradome libraries. Lanes 1–3: LMW DNA Ladder dilution series. Lane 4: sequencing sample (4 index-labeled degradome libraries were pooled together). Lane 5: 20 bp DNA ladder

8. When ordering the sequencing service, be sure that the proper sequencing primer is used (5′-TACACGTTCA GAGTTCTACAGTCCGAC-3′) [16]. It differs from the original Illumina sequencing primer.

3.11 Bioinformatic Analysis

To analyze the data of the degradome libraries, first remove the adapter sequences and then follow the steps of the used bioinformatic tools. There are several excellent bioinformatic tools available to analyze degradome sequencing such as PAREsnip [13], CleaveLand4 [14], or sPARTA [15].

4 Notes

1. Make sure the used RNA isolation method is suitable for the protocol (e.g., elution buffer, RNA size, etc.). To prepare a good quality degradome library for next-generation sequencing, a high-quality RNA sample is needed. Keep in mind that poor sample preparation results in poor quality data. Since poor results often become apparent only after expensive and time-consuming sequencing, it is important to pay extra attention to every step from the preparation of the RNA samples to the library preparation as well.
2. This isolation step will enrich the poly(A) mRNA of the sample and helps to reduce rRNAs contamination. As an initial step, it is recommended to use as much total RNA as possible according to the manufacturer's protocol, and start the ligation step with a high-quality poly(A) RNA.

3. This is an optional stopping point: cDNA can be stored for several days at −20 °C for later use.
4. Include 10% overage to cover pipetting errors. Also prepare water controls by adding nuclease-free water in place of the template.
5. Run a parallel 35 cycles PCR with 1 μL template as a control step. The product of this reaction should be seen as a smear in a 1.2% agarose gel. In case of contamination or in the absence of any visible product in this step, start over the protocol from Subheading 3.2, **step 1**.
6. This step is necessary for concentrating and desalting the DNA. It also helps to remove primers from amplified DNA. It is an optional stopping point.
7. 10× CutSmart buffer and 10× SAM are supplied with *Mme*I.
8. If the amount of cDNA from Subheading 3.4, **step 5** is lower than 5 μg, adjust the stoichiometric ratio between *Mme*I enzyme and substrate, since excess of *Mme*I blocks cleavage.
9. We use the Bio-Rad Mini–PROTEAN Tetra handcast systems with 1.0 mm spacers.
10. If multiple libraries are prepared simultaneously, it is recommended, for every PAGE purification step, to run each sample on separate gels to avoid contamination among samples. After 30 min of polymerization, pre-run the gel in 1× TBE buffer.
11. Using two different types of markers will help to detect the correct product size. It is critical in this step to have appropriate DNA size markers because no product band will be visible on the gel. It is recommended to leave at least one empty lane between the ladder and the sample to isolate the product properly. Dye migration in 12% polyacrylamide non-denaturing gel: bromophenol blue indicates dsDNA around 20 bp and xylene cyanol runs around with 70 bp dsDNA.
12. Handle ethidium bromide with caution: it is a potent mutagen. Always wear gloves. This dye is replaceable with other types of DNA dyes.
13. To avoid contamination, put a piece of Saran wrap between the gel and the UV device when cutting out the bands. Among these conditions, the expected product size partly overlaps with the xylene cyanol dye.
14. If the gel pieces did not go through the holes of the 0.5 mL tube, make further holes at the bottom and/or centrifuge for another 2–3 min. Elution time could take at least 6 h at room temperature.
15. Alternatively, request the oligonucleotide annealing of this service during the ordering process.

16. At this step, there is a visible band between 40 and 50 nt corresponding to unligated primers. Cut out the "empty" gel slice corresponding to 60 to 70 nt.
17. We use Illumina adapters from TruSeq Small RNA Library Preparation Kits. Using unique identifiers, like Illumina Indexes allow us to pool different samples together in a single run. Multi-sample pooling is a cost-effective method. The Illumina HiSeq 2000 platform allows to pool up to eight degradome libraries together in one sequencing lane.
18. The number of cycles is variable and depends on the tissue type or the organism. An additional optimalization step could be carried out using the remaining template from Subheading 3.8, **step 5**.
19. In case of contamination (adapter or other) or poor quality libraries, repeat the protocol from the final PCR amplification (*see* Subheading 3.9, **step 1**). In case this does not solve the problem, start over from an earlier step (*see* Subheading 3.2, **step 1** or Subheading 3.5, **step 1**).
20. For multiplexed sequencing, use different index primers for each sample. Check if they are compatible with each other at Illumina website.

Acknowledgments

We are thankful for Jeannette Bálint and Tibor Csorba for critically reading the manuscript. This work was supported by Hungarian National Research, Development and Innovation Office grants K101793 and K106170.

References

1. Rogers K, Chen X (2013) Biogenesis, turnover, and mode of action of plant microRNAs. Plant Cell 25(7):2383–2399. doi:10.1105/tpc.113.113159
2. Sun G (2012) MicroRNAs and their diverse functions in plants. Plant Mol Biol 80(1):17–36. doi:10.1007/s11103-011-9817-6
3. Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science 297(5589):2053–2056. doi:10.1126/science.1076311
4. Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 15(11):2730–2741. doi:10.1105/tpc.016238
5. Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science 303(5666):2022–2025. doi:10.1126/science.1088060
6. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190. doi:10.1126/science.1159151
7. Fei Q, Xia R, Meyers BC (2013) Phased, secondary, small interfering RNAs in post-transcriptional regulatory networks. Plant Cell 25(7):2400–2415. doi:10.1105/tpc.113.114652
8. Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids

Res 39(Web Server issue):W155–W159. doi:10.1093/nar/gkr319

9. Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res 18(10): 1602–1609. doi:10.1101/gr.080127.108
10. Scotto-Lavino E, Du G, Frohman MA (2006) 5′ end cDNA amplification using classic RACE. Nat Protoc 1(6):2555–2562. doi:10.1038/nprot.2006.480
11. German MA, Luo S, Schroth G, Meyers BC, Green PJ (2009) Construction of parallel analysis of RNA ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. Nat Protoc 4(3):356–362. doi:10.1038/nprot.2009.8
12. Baksa I, Nagy T, Barta E, Havelda Z, Varallyay E, Silhavy D, Burgyan J, Szittya G (2015) Identification of Nicotiana benthamiana microRNAs and their targets using high throughput sequencing and degradome analysis. BMC Genomics 16(1):1025. doi:10.1186/s12864-015-2209-6
13. Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittya G, Dalmay T, Moulton V (2012) PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. Nucleic Acids Res 40(13):e103. doi:10.1093/nar/gks277
14. Addo-Quaye C, Miller W, Axtell MJ (2009) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25(1):130–131. doi:10.1093/bioinformatics/btn604
15. Kakrana A, Hammond R, Patel P, Nakano M, Meyers BC (2014) sPARTA: a parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. Nucleic Acids Res 42(18):e139. doi:10.1093/nar/gku693
16. Pantaleo V, Szittya G, Moxon S, Miozzi L, Moulton V, Dalmay T, Burgyan J (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. Plant J 62(6):960–976. doi:10.1111/j.1365-313X.2010.04208.x

Chapter 8

Dissecting the Subnuclear Localization Patterns of Argonaute Proteins and Other Components of the RNA-Directed DNA Methylation Pathway in Plants

Cheng-Guo Duan and Jian-Kang Zhu

Abstract

RNA-directed DNA methylation (RdDM) is a nuclear pathway which is comprised of multiple main and accessory protein components, including two plant-specific DNA-dependent RNA polymerases, Pol IV and Pol V, and argonaute (AGO) proteins. Regulation in the RdDM pathway can be achieved via multiple mechanisms, including the spatial distribution of different RdDM components. Here we describe a protocol for dissecting the subnuclear localization of AGO proteins and other RdDM components, including nuclei extraction from seedlings, slide preparation, and subsequent immunostaining.

Key words RdDM, Argonaute, Immunostaining, Subnuclear localization, Microscopy, DNA methylation

1 Introduction

RNA-directed DNA methylation (RdDM) is a conserved epigenetic silencing pathway in plants [1]. RdDM plays significant roles in diverse processes including transposon silencing, pathogen defense, stress response, and genome stability [2–4]. In the RdDM pathway, two plant-specific RNA polymerases, Pol IV and Pol V, and many other proteins are required [2–6]. Among these other proteins, *Argonaute* (AGO) proteins directly bind small RNAs to target genomic regions with high sequence complementarity to the guide small RNA [7, 8]. In plants, the AGO family contains multiple members. For example, *Arabidopsis thaliana* encodes ten AGOs which are classified into different clades according to their phylogenetic origin. AGO4, AGO6, and AGO9 are thought to be involved in the RdDM pathway [9–11]. RdDM takes place in the nucleus and is regulated through the spatial segregation of different components. For example, AGO4 and AGO6 are co-localized with

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_8,

different RNA polymerases in different nuclear sub-compartments [12, 13]. Determining the subnuclear co-localization patterns of the different components is critical for understanding the complex regulation of the RdDM pathway in plants.

Here we describe an immunostaining protocol for dissecting the co-localization relationship of RdDM components in the nucleus. We use AGO (AGO4 and AGO6) proteins and the largest subunits of RNA polymerase (NRPD1 and NRPE1) proteins to test their co-localization patterns.

2 Materials

2.1 Plant Materials and Growth Conditions

The plants used in this study include *A. thaliana* Col-0, Flag-AGO4, Flag-AGO6, NRPD1-Flag, and NRPE1-Flag [8, 13], which are plants expressing the tagged proteins under their respective native promoters in their respective mutant backgrounds. All plants were grown on 1% sucrose containing 1/2 Murashige and Skoog (MS) medium under a long-day photoperiod (16 h light/8 h dark). Two-week-old seedlings were collected for nuclei extraction.

2.2 Solutions, Antibodies, and Other Materials

1. Nuclear extraction buffer 1 (NEB1): 10 mM Tris–HCl (pH 9.5), 10 mM potassium chloride (KCl), 500 mM sucrose, 4 mM spermidine, 10 mM spermine, 0.1% Triton-X 100, 8% formaldehyde, and 0.1% 2-mercaptoethanol. Make aliquots and store at 4 °C (*see* **Note 1**).
2. NEB2: 10 mM Tris–HCl (pH 9.5), 10 mM KCl, 1.7 M sucrose, 4 mM spermine, 10 mM spermine, 0.1% Triton-X 100, 8% formaldehyde, and 0.1% 2-mercaptoethanol. Make aliquots and store at 4 °C (*see* **Note 1**).
3. 10× PBST stock solution: 1.28 M sodium chloride (NaCl), 20 mM KCl, 80 mM disodium phosphate (Na_2HPO_4), 20 mM monopotassium phosphate (KH_2PO_4), and 0.1% Triton-X 100. Prepare 10× stock in water and autoclave.
4. Blocking solution: prepare 1% bovine serum albumin (BSA) in 1× PBST. Prepare it fresh and store at 4 °C.
5. DAPI: ProLong® Gold Antifade Mountant with DAPI (Thermo Fisher Scientific).
6. Nylon mesh: 20, 53, and 100 mm (Spectra/Mesh).
7. Razor blades.
8. Petri dishes.
9. Nail polish.
10. Glass slides (VWR Superfrost Plus) and cover glass.
11. Primary antibody: anti-Flag monoclonal antibody generated in mouse (Sigma–Aldrich), anti-AGO4 polyclonal antibody gen-

erated in rabbit (Agrisera), and anti-AGO6 polyclonal antibody generated in rabbit (Agrisera) (*see* **Note 2**).

12. Secondary antibody: Alexa Fluor® 488-conjugated goat anti-rabbit IgG (H + L) secondary antibody (Thermo Fisher Scientific) and Alexa Fluor® 568-conjugated donkey anti-mouse IgG (H + L) secondary antibody (Thermo Fisher Scientific) (*see* **Note 2**).

3 Methods

3.1 Nuclei Extraction and Slide Preparation

Figure 1 shows a diagram for the main experimental steps regarding nuclei extraction and slide preparation (*see* **Note 3**). Centrifugations are done in a microcentrifuge.

1. Plant growth: seeds are sterilized and sown on 1/2 MS medium. After 2-day vernalization at 4 °C, seeds are grown on a chamber with photoperiod of 16 h light/8 h dark at 23 °C for 2 weeks.
2. Homogenization: collect about 2 g of seedlings from 1/2 MS medium and transfer seedlings to petri dish on ice (*see* **Note 4**). Add 2 mL of NEB1 pre-chilled at 4 °C into the samples and chop seedlings with a razor blade to get a fine homogenate.
3. Fixing: add equal volume of 8% formaldehyde (prepared in NEB1 buffer) to the homogenate. Use wide-bore pipette tips to mix the homogenate gently. Cover the petri dish and leave it at 4 °C for 30–60 min to fix the homogenate.
4. Filtration: use wide-bore pipette tips to transfer the fixed homogenate onto one sheet of 100 μm nylon filter. Sequentially filter the homogenate through the 53 μm and 20 μm filters. Split the homogenate into 1.5 mL tubes.
5. Centrifuge the homogenate for 3 min at 2500 (600 × *g*) at 4 °C. Discard the supernatant and gently resuspend the pellet in 300 μL of NEB1. Layer 300 μL of resuspended pellet on the top of the 300 μL NEB2 in a 1.5-mL tube (*see* **Note 5**).
6. Centrifuge the assembly at top speed (10,000–13,000 rpm, 10,000–16,000 × *g*) at 4 °C for 1 h to pellet the nuclei.
7. Carefully remove the two layers of supernatant and gently resuspend the nuclei pellet in 40 μL of NEB1 to obtain the nuclei preparation.
8. Pipette 3–5 μL of nuclei onto a slide and mix with 3 μL of DAPI. Cover with a coverslip to check the quality of nuclei under a microscope.
9. Pipette 3 μL of nuclei onto the slide and spread into a thin film of 20 × 20 mm square by a piece of coverslip (*see* **Note 6**).

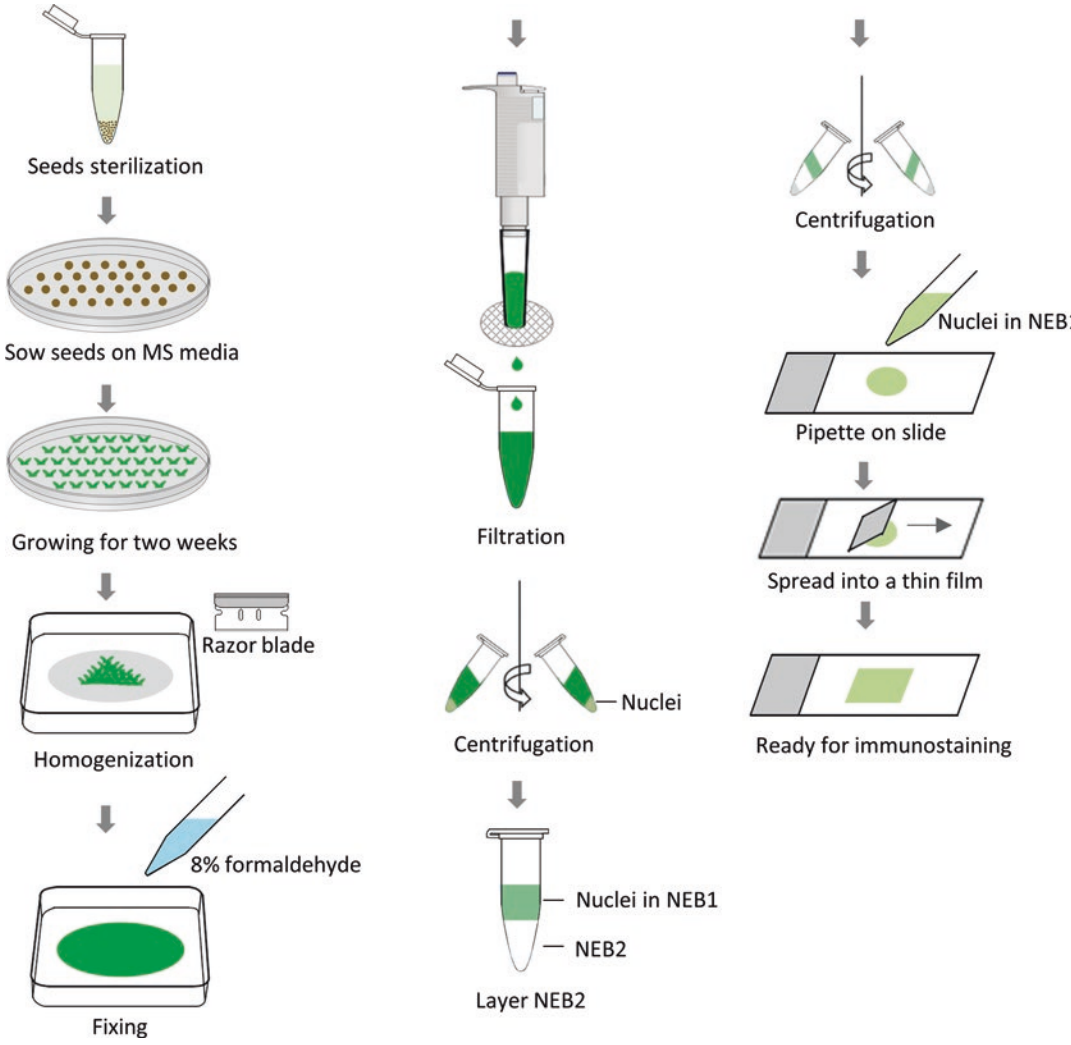


Fig. 1. A workflow for nuclei extraction and slide preparation. Diagram showing experimental steps regarding nuclei extraction and slide preparation to analyze the subnuclear localization of AGOs and other components of the *A. thaliana* RdDM pathway

10. Air-dry the slide at room temperature and store the dried slide on the cardboard slide storage book at 4 °C.

3.2 Subnuclear Localization of AGO4 and AGO6

1. Use *PAGO4::Flag-AGO4* and *PAGO6::Flag-AGO6* seedlings [8] to prepare the slides.
2. Put air-dried slides into a square plate. Cover the area of nuclei with 200 μL of 4% formaldehyde (prepared in PBST) to refix the nuclei for 30 min on the bench. Wash the slide with 1× PBST for 3–5 min.

3. Create a humid chamber by placing several layers of wet paper towel on the bottom of a square plate. Place the slide on the top of the wet towels and drop 200 μL of blocking solution to the nuclei area. Place a piece of plastic sheet on top of the blocking solution. Incubate the covered plate at 37 °C for 30 min.
4. Wash slides with 1× PBST for 3–5 min in a new square plate.
5. Transfer the slides to a new humid chamber. Drop 100 μL of primary antibody in blocking solution (1:200 for anti-Flag) to the nuclei area and place a piece of plastic sheet on top of the solution. Keep the covered plate on the bench in the dark overnight.
6. Wash slides with 1× PBST for 3–5 min in a new square plate.
7. Transfer the slides to a new humid chamber. Drop 200 μL of blocking solution and place a piece of plastic sheet on top of the solution. Incubate the covered plate at 37 °C for 30 min.
8. Rinse the slides with 1× PBST briefly. Drop 100 μL of Alexa Fluor® 568-conjugated donkey anti-mouse secondary antibody (1:400) in 1× PBST to nuclei area and place a piece of plastic sheet on top of the PBST solution. Put the covered plate inside a 37 °C incubator for 2 h.
9. Wash the slides with 1× PBST for 3–5 min in an empty square plate.
10. Drop 5 μL of ProLong® Gold Antifade Mountant with DAPI in the center of the nuclei area and apply a coverslip. Seal the slide with clear nail polish for microscope observation.

3.3 Subnuclear Colocalization of AGO4 and AGO6 Together with NRPD1 and NRPE1

1. Use *PNRPD1::NRPD1-Flag* and *PNRPE1::NRPE1-Flag* seedlings [13] to prepare the slides. Immunostain the slides first with anti-Flag primary antibody (1:200 dilution) and then label them with Alexa Fluor® 568-conjugated donkey anti-mouse secondary antibody (1:200 dilution) according to the procedure described above. Perform the second immunolabeling as follows.
2. Apply the blocking solution to the slides (*see* **step 3** in Subheading 3.2).
3. Wash the slides (*see* **step 4** in Subheading 3.2).
4. Incubate the slides with second primary antibody (*see* **step 5** in Subheading 3.2). Use anti-AGO4/anti-AGO6 as second primary antibodies (1:200 dilution for anti-AGO4 and 1:100 dilution for anti-AGO6).
5. Wash the slides (*see* **step 6** in Subheading 3.2).
6. Apply the blocking solution to the slides (*see* **step 7** in Subheading 3.2).

7. Incubate the slides with secondary antibody (*see* **step 8** in Subheading 3.2). Use Alexa Fluor® 488-conjugated goat anti-rabbit secondary antibody.
8. Wash the slides (*see* **step 9** in Subheading 3.2).
9. Drop 6 μL of ProLong® Gold Antifade Mountant with DAPI in the center of nuclei area, and cover with a coverslip. Seal the slide with clear nail polish for observation under a microscope.

4 Notes

1. Prepare NEB1 and NEB2 fresh every time needed. Ideally, all reagents and solutions should be freshly prepared before the experiment.
2. The dilution ratio for each primary and secondary antibody has to be optimized empirically.
3. All the steps during nuclei extraction are performed on ice.
4. Avoid bringing agar with roots when transferring seedlings from MS medium to petri dishes. The presence of excessive agar in seedlings will result in bad quality nuclei.
5. For nuclei extraction, if the nuclei pellet looks still green after density gradient extraction with NEB2, additional washes with NEB1 are recommended.
6. To maintain the film thin while spreading the nuclei on the slide via coverslip, you only need to do it once toward one orientation for each slide (Fig. 1).

Acknowledgments

This work was supported by the Chinese Academy of Sciences and by National Institutes of Health Grant R01GM070795 (to J.–K. Z).

References

1. Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10(2):94–108. doi:10.1038/nrg2504
2. Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15(6):394–408. doi:10.1038/nrg3683
3. Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11(3):204–220. doi:10.1038/nrg2719
4. Zhang H, Zhu JK (2011) RNA-directed DNA methylation. Curr Opin Plant Biol 14(2):142–147. doi:10.1016/j.pbi.2011.02.003
5. Pikaard CS, Haag JR, Ream T, Wierzbicki AT (2008) Roles of RNA polymerase IV in gene silencing. Trends Plant Sci 13(7):390–397. doi:10.1016/j.tplants.2008.04.008

6. Wierzbicki AT, Haag JR, Pikaard CS (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell 135(4):635–648. doi:10.1016/j.cell.2008.09.035
7. Mallory A, Vaucheret H (2010) Form, function, and regulation of ARGONAUTE proteins. Plant Cell 22(12):3879–3889. doi:10.1105/tpc.110.080671
8. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC (2010) The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22(2):321–334. doi:10.1105/tpc.109.072199
9. Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299(5607):716–719. doi:10.1126/science.1079695
10. Zheng X, Zhu J, Kapoor A, Zhu JK (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J 26(6):1691–1701. doi:10.1038/sj.emboj.7601603
11. Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature 464(7288):628–632. doi:10.1038/nature08828
12. Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell 126(1):93–106. doi:10.1016/j.cell.2006.05.032
13. Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS (2006) The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell 126(1):79–92. doi:10.1016/j.cell.2006.05.031

Chapter 9

Detection of Argonaute 1 Association with Polysomes in *Arabidopsis thaliana*

Cécile Lecampion, Elodie Lanet, and Christophe Robaglia

Abstract

Argonaute (AGO) proteins play a key role in RNA silencing mechanisms. RNA silencing affects both RNA degradation and translation. The characterization of translation-associated RNA silencing mechanisms and components often requires polysome isolation and analysis. In this chapter, we describe the identification of AGO1 association with polysomes through polysome fractionation on sucrose gradient, preparation of proteins by filtration and concentration, and immunoblotting.

Key words RNA silencing, Argonaute, Polysomes, Immunoblotting, Sucrose gradients, *Arabidopsis thaliana*, MicroRNA

1 Introduction

RNA silencing is an important mechanism for the regulation of posttranscriptional eukaryotic gene expression. Small RNAs bound to argonaute (AGO) proteins form a silencing effector complex (called RISC for RNA-induced silencing complex) that targets messenger RNA (mRNA) for cleavage or translational repression [1]. In animals, where poor complementarity usually exists between microRNAs (miRNAs) and mRNA targets, RISC binding mostly leads to translational repression but also to mRNA deadenylation and decay [2]. In plants, where complementarity is high between miRNA and mRNA, RISC binding often leads to target RNA cleavage and degradation. However, translational repression also occurs [3–5]. The precise mechanisms leading to translation repression are not completely elucidated. The dissociation of eukaryotic initiation factor 4F (eIF4F) translation initiation complex and of poly(A)-binding protein (PABP) from translated mRNA has been observed in animals [2]. In plants, translational repression was found associated

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_9,

with endoplasmic reticulum, and miRNA and AGO were found largely associated with polysomes [6].

The analysis of translational repression mechanisms often requires the characterization of defined miRNAs or proteins in association with polysomes. Polysomes are separated using sucrose gradient centrifugation, and fractions to be analyzed always contain high sucrose concentration that requires specific treatment for further processing. In this chapter, we will describe a method developed to identify AGO association with polysomes. We anticipate that this method can be applied to other proteins for which antibodies are available.

Detection of AGO1 association with polysomes relies on three main steps: polysome fractionation [7], protein extraction from the collected fraction [6], and immunodetection. The polysome fractionation method described here is a simple method that only requires equipment commonly found in laboratories. It relies on the separation on sucrose gradient of the mRNA associated with ribosomes. Sucrose is then removed from the fraction and macromolecules are concentrated. Finally, proteins are resolved by SDS–PAGE gel electrophoresis, and AGO1 proteins are detected by immunoblotting using an anti-AGO1 antibody.

2 Materials

Use RNase/DNase-free reagents to prepare solutions (*see* **Note 1**).

2.1 Polysome Fractionation

1. 10× Salt solution: 400 mM Tris–HCl pH 8.4, 200 mM potassium chloride (KCl), and 100 mM magnesium chloride ($MgCl_2$). To prepare a 100 mL solution, weight 4.85 g Tris–HCl and add 1.49 g of KCl and 2.03 g of $MgCl_2$ hexahydrate. Add 80 mL ultrapure water and adjust the pH to 8.4 with KOH. Add water to a volume of 100 mL. Autoclave the solution.
2. Sucrose solution: 2 M Sucrose in 1× salt solution. To prepare a 200 mL solution, weight 137 g of sucrose and dissolve in 1× salt solution.
3. 4× Polysome buffer: salt solution, 5.26 mM EGTA, 0.5% octylphenoxy poly(ethyleneoxy)ethanol, branched, 50 µg/mL cycloheximide, 50 µg/mL chloramphenicol, and 300 µg/mL heparin. For a 10 mL solution, mix 2.5 mL of 10× salt solution, 20 mg EGTA, 50 µL octylphenoxy poly(ethyleneoxy) ethanol, branched, 10 µL cycloheximide (50 mg/mL), 10 µL chloramphenicol (50 mg/mL), and 300 µL heparin (10 mg/mL). Adjust to 10 mL.
4. Ultracentrifuge tube, thin wall, polypropylene, 13.2 mL.

2.2 Protein Extraction

1. Vivaspin 20 10 kDa ultrafiltration unit (Sartorius).
2. 2× Laemmli buffer: 120 mM Tris–HCl pH 6.8, 4% SDS, 20% glycerol, 0.1% bromophenol blue, and 0.2% β-mercaptoethanol.

2.3 SDS–PAGE Gel Electrophoresis and Immunoblotting

1. Stacking gel: 5% acrylamide/bis-acrylamide 37.5:1, 125 mM Tris–HCl pH 6.8, 0.1% SDS, 0.1% ammonium persulfate, and 0.05% TEMED.
2. Running gel: 10–15% acrylamide/bis-acrylamide 37.5:1, 375 mM Tris–HCl pH 8.8, 0.1% SDS, 0.1% ammonium persulfate, and 0.05% TEMED.
3. TGS buffer: 12 mM Tris–HCl pH 8.3, 96 mM glycine, and 0.1% SDS.
4. Coomassie blue solution: 10% acetic acid, 25% ethanol, and 0–125% Coomassie blue.
5. TGS/ethanol buffer: TGS buffer with 20% ethanol.
6. TBST: 50 mM Tris–HCl pH 7.5, 150 mM NaCl, and 0.1% Tween 20.
7. Milk powder.

2.4 Antibodies

1. Primary antibody: anti-AGO1 against AGO1 peptide MVRKRRTDAPSC [8] (other anti-AGO1 antibodies are available at Agrisera #AS09 527–100).
2. Secondary antibody: goat anti-rabbit IgG (H + L) horseradish peroxidase conjugated (Promega).

3 Methods

To avoid any RNA degradation, keep all samples at 4 °C and work in RNase/DNase-free conditions with RNase/DNase-free solutions.

3.1 Polysome Fractionation

1. Gradient making: gradients are made of four layers of sucrose solution of different concentrations (50%, 35%, and two layers of 20%). Prepare the solutions by diluting 2 M sucrose solution in 1× salt solution as described in Table 1.

   Pour the layer in a 13.2 mL ultracentrifuge tube according to Table 2. Freeze each layer in a −40 °C or a −80 °C freezer before adding the next one. This step will prevent from mixing or disturbing the layers (*see* **Note 2**). On the day of the experiment, remove the necessary number of gradients (*see* **Note 3**) from the freezer (*see* **Note 4**) and, if not done before, add the last 20% layer and let the gradient thaw in the cold room.
2. Prepare the polysome buffer.

Table 1
Sucrose solution for gradient layers (for six gradients)

Final sucrose concentration (%)	Sucrose 2 M (mL)	1× Salt solution (mL)	Final vol. (mL)
50	8.8	3.2	12
35	12.9	12.1	25
20	7.4	17.6	25
20	5.8	14.2	20

Table 2
Volume of each gradient layer

Sucrose layer	50%	35%	20%	20%
Vol. (mL)	1.85	3.65	3.65	1.35

3. Use 6-day-old *Arabidopsis thaliana* seedlings, grown on ½ Murashige and Skoog [9] medium supplemented with 1% sucrose. Harvest the seedlings, quickly freeze them in liquid nitrogen, and grind them thoroughly. To 300 mg of the resulting powder, add 2.4 mL of precooled polysome buffer. Homogenize with a pipette tip and transfer to 1.5 mL tubes. Centrifuge at 16,000 × *g* for 15 min at 4 °C in a microcentrifuge to pellet debris (*see* **Note 5**).
4. Carefully load the supernatant on the top of a gradient (one per 1.5 mL tube) without disturbing the surface. Transfer to precooled buckets and ultracentrifuge at 175,000 × *g* for 2 h and 45 min at 4 °C.
5. Gradients are collected from the bottom to the top, using a capillary tube connected to a UV cuvette (path length: 1 mm). Set the fraction collector speed to collect 2 mL fractions. Read the absorbance continuously at 260 nm (Fig. 1).

3.2 Protein Purification

Six identical gradients are run in parallel and separated into six fractions. All the corresponding fractions are pooled before concentration.

1. Proteins from each of the six pools are loaded on the ultrafiltration unit to remove sucrose and to concentrate the sample. The final volume is adjusted to 20 mL with water (*see* **Note 6**).
2. Samples are centrifuged three times at 4000 × *g* for 30 min at 4 °C in a swinging bucket.

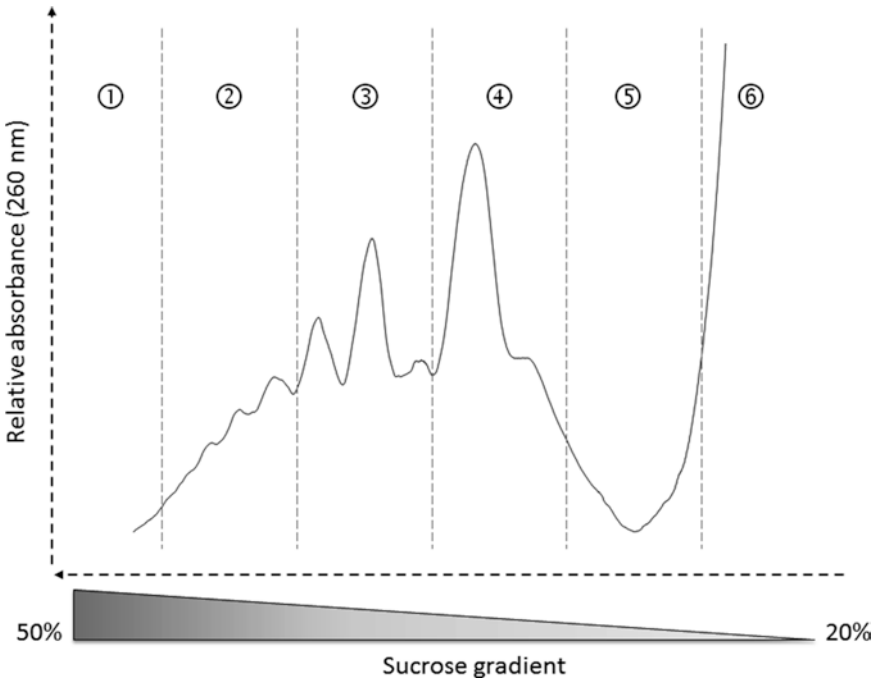


Fig. 1 Polysome profiling of 6-day-old *A. thaliana* seedlings. The first part of the profile corresponds to the polysomal fraction (areas 1–3). The main peak (area 4) corresponds to the monosome and the last part (areas 5–6) corresponds to the non-polysomal fractions

3. 300 µL of the concentrated sample is recovered at the bottom of the concentrate pocket. The volume is adjusted to 500 µL with water (*see* **Note 7**).
4. Proteins are again concentrated using a vacuum concentrator for 3 h with no heat (*see* **Note 8**).
5. The final volume of 50 µL is collected, mixed with 2× Laemmli buffer, and heated for 5 min at 99 °C.

3.3 SDS–PAGE Gel Electrophoresis and Immunoblotting

1. Prepare running and stacking gel and cast two mini gels (8.6 × 6.8 cm).
2. Load 10 µL of protein.
3. Run the gel in TGS buffer at 60 V in the stacking gel, and then increase to 100 V until the bromophenol blue exits the gel.
4. Stain one gel with Coomassie blue solution.
5. Transfer the second gel to a nitrocellulose membrane in TGS/ethanol buffer (*see* **Note 9**) for 1 h at 100 V.
6. Block the nitrocellulose membrane overnight in TBST plus 5% milk powder at 4 °C.
7. Incubate membrane with anti-AGO1 antibody (1/8000) in TBST plus 5% milk powder for 2 h at 4 °C. Rinse the membrane twice with TBST.

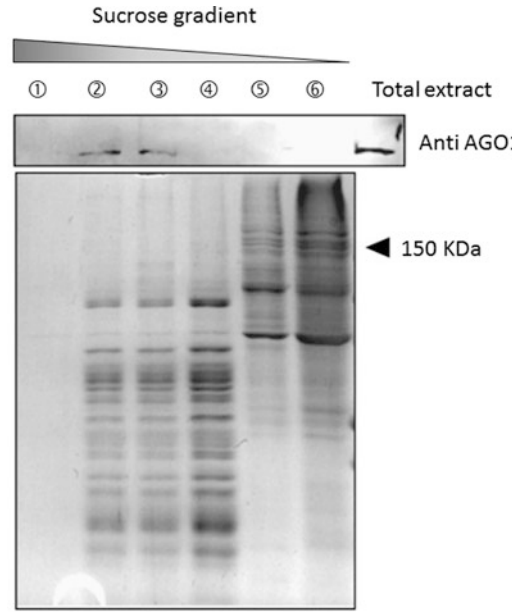


Fig. 2 AGO1 is associated with polysomes. Proteins from the six fractions were resolved by SDS–PAGE gel electrophoresis (low panel is the *Coomassie blue staining*) and transferred to a nitrocellulose membrane for immunoblotting with an anti-AGO1 antibody (copyright by the American Society of Plant Biologists)

8. Incubate the membrane with the secondary antibody (1/2000) in TBST plus 5% milk powder at 4 °C for 1 h.
9. Develop using a ECL Western blotting substrate as per supplier specifications (Fig. 2).

4 Notes

1. After sterilization, all solutions can be kept at room temperature for months.
2. The last 20% of the layer can be added either in advance (in this case freeze it like the other layers) or extemporaneously.
3. To get a good signal in the Western blot, we recommend to prepare six gradients per sample and to pool each corresponding fraction. For protein purification, each of the six fractions is made of 12 mL (6 × 2 mL).
4. Frozen gradients can be kept for at least 6 months in a −40 °C or −80 °C freezer.
5. All those steps have to be performed quickly to prevent samples from warming.

6. Polysomal fractions contain high sucrose concentrations. This dilution is helpful to allow the solution to cross the concentrator membrane.
7. At this step, samples can be kept at −20 °C.
8. Using heat during vacuum concentration would favor evaporation of the buffer but may damage proteins.
9. For this step, we use ethanol (instead of methanol that is commonly used for electroblotting) because it is as efficient as methanol but does not have toxic effects.

References

1. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297. doi:10.1016/S0092-8674(04)00045-5
2. Iwakawa HO, Tomari Y (2015) The functions of MicroRNAs: mRNA decay and translational repression. Trends Cell Biol 25(11):651–665. doi:10.1016/j.tcb.2015.07.011
3. Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. Science 301(5631):336–338. doi:10.1126/science.1085242
4. Pillai RS, Bhattacharyya SN, Filipowicz W (2007) Repression of protein synthesis by miRNAs: how many mechanisms? Trends Cell Biol 17(3):118–126. doi:10.1016/j.tcb.2006.12.007
5. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190. doi:10.1126/science.1159151
6. Lanet E, Delannoy E, Sormani R, Floris M, Brodersen P, Crete P, Voinnet O, Robaglia C (2009) Biochemical evidence for translational repression by Arabidopsis microRNAs. Plant Cell 21(6):1762–1768. doi:10.1105/tpc.108.063412
7. Lecampion C, Floris M, Fantino JR, Robaglia C, Laloi C (2016) An easy method for plant Polysome profiling. J Vis Exp 114. doi:10.3791/54231
8. Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell 19(3):421–428. doi:10.1016/j.molcel.2005.06.014
9. Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15(3). doi:10.1111/j.1399-3054.1962.tb08052.x

Chapter 10

Functional Analysis of Arabidopsis ARGONAUTEs in Meiosis and DNA Repair

Marina Martinez-Garcia and Mónica Pradillo

Abstract

Plant ARGONAUTE (AGO) proteins regulate a wide range of cellular and developmental functions. Recent findings highlight their role during homologous recombination, a basic mechanism to repair double-strand DNA lesions (in somatic cells) and programmed DNA breaks (in meiocytes). This chapter contains an exhaustive description of procedures applied to analyze meiotic chromosome behavior (cytogenetic techniques) and DNA repair capacity (genotoxicity assays) in AGO-deficient *Arabidopsis thaliana* mutants.

Key words Arabidopsis, ARGONAUTEs, Small RNAs, Chromosomes, DNA repair, FISH, Genotoxic agents, Homologous recombination, Meiosis

1 Introduction

ARGONAUTEs (AGOs) are key players in all known small RNA-directed pathways [1]. The genome of *Arabidopsis thaliana* (Arabidopsis) contains ten *AGO* genes, designated *AGO1* to *AGO10*, belonging to three functional groups: RNA slicers, small RNA binders, and chromatin modifiers [2]. Mutants corresponding to these genes have been extensively analyzed through different approaches [3].

Here we focus on cytological methodologies to analyze the chromosome dynamics in pollen mother cells (PMCs) from *ago* mutants, since several of these mutants present abnormal meiosis [4, 5]. This chapter describes the basic preparation of male meiotic chromosomes by spreading to analyze meiosis progression, chromatin condensation, and chromosome segregation after 4,6-diamidino-2-phenylindole (DAPI) staining. Furthermore, the protocol includes localization of rDNA repetitive sequences (5S and 45S) by fluorescence in situ hybridization (FISH) to distinguish each chromosome. During the first meiotic prophase I, homologous chromosomes are physically connected forming stable

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_10,

bivalents. These connections, named chiasmata, represent the cytological manifestations of reciprocal exchange events (crossovers, COs). FISH can be applied in order to determine chiasma frequency per chromosome and per cell [6–8]. Different groups have different preferences for cytological versus genetic assessment of meiotic recombination. Both procedures have their advantages, and both are broadly accepted in the literature. Although chiasma scoring is thought to provide a slight underestimation of the actual number of COs, it truly fits with values obtained by genetic recombination measurements [6, 7], meiotic tetrads analyses [9], whole genome sequencing [10], or "pollen typing" [11].

Here we also include a way to trace the epigenetic modification 5-methylcytosine (5-mC) (addressed by immunolocalization) at microscopic level. This DNA modification is present in pericentromeric heterochromatic regions in plant chromosomes [4, 5, 12], and its pattern distribution is altered in *ago4–1* [4].

Finally, to define the function of AGO proteins in DNA repair as well as in meiotic homologous recombination (HR), we provide several methods to check the sensitivity of *ago* mutants to different DNA damaging agents: cisplatin [cis–diamminedichloroplatinum (II), CDDP], gamma rays, mitomycin C (MMC), and UV-C. These agents generate a wide range of DNA lesions including interstrand and intrastrand cross-links, single- and double-strand breaks (SSBs and DSBs, respectively), dipyrimidine photoproducts, abasic sites, and oxidative products. HR, nucleotide excision repair (NER), and base excision repair (BER) are the main mechanisms involved in removing these lesions [13].

Analyses conducted by the methodologies detailed here are routinely applied in the characterization of mutants defective in meiosis and HR [14, 15]. However, they are not commonly used in studies related to mutants affected in small RNA-directed pathways. This chapter defines our current protocols, which have contributed to reveal the influence of AGOs in meiosis progression and DNA repair.

2 Materials

2.1 Plant Material

Obtain *Arabidopsis* seeds from collections of T–DNA insertional mutants, such as those from the Salk Institute Genomic Analysis Laboratory (SIGnAL, http://signal.salk.edu/cgi–bin/tdnaexpress) [16] (*see* **Note 1**). Be aware of the genetic background of the plants to use an appropriate control.

2.2 Collection of Inflorescences and Chromosome Preparation by Spreading

1. Fresh ice-cold Carnoy's fixative: 60% ethanol, 30% chloroform, and 10% glacial acetic acid.
2. Fresh 3:1 fixative: 75% ethanol and 25% glacial acetic acid.
3. Citrate buffer: prepare a stock solution (10×) with 40 mL of 0.1 M sodium citrate and 60 mL of 0.1 M citric acid, pH 4.6 (*see* **Note 2**).

4. Enzyme mixture: 0.3% (w/v) cellulase, 0.3% (w/v) pectolyase, and 0.3% (w/v) cytohelicase in citrate buffer (*see* **Note 3**).
5. 60% acetic acid: prepare a dilution with distilled water. Store at 4 °C.
6. Vectashield antifade mounting medium with 1 μg/mL 6-diamidino-2-phenylindole (DAPI). Store at 4 °C.
7. Moist chamber for enzyme digestion.
8. Microscope slides and coverslips.
9. Watch glasses, glass capillary tubes, dissecting needles, and fine forceps.
10. Heating block (45 °C).
11. Incubator (37 °C).
12. Stereo microscope.
13. Light microscope with phase contrast.
14. Fluorescence microscope and image acquisition software.
15. Adobe Photoshop software.

2.3 Probes and Probe Labeling Using Nick Translation Mix

1. Probes: 45S rDNA (plasmid pTa71) [17] and 5S rDNA (plasmid pCT4.2) [18] (*see* **Note 4**).
2. Nucleotides: dATP, dCTP, dGTP, dTTP, biotin-dUTP, and digoxigenin-dUTP.
3. Nick translation mix (DNA Polymerase I and DNase I).
4. Ethylenediaminetetraacetic acid (EDTA) 0.5 M, pH 8.
5. Eppendorf tubes.
6. Thermocycler.

2.4 Fluorescence In Situ Hybridization

1. 20× SSC (saline–sodium citrate) buffer: 0.3 M sodium chloride (NaCl), 0.03 M sodium citrate, pH 7. Store at room temperature. Prepare a working dilution (2×) with distilled water.
2. Fresh pepsin solution: 0.01% (w/v) in 0.01 M hydrogen chloride (HCl). Preheat 100 mL in a Coplin jar at 37 °C before use.
3. 4% paraformaldehyde: prepare from powder by stirring for around 30 min on a heating plate (~60 °C) under a fume hood. Add five drops of sodium hydroxide (NaOH) 1 M in 100 mL of hot distilled water. Adjust pH to 8. Subsequent filtration through filter paper is suitable. Store at 4 °C (*see* **Note 5**).
4. Ethanol series (70%, 90%, 100%).
5. Hybridization mix: 1 g dextran sulfate, 5 mL deionized formamide, and 1 mL 20× SSC. Dissolve at 65 °C; cool and adjust pH to 7. Aliquot and store at −20 °C.
6. 50% deionized formamide—2× SSC: deionize formamide with resin beads, and mix with 20× SSC to get a final concentration

of 2× SSC (150 mL of deionized formamide, 30 mL of 20× SSC, 120 mL of distilled water). Preheat at 45 °C three Coplin jars with 100 mL.

7. 2× SSC: prepare from 20× SSC (*see* **item 1** in Subheading 2.3), and preheat at 45 °C one Coplin jar with 100 mL.
8. Fresh 4T: 4× SSC and 0.05% (v/v) Tween 20. Preheat at 45 °C one Coplin jar with 100 mL.
9. Antibodies: anti-digoxigenin or streptavidin (for biotin-labeled probes) conjugated with FITC and Cy3, respectively. Dilute at the concentration of 5 ng/μL in TNB buffer. This buffer is made from 100 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5% (w/v) Boehringer blocking reagent. TNB can be stored at −20 °C.
10. Vectashield antifade mounting medium with 1 μg/mL DAPI. Store at 4 °C.
11. Coverslips and parafilm.
12. Eppendorf tubes.
13. Coplin jars.
14. Moist chamber.
15. Incubator (37 °C, 60 °C).
16. Heating block (72 °C, 80 °C).
17. Thermocycler.
18. Fluorescence microscope with optical filters for DAPI, FITC, and Cy3 fluorochromes and image acquisition software.
19. Adobe Photoshop software.

2.5 5-Methylcytosine Immunolocalization

1. 1% paraformaldehyde (*see* **item 3** in Subheading 2.4).
2. Ethanol series (70%, 90%, 100%).
3. HB50 buffer: 50% deionized formamide, 2× SSC, 50 mM sodium phosphate, pH 7.0.
4. 2× SSC: prepare from 20× SSC (*see* **item 1** in Subheading 2.4), and cool on ice one Coplin jar with 100 mL.
5. 10× phosphate-buffered saline (PBS): solution A (16.02 g $Na_2HPO_4{\cdot}2H_2O$ + 73.84 g NaCl in 900 mL of distilled water) + solution B (2.76 g $NaH_2PO_4H_2O$ + 16.56 g NaCl in 100 mL of distilled water). Adjust pH to 7 and autoclave. Dilute at 1:10 with distilled water before use. Store at room temperature.
6. Blocking buffer: dissolve 1% bovine serum albumin (BSA) and 0.1% (v/v) Triton X–100 in 1× PBS.
7. TNT buffer: 100 mM Tris–HCl pH 7.5, 150 mM NaCl, and 0.05% (v/v) Tween 20.
8. Primary antibody: anti-5-methylcytosine rose in mouse (1:50) diluted in blocking buffer.

9. Secondary antibody: rabbit anti-mouse antibody FITC (1:50) diluted in blocking buffer.
10. Vectashield antifade mounting medium with 1 μg/mL DAPI. Store at 4 °C.
11. Coverslips and parafilm.
12. Coplin jars.
13. Moist chamber.
14. Incubator (37 °C).
15. Heating block (60 °C, 80 °C).
16. Fluorescence microscope and image acquisition software.
17. Adobe Photoshop software.

2.6 Media Plates and Genotoxic Agents

1. Germination media (GM): Murashige and Skoog salt mixture (1×) in 1% sucrose. About 1% agarose can be added for solid media. Additionally, add vitamins: 100 mg/L inositol, 1 mg/L thiamine, 0.5 mg pyridoxine, 0.5 mg/L nicotinic acid, and 0.5 g/L methyl ester sulfonate (MES). Adjust pH to 5.7 and autoclave.
2. CDDP at 5, 30, 50, and 75 μM (*see* **Note 6**). It is recommended to add cisplatin after sterilization of media.
3. Gamma rays from a ^{137}Cs source. Apply a range from 50 to 500 Gy.
4. MMC at 3, 6, 9, and 12 μg/mL (*see* **Note 6**). It is recommended to add MMC after sterilization of media.
5. UV–C irradiation with power to achieve 300 J/m^2.
6. Petri plates (standard and square), 24-well plates, glass capillary tubes, rubber bulbs, and micropore tape.
7. Sowing template grids (Fig. 1).
8. Laboratory oven.
9. Laboratory balance scale.
10. High-resolution photographic camera.
11. Image J software.

3 Methods

3.1 Plant Growth

1. For cytological analyses, seeds should be sown in pots containing a soil mixture of vermiculite and commercial soil (3:1) and grown in a plant growth chamber at 18–20 °C with a 16 h light/8 h dark cycle (*see* **Note 7**). Plants produce buds after about 6 weeks.

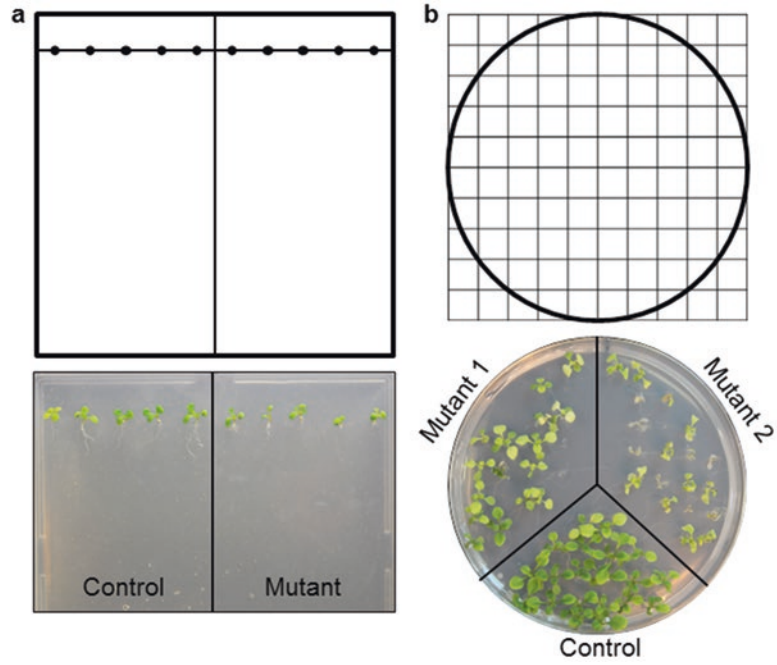


Fig. 1 Template grids and examples of Arabidopsis seedlings. (**a**) 12 cm square plate with control and mutant plants. (**b**) 9 cm Petri dish with a control and two different mutant lines

2. For DNA repair analysis, seeds are washed 10 min in 2.5% sodium hypochlorite. Rinse three times in sterile water for 5 min, and leave them overnight at 4 °C in darkness.

3.2 Chromosome Preparation by Spreading

In order to get a fine-tuned characterization, it is very important to obtain high-quality chromosome spreads. This protocol has been previously developed on several publications [6, 7, 19]. We include here slight modifications and variations that should be applied to get better results.

1. Inflorescences prior to flower opening from healthy plants are cut and fixed in fresh Carnoy's fixative. Leave them on the bench overnight at room temperature. Replace the fixative during the following days as many times as necessary to get totally white inflorescences (*see* **Note 8**). Fixed samples can be stored at −20 °C for at least 2 years.
2. Transfer inflorescences to a watch glass. Individualize buds by using a needle and forceps, and remove any with yellow anthers that contain pollen.
3. Wash three times (5 min each wash) in 3:1 fixative, followed by citrate buffer (three times, 5 min each wash). Then incubate in the enzyme mixture inside a moist chamber at 37 °C, for 2 h (*see* **Note 9**).

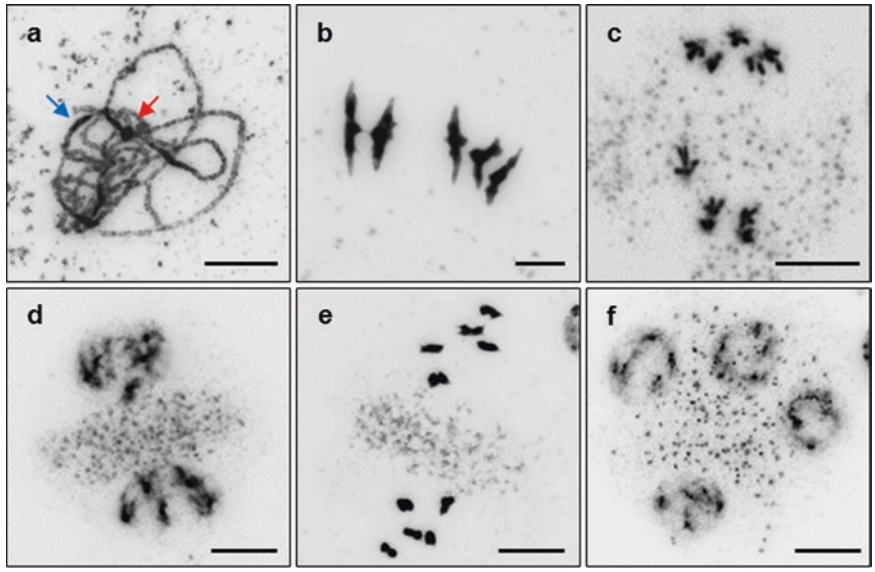


Fig. 2 Representative images of different meiotic stages in pollen mother cells. (**a**) Pachytene with full synapsis. Nucleolus organizing regions are indicated by a red arrow. The *blue arrow* points out a pericentromeric heterochromatic region. (**b**) Typical wild-type metaphase I with five bivalents. (**c**) Anaphase I showing the segregation of five chromosomes to each pole. (**d**) Prophase II with balanced nuclei. (**e**) Metaphase II with five chromosomes in each nucleus. (**f**) Tetrad containing four meiotic products. Bars represent 5 μm

4. Replace the enzyme mixture with ice-cold citrate buffer to stop the enzyme reaction.
5. A single bud is transferred to a clean slide with a minimal amount of buffer and tapped out using a fine needle to produce a cell suspension (*see* **Note 10**).
6. Add 10 μL 60% acetic acid to the cell suspension, and place the slide on the hot plate at 45 °C for 1 min. Add a further drop of 10 μL 60% acetic acid on the suspension. This way cells dissociate to form a monolayer.
7. Refix the material with 100 μL of ice-cold 3:1 fixative as a circle around the suspension on the slide. Add a further drop of 100 μL of fixative and air-dry (*see* **Note 11**).
8. It is convenient to check the meiotic stage of the buds before DAPI staining under phase contrast microscope. This can help to establish a possible link between meiotic stages and bud size (*see* **Notes 12** and **13**).
9. Finally, stain the slides with DAPI in Vectashield antifade mounting medium. Capture images at the highest-quality settings, and adjust brightness and contrast with Adobe Photoshop software (*see* **Note 14**) (Fig. 2).

3.3 Preparation of Probes Using Nick Translation

1. To produce the probes, combine the following in a 200 μL tube: 1 μg of DNA (45S rDNA, plasmid pTa71; or 5S rDNA, plasmid pCT4.2), a nick translation mixture, and a reaction mix containing dNTPs and digoxigenin–dUTP (for 45S rDNA) or dNTPs and biotin–dUTP (for 5S rDNA). Digoxigenin is bigger than dTTP; for this reason the proportion between dTTP and digoxigenin should be 2 dTTP:1 digoxigenin. Biotin is smaller, the proportion should be 1 dTTP:10 biotin (*see* **Note 15**). Use a nick translation labeling kit and follow the manufacturer's instructions.
2. Incubate for 90–120 min (according to the manufacturer's instructions) at 15 °C, and place reaction on ice afterward.
3. Stop the reaction by adding 1 μL of 0.5 M EDTA, pH 8, and heating to 65 °C for 10 min. Place on ice for 5 min prior to further use, or store at 4 °C.

3.4 Fluorescence In Situ Hybridization

1. Pretreat slides obtained by chromosome spreading by washing in 2× SSC for 10 min at room temperature.
2. Digest material on the slide with pepsin at 37 °C for 90 s (*see* **Note 16**).
3. Wash again in 2× SSC at room temperature for 10 min.
4. Fix the material on the slides in 4% paraformaldehyde at room temperature for 10 min.
5. Rinse the slides in distilled water, and then dehydrate using increasing concentrations of ethanol (70%, 90%, and 100%), for 2 min each wash.
6. Air-dry for at least 1 h. Continue the protocol or store at 4 °C.
7. Prepare 20 μL of the probe mixture per slide: 14 μL hybridization mix and 3 μL of each labeled probe. Denature the probe by heating for 10 min at 80 °C. Cool on ice for 5 min.
8. Add the mixture on the slides (cut the tip) and cover with a coverslip. Heat the slide on a hot plate for 4 min at 72 °C.
9. Hybridize the probe and chromosomes by incubating overnight at 37 °C in a moist chamber.
10. Blow over the coverslip, and wash the slides three times, 5 min each, in 50% formamide—2× SSC at 45 °C.
11. Wash once in 2× SSC at 45 °C, 5 min.
12. Wash twice in 4T, 5 min each, once at 45 °C, and once at room temperature.
13. Add 50 μL of a solution containing fluorescent antibodies to each slide, cover with parafilm, and incubate in a moist chamber at 37 °C, in darkness for 1 h.
14. Wash again the slides in 4T for three times, 5 min each.

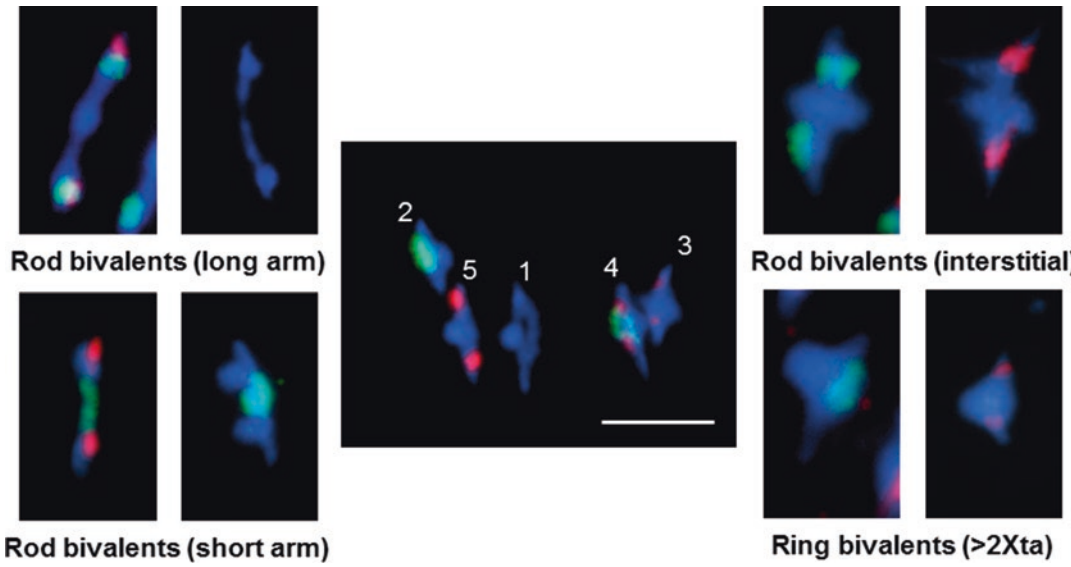


Fig. 3 Chromosome identification and chiasma analysis by FISH. While 45S rDNA regions (*green*) are located on chromosomes 2 and 4, 5S rDNA regions (*red*) appear on chromosomes 3, 4, and 5 in Arabidopsis Col-0 accession. An example of metaphase I with five ring bivalents (with two chiasmata, one per arm) is displayed in the middle. Some representative individual bivalents are also showed: rod bivalents with a single (*distal/interstitial*) chiasma (*left side*); rod bivalents with an interstitial chiasma (*top right*); ring bivalents with more than two chiasmata (*bottom right*). Xta: chiasmata. Bar represents 5 μm

15. Lastly, mount the slides in 10 μL DAPI/Vectashield. Capture images at the highest-quality settings, merge channels, and adjust brightness and contrast with Adobe Photoshop software.
16. Estimate chiasma frequency by counting chiasmata at metaphase I (*see* **Note 17**) (Fig. 3).

3.5 5-Methylcytosine Immunolocalization

1. Dry slide preparations at 60 °C (30 min).
2. Fix in 1% paraformaldehyde (10 min).
3. Dehydrate the slides through an alcohol series of 70%, 90%, and 100% ethanol, 2 min in each solution.
4. Denature in HB50 at 80 °C for 2 min. This step increases the accessibility of the modified bases to antibodies.
5. Wash in ice-cold 2× SSC twice, 5 min each.
6. Incubate for 1 h in blocking buffer.
7. Wash three times in 1× PBS and once in TNT, 5 min each.
8. Incubate with the 5-mC antibody and cover with parafilm in a moist chamber at 37 °C, 30 min.
9. Wash the slides three times in TNT, 5 min each, at room temperature.

10. Add secondary antibody, cover with parafilm, and incubate for 30 min at 37 °C in a moist chamber.
11. Finally, wash the slides again three times in TNT, 5 min each, at room temperature in darkness.
12. Counterstain with DAPI in Vectashield (*see* **Note 18**).
13. Capture images at the highest-quality settings, and adjust brightness and contrast with Adobe Photoshop software.

3.6 Cisplatin Treatment

1. Sow sterile wild-type (WT) and *ago* plants in 9 cm Petri plates using a template grid (Fig. 1b). Plates have GM with different doses of the agent: a suitable range would be 0, 15, 30, 50, and 75 μM. Assure at least 200 plants per line and per dose (this could be done by putting 40 WT seeds and 40 mutant seeds in 5–7 plates) (*see* **Note 19**). Sowing seeds could be done by dropping individual seeds with a sterile glass capillary tube (*see* **Note 20**). Leave the plates to dry for a few minutes before closing them with micropore tape. Use this as general advice for the remaining methods.
2. Grow plants in long-day conditions (16 h light, 8 h darkness) at 20–22 °C for 14 days.
3. Take representative pictures. Quantify germinated seeds and the number of true leaves per dose per line (*see* **Note 21**). Measure the total fresh weight of plants per dose per line.
4. Leave plants O/N in an oven at 50 °C to dry the seedlings. Measure the total dry weight of plants per dose per line. If the experiment is repeated several times, statistical analysis would be possible with fresh and dry weight data.
5. Normalize data to the control conditions (0 μM) for each line before performing the statistical tests between WT and mutant plants per dose.

3.7 Gamma Ray Treatment

1. Irradiate sterile WT and *ago* seeds in a range of 0, 100, 200, 300, 400, and 500 Gy during an exposure time which depends on the power of the ^{137}Cs source. Leave seeds vernalizing overnight. Sow at least 200 seeds per line in 9 cm Petri dishes with GM using a template grid (Fig. 1b).
2. Grow plants for 14 days (*see* Subheading 3.6). Take representative pictures; evaluate the number of germinated seeds, fresh and dry weight (*see* Subheading 3.6).

3.8 Mitomycin C Treatment

1. Sow sterile seeds in 9 cm Petri dishes with GM (without MMC), and let them grow for 5 days in normal conditions (*see* **Note 22**).
2. Transfer two-leaf plants into sterile 24-well plates with liquid GM. Put an individual plant in each well containing 1.8 mL of

liquid GM (without MMC). Leave them for 24 h growing in normal conditions.

3. Add 0.2 mL of MMC in liquid GM. Use an appropriate concentration to achieve the desired final dose (i.e., 3, 6, 9, and 12 μg/mL).
4. After 12 days of growing, take representative pictures, count the number of leaves, and measure fresh and dry weight as described before.

3.9 UV-C Treatment

1. Sow sterile seeds in 10 or 12 cm square GM plates (Fig. 1a) including both (control and *ago* seeds) in the same plates. Grow the plants putting plates vertically to assure gravitropism, and facilitate root measurements (*see* **Note 23**).
2. Irradiate two-leaf plants after 5 days of being sowed (*see* **Note 24**). Cover the plates with aluminum foil for 6 days to avoid endogenous photoreactivation (*see* **Note 25**).
3. Take pictures of each plate daily or in alternate days until 6 days after irradiation using a black background to highlight roots over GM. Measure root length manually using Image J CTRL + M tool (with a reference) or download a specific program such as BRAT (http://www.plant–image–analysis.org/software/brat).

4 Notes

1. Beware of chromosome rearrangements. Chromosome translocations appear in around 10% of T-DNA lines. These rearrangements produce defects in meiosis and a reduction in fertility. The cytological analysis presented allows the identification of possible translocations.
2. Working dilution (1×) should be prepared fresh with distilled water, and the stock can be stored at 4 °C.
3. A stock solution containing 1% of each enzyme in citrate buffer is stored at −20 °C. Working dilution (1×) should be prepared fresh with citrate buffer (1×).
4. Alternatively, the use of centromeres (pAL1) and telomeres (pLT11) should provide additional information for the cytological characterization of the mutants.
5. It is possible to reuse the solution for at least five times.
6. Concentration of damaging agents might be lower or higher depending on the hypersensitivity of the mutant. Some lines require different doses to show a phenotype. Described here are the concentrations and doses used in [5].
7. Sow plants at a low density in order to obtain healthy plants that will contain PMCs at all meiotic stages.

8. Buds must be fixed for at least 1 week prior to digestion for best results.
9. In order to get a better digestion, avoid floating buds.
10. Beware that the sample does not dry out. Add more buffer if necessary.
11. Slides could be stored at 4 °C before DAPI staining for at least 2 months.
12. The link between meiotic stages and bud size is disturbed in *ago* mutants.
13. At this point slides could be used directly for FISH (if they contain chromosomes at the desired stage). If the slides are stained with DAPI, coverslips and mounting medium should be removed by washing overnight in 4T buffer.
14. We prefer to use gray-scale images for a better distinction of chromosome morphology.
15. We prefer to use red for detecting 5S rDNA and green for 45S rDNA (nucleolus organizing regions, NORs), since 5S rDNA signals are smaller and red is usually clearly distinguished than green. For chromosome identification, *see* Fig. 1 from [7].
16. Increase the time if chromosomes appear to be covered by cytoplasm, but avoid degrading chromatin with a long treatment.
17. To get an accurate estimation of chiasmata, a broad cytological expertise is needed.
18. After this protocol, slides can be processed for FISH.
19. Sow extra seeds to overcome problems of germination and microbial contamination of the plates.
20. Alternatively, dry sterile seeds can be sowed with a sterile wooden toothpick.
21. It is quicker to count the total number of leaves, and then subtract two (cotyledon leaves) per plant.
22. Plan in advance the number of plants needed for the MMC analysis to sow them first in GM plates, and then transfer them to 24-well plates of liquid GM.
23. Sometimes it is recommendable to add extra percentage of agarose to help maintain the medium in vertical position.
24. Consider the power of the UV–C lamp used before irradiation. Calculate the time and the distance needed to get a dose that can trigger response, 300 J/m^2.
25. Photoreactivation can reverse DNA damage coming from UV light. Plants have a specific enzyme that use energy from light to correct 6–4 subproducts and cyclobutane pyrimidine dimers (CPD) created by UV irradiation [13].

Acknowledgments

MMG and MP acknowledge the support of the European Union by the advanced grant Marie Curie Initial Training Network (ITN) "COMREC" (Grant agreement number: 606956). In addition, the work in MP laboratory is supported by a grant from the Ministerio de Economía y Competitividad of Spain (AGL2015–67349–P). The authors thank Juan Luis Santos and Eugenio Sánchez–Morán for helpful discussions about the importance of accurate chiasma quantification in the analysis of meiotic mutants, and Alexander Knoll and Holger Puchta for sharing protocols related to sensitivity assays using genotoxic agents.

References

1. Borges F, Martienssen RA (2015) The expanding world of small RNAs in plants. Nat Rev Mol Cell Biol 16(12):727–741. doi:10.1038/nrm4085
2. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC (2010) The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22(2):321–334. doi:10.1105/tpc.109.072199
3. Zhang H, Xia R, Meyers BC, Walbot V (2015) Evolution, functions, and mysteries of plant ARGONAUTE proteins. Curr Opin Plant Biol 27:84–90. doi:10.1016/j.pbi.2015.06.011
4. Oliver C, Santos JL, Pradillo M (2016) Accurate chromosome segregation at first meiotic division requires AGO4, a protein involved in RNA-dependent DNA methylation in *Arabidopsis thaliana*. Genetics 204(2):543–553. doi:10.1534/genetics.116.189217
5. Oliver C, Santos JL, Pradillo M (2014) On the role of some ARGONAUTE proteins in meiosis and DNA repair in *Arabidopsis thaliana*. Front Plant Sci 5:177. doi:10.3389/fpls.2014.00177
6. Sanchez Moran E, Armstrong SJ, Santos JL, Franklin FC, Jones GH (2001) Chiasma formation in *Arabidopsis thaliana* accession Wassileskija and in two meiotic mutants. Chromosom Res 9(2):121–128
7. Sanchez-Moran E, Armstrong SJ, Santos JL, Franklin FC, Jones GH (2002) Variation in chiasma frequency among eight accessions of *Arabidopsis thaliana*. Genetics 162(3):1415–1422
8. Lopez E, Pradillo M, Oliver C, Romero C, Cunado N, Santos JL (2012) Looking for natural variation in chiasma frequency in *Arabidopsis thaliana*. J Exp Bot 63(2):887–894. doi:10.1093/jxb/err319
9. Copenhaver GP, Browne WE, Preuss D (1998) Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads. Proc Natl Acad Sci U S A 95(1):247–252
10. Lu P, Han X, Qi J, Yang J, Wijeratne AJ, Li T, Ma H (2012) Analysis of Arabidopsis genome-wide variations before and after meiosis and meiotic recombination by resequencing Landsberg erecta and all four products of a single meiosis. Genome Res 22(3):508–518. doi:10.1101/gr.127522.111
11. Drouaud J, Khademian H, Giraut L, Zanni V, Bellalou S, Henderson IR, Falque M, Mezard C (2013) Contrasted patterns of crossover and non-crossover at *Arabidopsis thaliana* meiotic recombination hotspots. PLoS Genet 9(11):e1003922. doi:10.1371/journal.pgen.1003922
12. Mathieu O, Picard G, Tourmente S (2002) Methylation of a euchromatin-heterochromatin transition region in *Arabidopsis thaliana* chromosome 5 left arm. Chromosom Res 10(6):455–466
13. Manova V, Gruszka D (2015) DNA damage and repair in plants–from models to crops. Front Plant Sci 6:885. doi:10.3389/fpls.2015.00885
14. Bonnet S, Knoll A, Hartung F, Puchta H (2013) Different functions for the domains of the *Arabidopsis thaliana* RMI1 protein in DNA cross-link repair, somatic and meiotic recombination. Nucleic Acids Res 41(20):9349–9360. doi:10.1093/nar/gkt730
15. Pradillo M, Knoll A, Oliver C, Varas J, Corredor E, Puchta H, Santos JL (2015) Involvement of the cohesin cofactor PDS5 (SPO76) during meiosis and DNA repair in *Arabidopsis thaliana*. Front Plant Sci 6:1034. doi:10.3389/fpls.2015.01034

16. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301(5633):653–657. doi:10.1126/science.1086391

17. Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7(7):1869–1885

18. Campell BR, Song Y, Posch TE, Cullis CA, Town CD (1992) Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. Gene 112(2):225–228

19. Armstrong SJ, Sanchez-Moran E, Franklin FC (2009) Cytological analysis of *Arabidopsis thaliana* meiotic chromosomes. Methods Mol Biol 558:131–145. doi:10.1007/978-1-60761-103-5_9

Chapter 11

Isolation and Characterization of *ARGONAUTE* Mutants in *Chlamydomonas*

Tomohito Yamasaki

Abstract

Random insertional mutagenesis and subsequent reverse genetic screening allow the isolation of mutants of interest. Here I describe the protocol for generating a tag insertion line and subsequent PCR-based screening for *ARGONAUTE* mutants as an example of a reverse genetic screen for the unicellular green alga *Chlamydomonas reinhardtii*.

Key words *Chlamydomonas*, Mutagenesis, ARGONAUTE, miRNA, PCR, Genomic DNA extraction

1 Introduction

ARGONAUTE (AGO) proteins incorporate microRNAs to form an RNA-induced silencing complex that recognizes target transcripts and induces endonucleolytic cleavage and/or translational repression [1]. The genome of the unicellular green alga *Chlamydomonas reinhardtii* harbors three *AGO* genes (*AGO1–3*) [2]. However, their individual physiological significance and exact functions are largely unknown [3–6]. Here I describe the method for a PCR-based reverse genetic screen for isolating *AGO* mutants that is a modified version of the original screening method [7]. The modified method is composed of five steps as follows: (1) preparation of the *aph7″* marker gene as an insertion DNA tag, (2) generation of a tag insertion library by nuclear transformation, (3) preparation of genomic DNA pools comprising genomic DNA of thousands of different transformants, (4) PCR-based screening in which one primer anneals to the *aph7″* gene and another anneals to the target *AGO* gene, and (5) validation of the isolated mutants. Building on the original method, the conditions for nuclear transformation were optimized to achieve more efficient generation of the single-copy transgenic library. The sensitivity and specificity of

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_11,

the PCR screening were improved using a thermostable DNA polymerase that was optimized for amplification of high G/C and long fragments and hexadecyltrimethylammonium bromide (CTAB) for purifying genomic DNA. In addition, some tricks for conducting a high-throughput PCR screening are described. With this method, I have isolated mutants of all three *AGO* paralogs and other specific genes of interest from ~50,000 tag insertion lines [6] (data not shown).

2 Materials

2.1 PCR

1. KOD FX Neo DNA polymerase (Toyobo). 2× PCR buffer for KOD FX Neo and 2 mM dNTPs mixture are packaged together (*see* **Note 1**).
2. pHyg3 plasmid (20 pg/μL) [8]. This plasmid harbors the *aph7″* expression cassette, which confers hygromycin B resistance and is available for *Chlamydomonas* from the Chlamydomonas Resource Center (http://www.chlamycollection.org/).
3. Thermal cycler.
4. Primers for amplification of the *aph7″* expression cassette: pHyg3-F1 primer (5′-CACACAGGAAACAGCTATGAC CATG-3′) and pHyg3-R1 (5′-CGTTGTAAAACGACGGCC AGTG-3′). For PCR screening for *ARGONAUTE* mutations: aph7-F1 (5′-GACGTCTATGCGGGAGACTC-3′), aph7-F2 (TGGTGCAACTGCATCTCAAC-3′), aph7-R1 (5′-CGAAT CAATACGGTCGAGAAGTAACAG-3′), AGO1-Fo (5′-CAG GCACCACTAGTGTTATAGAAGG-3′), AGO1-Fi (5′-CCTA GGTCTACATCTGAATGCTTGG-3′), AGO1-Ri (5′-CCAG CGCACATGTTTTAGATTCATG-3′), AGO1-Ro (5′-CAAC TTGCCGGACAGTTTGTAG-3′), AGO2-Fo (5′-GCTCTT TTAATGCCCGCTTTGAG-3′), AGO2-Fi (5′-GAGGCTAA GTTCATTCATATACTTTGAGAG-3′), AGO2-Ri (5′-GGG GAACTGAAGTAAAAGAGACTGC-3′), AGO2-Ro (AGAA CGAAGCGATGAAATCACTTG-3′), AGO3-Fo (5′-TGTG AGATAAAGCCTCGAGAGCTC-3′), AGO3-Fi (5′-GAATG GTTAGCCGATTCCATGGAC-3′), AGO3-Ri (5′-ACACA GCCAAAGAGCAAAGGTG-3′), and AGO3-Ro (5′-GTAAC ACTAGTACGGTACCACGTCG-3′) (Fig. 1) (*see* **Note 2**). All primers were used at concentrations of 5 μM.

2.2 Gel Purification of PCR Amplicons

1. Submarine electrophoresis system.
2. Agarose powder.
3. Tris-acetate-ethylenediaminetetraacetic acid (EDTA) buffer (TAE): 40 mM Tris-base, 40 mM acetic acid, 1 mM EDTA (pH 8.0). To prepare 1 L of 50× TAE buffer, weigh 242 g

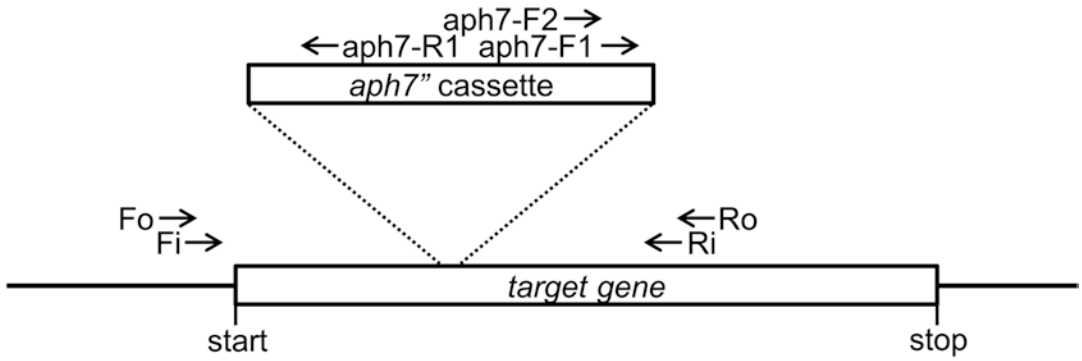


Fig. 1 Position of primers. The location and orientation of gene-specific primers for a target gene and *aph7″* as arrows. To screen a mutant strain that has a tag insertion within ~2 kb downstream of the start codon, gene-specific forward primers anneal immediately upstream of the start codon, and gene-specific reverse primers anneal ~2 kb downstream of the start codon. The positions of Fo and Ro primers should be shifted 200–500 bp outside the positions of Fi and Ri primers in order to see size difference on an agarose gel

Tris-base and dissolve in 800 mL water. Add 57.1 mL acetic acid and 100 mL of 0.5 M EDTA (pH 8.0). Make up to 1 L with water. Dilute 50-fold prior to use.

4. PCR product purification kit (e.g., NucleoSpin gel and PCR cleanup, Takara).
5. Microvolume spectrophotometer (e.g., Nanodrop, Thermos Fisher Scientific).
6. UV transilluminator.
7. Ethidium bromide solution (10 mg/mL).

2.3 Chlamydomonas Transformation

1. Phosphate buffer II. To prepare 100 mL of phosphate buffer II, weight 10.8 g of dipotassium phosphate (K_2HPO_4) and 5.6 g of monopotassium phosphate (KH_2PO_4), and dissolve in 80 mL ultrapure water. Make up to 100 mL with ultrapure water.
2. Solution A. To prepare 500 mL of solution A, weigh 20 g of ammonium chloride (NH_4Cl), 5 g of magnesium sulfate heptahydrate ($MgSO_4{\cdot}7H_2O$), and 2.5 g of calcium chloride dehydrate ($CaCl_2{\cdot}2H_2O$) and dissolve in 450 mL ultrapure water. Make up to 500 mL with ultrapure water.
3. Hutner's trace elements. Dissolve 50 g EDTA disodium salt in 250 mL of boiling water. Dissolve 22 g of zin sulfate heptahydrate ($ZnSO_4{\cdot}7H_2O$) in 100 mL ultrapure water. Dissolve 11.4 g of boric acid (H_3BO_3) in 200 mL ultrapure water. Dissolve 5.06 g of manganese chloride tetrahydrate ($MnCl_2{\cdot}4H_2O$) in 50 mL ultrapure water. Dissolve 1.61 g of cobalt chloride hexahydrate ($CoCl_2{\cdot}6H_2O$) in 50 mL ultrapure water. Dissolve 1.57 g of copper sulfate pentahydrate

($CuSO_4 \cdot 5H_2O$) in 50 mL ultrapure water. Dissolve 1.10 g of ammonium molybdate hydrate [$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$] in 50 mL ultrapure water. Dissolve 4.99 g of iron sulfate heptahydrate ($FeSO_4 \cdot 7H_2O$) in 50 mL ultrapure water. $FeSO_4$ solution should be prepared lastly to avoid oxidation. Mix all the solutions except EDTA solution. Boil the mixture on a hot plate with magnetic stirrer, and then add EDTA solution. The color of mixture will be green. After dissolving everything, cool to 70 °C on a hot plate with magnetic stirrer. Dissolve 17 g of potassium hydroxide (KOH) in 50 mL ultrapure water, and make up to 85 mL with ultrapure water. Warm KOH solution to 70 °C and add to the mixture. Cool the mixture to room temperature, and make up to 1 L with ultrapure water. Transfer the mixture to a 2 L flask, close the flask with a cotton plug, and swirl it once a day for 2 weeks. The color of mixture will turn purple. Filter the mixture through two layers of filter paper (e.g., Whatman #1 filter paper). Store the solution at 4 °C.

4. Tris-acetate-phosphate (TAP) medium. To prepare 1 L of TAP, weigh 2.42 g Tris base and dissolve in 800 mL ultrapure water. Add 1 mL of phosphate buffer II, 10 mL of solution A, 1 mL of Hutner's trace elements, and 1 mL of glacial acetic acid. Make up to 1 L with ultrapure water and sterilize by autoclaving.
5. Hemocytometer.
6. Square pulse electroporator (e.g., NEPA-21 electroporator, Nepagene [9]) (*see* **Note 3**).
7. Electroporation cuvette (2 mm gap).
8. TAP-sucrose medium. Dissolve sucrose at a concentration of 50 mM in the TAP medium and autoclave.
9. *Chlamydomonas* wild-type strain CC-125.
10. TAP agar (1.5%) plate containing 15 μg/mL of hygromycin B.
11. Toothpick (sterile).
12. 96-well culture plates (sterile).
13. Plant growth chamber.
14. Rotary shaker.
15. 15 mL conical tube (sterile).
16. 50 mL conical tube (sterile).

2.4 DNA Extraction

1. 200 μL 12-channel pipette.
2. 1.5 mL tubes.
3. 3 M sodium chloride (NaCl).
4. TEN buffer: 10 mM Tris–HCl (pH 8.0), 10 mM EDTA (pH 8.0), 150 mM NaCl.
5. Proteinase K (20 mg/mL).

6. Sodium dodecyl sulfate (SDS)-EB buffer: 0.5% SDS, 250 mM NaCl, 100 mM EDTA (pH 8.0), 100 mM Tris–HCl (pH 8.0). Immediately before use, add 5 μL of proteinase K (20 mg/mL) to 1 mL of SDS-EB buffer.
7. Neutralized phenol (pH 8.0).
8. Chloroform.
9. 3-Methyl-1-butanol.
10. 2-Mercaptoethanol.
11. Tube rotator.
12. CTAB/NaCl buffer: 10% (w/v) CTAB, 0.7 M NaCl.
13. CTAB precipitation buffer: 1% (w/v) CTAB, 50 mM Tris–HCl (pH 8.0), 0.1 mM EDTA (pH 8.0).
14. RNase A (10 mg/mL).
15. High salt TE buffer: 10 mM Tris–HCl (pH 8.0), 1 mM EDTA (pH 8.0), 1 M NaCl, RNase A (100 μg/mL). Immediately before use, add 1/100 volume of concentrated RNase A (10 mg/mL).
16. Isopropyl alcohol.
17. 70% ethanol.
18. Tris-EDTA buffer: 10 mM Tris–HCl (pH 8.0), 10 mM EDTA (pH 8.0).
19. Microvolume spectrophotometer (e.g., Nanodrop, Thermos Fisher Scientific).

3 Methods

Unless otherwise noted, experiments are performed at room temperature.

3.1 Preparation of the Insertion DNA Tag by PCR

1. Mix 100 μL of 2× PCR buffer for KOD FX Neo, 40 μL of 2 mM dNTPs mixture, 8 μL of pHyg3-F1 primer, 8 μL of pHyg3-R1 primer, 10 μL of pHyg3, 1 μL of KOD FX Neo DNA polymerase, and 33 μL of distilled water. Dispense 50 μL of the mixture into four PCR tubes.
2. Amplify the *aph7″* expression cassette by PCR (Table 1).
3. Resolve the amplicon (1801 bp) by 1% agarose/TAE gel electrophoresis.
4. Cut out the appropriately sized DNA band from the gel, and recover the DNA using the PCR product purification kit.
5. Determine the concentration of the DNA fragment, and adjust to 50 ng/μL with ultrapure water.

Table 1
Program for *aph7″* cassette amplification

1 cycle	94 °C	2 min
30 cycles	98 °C 60 °C 68 °C	10 s 30 s 1 min
1 cycle	4 °C	Indefinite period

3.2 Generation of a Tag Insertion Library

1. Preculture *Chlamydomonas* cells in 5 mL TAP medium in a test tube. Inoculate a small number of cells into 5 mL TAP (*see* **Note 4**). Grow the cells for 2–3 days with shaking (140 rpm) under continuous dim light.
2. Count cells and determine cell density using a hemocytometer. Inoculate 5×10^6 cells into 200 mL of TAP medium in a 500 mL flask. Grow *Chlamydomonas* cells for 2 days with shaking (140 rpm) under continuous dim light to mid-log phase (~$2–4 \times 10^6$ cells/mL).
3. Using a 50 mL tube, collect 1×10^8 cells by centrifugation at $2000 \times g$ for 5 min. Decant and discard the TAP medium, and suspend the cell pellet in 20 mL of TAP-sucrose medium to wash the cells.
4. Centrifuge at $2000 \times g$ for 3 min. Decant and discard the TAP-sucrose medium.
5. Suspend the pellet in 1 mL of TAP-sucrose medium. The final cell density will be ~1×10^8 cells/mL.
6. Transfer a 125 μL aliquot of the cell suspension to a 2 mm gap cuvette for a no-DNA negative control.
7. Add 7 μL of the *aph7″* DNA fragment solution to the remaining 0.875 mL cell suspension (*see* **Note 5**).
8. Dispense 125 μL of the cell suspension into seven 2 mm gap cuvettes. Seven cuvettes are used for transformation, and one cuvette is used for the no-DNA control.
9. Perform electroporation. The settings are described in Table 2 (*see* **Note 6**).
10. Transfer the electroporated cells from the individual cuvettes into 5 mL of TAP-sucrose in 15 mL individual conical tubes.
11. Incubate the 15 mL tubes for 16–20 h under dim light to allow for the expression of *aph7″*.
12. Centrifuge the tubes at $2000 \times g$ for 5 min, discard the TAP-sucrose medium, and suspend the cells in 5 mL of TAP to wash the cells.

Table 2
Settings for NEPA-21 electroporation

	Voltage (V)	Pulse length (ms)	Pulse interval (ms)	Number of pulses	Decay rate (%)	Polarity
Poring pulse	300	8	50	2	40	+
Transfer pulse	20	50	50	1	40	+/−

13. Centrifuge at 2000 × *g* for 5 min, and suspend the cells in 500 μL of TAP in individual tubes.
14. Plate 100 μL of the cell suspension onto a TAP agar plate containing hygromycin B. Each suspension originating from a single electroporation cuvette is dispensed onto five selective plates.
15. Place the plates under dim light for 8–10 days. Typically, 100–200 hygromycin-resistant colonies appear per plate (*see* **Note 7**).
16. Pick the colonies with toothpick and release into a well of a 96-well culture plate, in which each well is filled with 150 μL of TAP (*see* **Note 8**).
17. Incubate the plates under dim light for 5–7 days.
18. Replicate the 96-well culture plates using a 12-channel pipette. Suspend the cell cultures in each well by pipetting, take 20 μL of cell culture, and inoculate into a new 96-well culture plate, in which each well is filled with 130 μL of TAP.
19. Place the replica plates under dim light for 3–5 days and keep the original plates.

3.3 Preparation of the Genomic DNA Pool

1. Combine 50 μL of cell culture from each well of a replica plate, and pour the mixture into a 15 mL tube. The mixture will contain 96 different transgenics (*see* **Note 9**).
2. Centrifuge at 2000 × *g* for 5 min, discard TAP medium, suspend the cells in 1 mL of TEN buffer, and then transfer the cell suspension to a 1.5 mL tube using a 1 mL pipette.
3. Centrifuge at 5000 × *g* for 5 min, and then discard supernatant using 1 mL pipette. Store the pellets at −20 °C to stop the procedure at this point, or suspend the cells in 0.15 mL of SDS-EB to proceed to the next step (*see* **Note 10**).
4. Incubate the cell suspension at 55 °C for 30 min.
5. Add 5 μL of 2-mercaptoethanol and 150 μL of phenol/chloroform/3-methyl-1-butanol (25:24:1).
6. Vortex the solution to completely emulsify, and then immediately start to rotate using a tube rotator. Rotate at top speed for 30 min.

7. Centrifuge at 12,000 × *g* for 5 min.
8. Transfer the upper aqueous phase to a new 1.5 mL tube, and add an equal volume of phenol/chloroform/3-methyl-1-butanol (25:24:1). Do not take the lower organic phase.
9. Vortex the solution to completely emulsify, and then immediately start to rotate using a tube rotator. Rotate at top speed for 30 min.
10. Centrifuge at 12,000 × *g* for 5 min.
11. Transfer the upper aqueous phase to a new 1.5 mL tube. Do not take the lower organic phase and the white intermediate layer. Add 1/6 volume of 3 M NaCl and mix completely. Then add 1/10 volume of CTAB/NaCl buffer, and mix completely (*see* **Note 11**).
12. Incubate the solution at 55 °C for 5 min.
13. Add an equal volume of chloroform/3-methyl-1-butanol (24:1).
14. Vortex the solution for 10 s, and then immediately start to rotate using a tube rotator. Rotate at top speed for 10 min.
15. Centrifuge at 12,000 × *g* for 5 min.
16. Transfer the upper aqueous phase to a new 1.5 mL tube, add two volumes of CTAB precipitation buffer, and mix well (*see* **Note 12**).
17. Incubate the solution at 65 °C overnight (*see* **Note 13**).
18. Centrifuge at 2000 × *g* for 2 min to precipitate the white aggregates and completely discard the supernatant (*see* **Note 14**).
19. Suspend the white pellet with 50 μL of the high salt TE buffer (*see* **Note 15**).
20. Incubate the solution at 37 °C for 2 h to degrade contaminating RNA.
21. Add 50 μL of isopropyl alcohol and mix well (*see* **Note 16**).
22. Centrifuge at 12,000 × *g* for 10 min.
23. Discard the supernatant and add 0.5 mL of 70% ethanol.
24. Centrifuge at 12,000 × *g* for 5 min at 4 °C.
25. Discard the supernatant.
26. Open the lid of the tube and leave for 5 min to air dry the DNA pellet.
27. Suspend the DNA pellet in 32 μL of water.
28. Use 2 μL of the DNA solution to check the concentration of the genomic DNA pool, and adjust the DNA concentration to 100 ng/μL in water (*see* **Note 17**).
29. Resolve 100 ng of the genomic DNA pool using 0.8% agarose/TAE gel electrophoresis to check the quality of DNA (Fig. 2) (*see* **Note 18**).

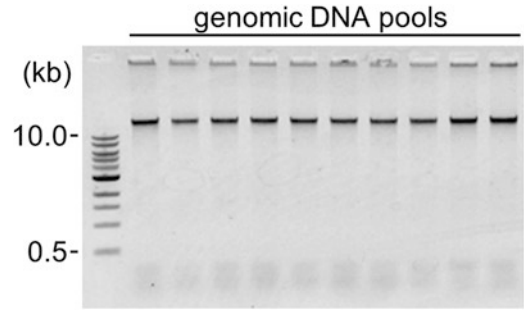


Fig. 2 Quality check of genomic DNA pools. On a 0.8% agarose/TAE gel, 100 ng of genomic DNA pools were resolved. Genomic DNA appears as a band larger than 10 kb. Partially degraded rRNA and/or tRNA appears as a smeared band smaller than 0.5 kb in negligible quantities

Table 3
Program for PCR screening of *AGO* mutants

1 cycle	94 °C	5 min
35 cycles	98 °C 60 °C 68 °C	10 s 30 s 3 min
1 cycle	4 °C	Indefinite period

30. Combine 20 μL of ten different genomic DNA pools to generate a DNA super pool. Individual DNA super pools contain 960 different genomic DNAs.

3.4 PCR Screening

1. Mix 10 μL of 2× PCR buffer for KOD FX Neo, 4 μL of 2 mM dNTPs mixture, 0.8 μL of aph7-F1 primer, 0.8 μL of one of gene-specific primers (AGO1-Fi, AGO1-Ri, AGO2-Fi, AGO2-Ri, AGO3-Fi, or AGO3-Ri), 1 μL of genomic DNA super pool, 0.1 μL of KOD FX Neo DNA polymerase, and 3.3 μL of water (*see* **Note 19**).
2. Perform PCR. The conditions are described in Table 3.
3. Resolve the PCR products using 1% agarose/TAE gel electrophoresis, and find a specific and intense PCR product (Fig. 3a) (*see* **Note 20**).
4. Perform PCR for the individual ten DNA pools that compose the super pool. Find the specific signal with the same PCR conditions used for the super pool PCR (Fig. 3b) (*see* **Note 21**).
5. Combine 10 μL of the cell culture from individual rows and columns of the 96-well replica plate that was used for genomic DNA preparation.

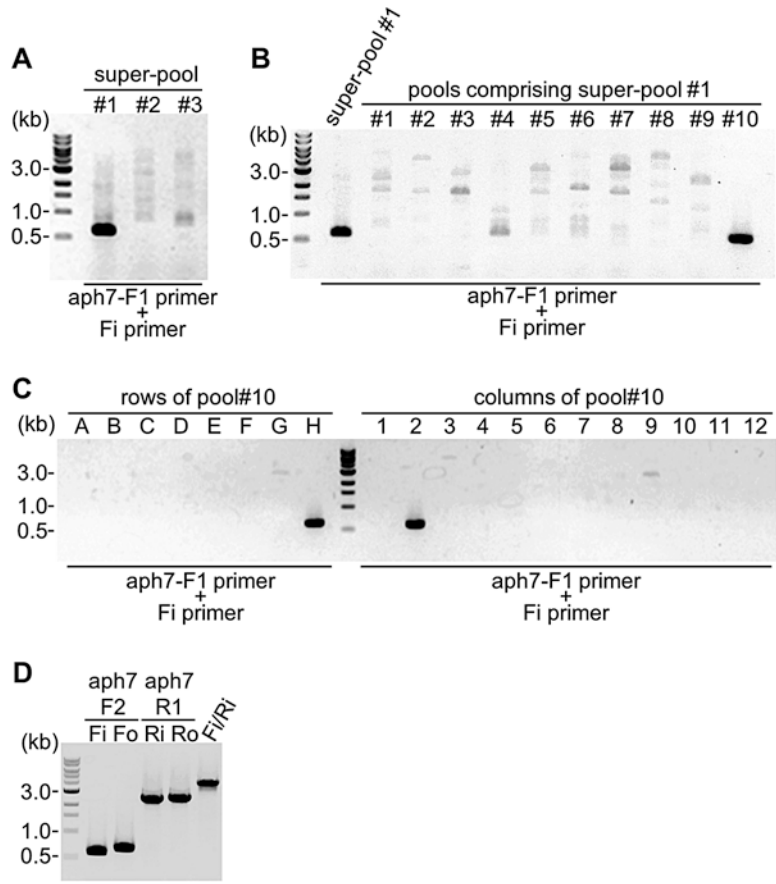


Fig. 3 Example of the results of each step of PCR screening. (**a**) PCR screening for three different super pools. The specific band derives from super pool #1. Since the aph7-F1 primer and gene-specific Fi primer generate the band, the DNA tag is inserted in the opposite direction relative to the target gene. (**b**) PCR screening for ten genomic DNA pools that comprise super pool #1. The specific band derives from pool #10. (**c**) PCR screening of rows and columns of a 96-well plate, which corresponds to pool #10 of super pool #1. Row "H" and column "2" generate the specific bands. (**d**) Final validation PCR for a mutant candidate from specified well "H2" in pool #10 of super pool #1

6. Centrifuge at 1000 × *g* for 1 min; discard supernatant using a pipette.
7. Add 50 μL of Tris-EDTA buffer and suspend the pellet.
8. Centrifuge at 1000 × *g* for 1 min; discard supernatant using a pipette.
9. Add 50 μL Tris-EDTA buffer and suspend the pellet by pipetting.
10. Incubate the cell suspension at 98 °C for 5 min, and then cool to room temperature (*see* **Note 22**).

11. Transfer the suspension to a new 1.5 mL tube, and centrifuge at 12,000 × *g* for 2 min.
12. Transfer 10 μL of the supernatant to a new 1.5 mL tube and add 40 μL of water.
13. Perform PCR for individual row and column pools with the same conditions used for the super pool DNA to find the specific signal (Fig. 3c). Identify the exact position of the well that generates the specific signal (*see* **Note 23**).
14. Take 50 μL of cell culture from the specified well of the original 96-well culture plate (not the replica plate), and extract the genomic DNA (*see* Subheading 3.4, **steps 8–12**).
15. Perform PCR on the genomic DNA of the transgenic in the specified well using the aph7-F2 primer and the appropriate "Fo" or "Ro" primer. In addition, perform PCR using the aph7-R1 primer and the appropriate "Fo" or "Ro" to check the opposite side of the DNA tag (Figs. 1 and 3d) (*see* **Note 24**).

3.5 Validation

1. Take 10 μL of the cell culture from the specified well of the original 96-well culture plate, and streak on a TAP agar plate containing hygromycin B for single colony isolation. Place the plate under dim light for 8–10 days.
2. Pick eight colonies, and transfer to a new TAP agar plate containing hygromycin B, and then place the plate under dim light for 5–6 days.
3. Take cells from the agar plate in an amount approximately the size of a rice grain using a toothpick, and suspend in 50 μL of Tris-EDTA buffer.
4. Extract DNA from the individual clones (*see* Subheading 3.4, **steps 8–12**).
5. Perform PCR on the eight clones with aph7-F1 and gene-specific primers (Fi or Ri) to amplify the specific bands.
6. Perform PCR on the eight clones with the gene-specific Fi and Ri primer set to ensure that no endogenous DNA fragment is amplified at the original size (Fig. 3d).
7. Clone the specific band, and check the DNA sequence to determine the exact boundaries between both ends of the inserted *aph7″* cassette and flanking genomic DNA regions.
8. Select one mutant clone as a mutant strain, and grow the strain till the late log phase in 50 mL of TAP medium with shaking (140 rpm) under continuous dim light (~6–8 × 10^6 cells/mL).
9. Isolate the genomic DNA of the mutant (*see* Subheading 3.3, **steps 2–29**), and perform a Southern blot to check the copy number of the integrated DNA tag. If the genome contains more than one copy, a backcross should be performed.

4 Notes

1. The thermostable DNA polymerase used in this method should be optimized for amplification of high G/C and long fragments from crude samples. Standard DNA polymerases, such as merely "Taq polymerase" or "KOD polymerase," are not sensitive enough to detect the tag insertions within the target genes with high fidelity.
2. Any insertional mutants of nuclear genes of interest can be screened using specific primer sets (Fo, Fi, Ri, and Ro) for individual genes. Since the tag DNA is integrated into the genome in two different potential orientations, gene-specific forward and reverse primers are needed to identify the insertion site. To design gene-specific primers that reduce the chance of nonspecific amplification, perform a BLAST search with the primer sequences in Phytozome 11 (https://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST&method=Org_Creinhardtii) with modified parameters as follows: Expect (E) threshold, 50; Allow gaps, unchecked. Sequences that match a large number of genomic positions with high homology should not be used as primer sequences.
3. Use a square pulse electroporator to obtain maximal transformation efficiency without cell wall removal.
4. The color of the culture medium should be quite pale green after inoculation with *Chlamydomonas* cells.
5. Use 50 ng of the tag DNA for every 125 μL of cell suspension in a cuvette, increasing the DNA concentration results in increased numbers of integrated marker gene copies. Almost all transgenics that are generated under these conditions harbor a single copy of the *aph7″* cassette in their genomes.
6. The settings described in Table 2 are optimized for the CC-125 strain. In the case of the other strains, adjustment of the voltage (250–300 V) and pulse length (4–8 ms) for the poring pulse may be needed to obtain maximal transformation efficiency.
7. Approximately 500–1000 colonies will be generated from a single cuvette.
8. Prepare at least thirty 96-well plates. These 30 plates constitute three genomic DNA super pools that are needed to compare band patterns and discriminate a specific band among them on an agarose gel. Success of the mutant screening depends on the scale of the mutant library. Continue to generate transformants until the mutants are found.
9. Use a 12-channel pipette. It is not necessary to change to sterilized tips for every row of the 96-well plates; simply wash the tips by pipetting in a 1 L bucket filled with 0.01% Tween 20.

However, tips should be changed for every 96-well plate. Collect cell cultures from each well on the inside of the lid of the 96-well culture plate, and transfer to a 15 mL tube using a 1 mL pipette. After transferring the collected cell culture, wipe off the cells inside of the lid with paper, spray 70% ethanol to kill the cells, wipe off the dead cells, and then put the lid back.

10. Invert and flip the tube to completely suspend the cell pellet. The suspension will be pulpy and look too dense, but do not use more than 0.2 mL of SDS-EB buffer in order to achieve efficient recovery of the genomic DNA.
11. It is difficult to take a precise volume of viscous CTAB/NaCl buffer. It is not necessary to take exactly 1/10 volume of the buffer, but slowly take up the buffer, and eject by pipetting. After addition, the solution will become white in color.
12. The solution will become faint white in color.
13. CTAB aggregates with nucleic acids and forms white flakes after incubation. It may appear within 1–2 h after starting the incubation, but incubation should be continued to complete aggregation.
14. The pellet is fragile. Remove the supernatant by carefully pipetting, and then spin the tubes again for a few seconds and carefully remove residual supernatant by pipetting.
15. Flipping tubes several times will dissolve aggregates.
16. After mixing, cotton-like aggregated nucleic acids appear.
17. If the DNA concentration is lower than 100 ng/μL, adjust the concentrations of all the DNA pools to the lowest concentration among them. Adjusting to the same concentration is important.
18. Make sure that there is little to no contamination by degraded RNA, which should appear as an obscure band around 300 bp (Fig. 2).
19. Use 100 ng of super pool genomic DNA for each reaction. If the DNA concentration is lower than 100 ng/μL, add an appropriate volume of super pool DNA.
20. Ignore faint bands. The desired band should appear as an intense band, if present (Fig. 3a, super pool #1).
21. Ignore nonspecific faint bands. A specific and intense band will appear from the one of the pools, if present (Fig. 3b, pool #10).
22. PCR tubes and a thermal cycler can be used for washing and heating the cells.
23. The specified well contains a candidate for an *ARGONAUTE* mutant (Fig. 3c, lanes H and 2).
24. Check the size of the amplicons based on the positions of each primer (Figs. 1 and 3d). An amplicon generated by the

gene-specific Fi and Ri primers from a mutant is ~1.8 kb longer than that from the wild-type strain, which corresponds to the added size of the DNA tag. In the case of the tag being inserted outside the region bounded by the Fo and Ro primer set, design new primers beyond the insertion site, and perform validation PCR.

References

1. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
2. Casas-Mollano JA, Rohr J, Kim EJ, Balassa E, van Dijk K, Cerutti H (2008) Diversification of the core RNA interference machinery in *Chlamydomonas reinhardtii* and the role of DCL1 in transposon silencing. Genetics 179(1):69–81. doi:10.1534/genetics.107.086546
3. Cerutti H, Ma X, Msanne J, Repas T (2011) RNA-mediated silencing in algae: biological roles and tools for analysis of gene function. Eukaryot Cell 10(9):1164–1172. doi:10.1128/EC.05106-11
4. Kim EJ, Ma X, Cerutti H (2015) Gene silencing in microalgae: mechanisms and biological roles. Bioresour Technol 184:23–32. doi:10.1016/j.biortech.2014.10.119
5. Voshall A, Kim EJ, Ma X, Moriyama EN, Cerutti H (2015) Identification of AGO3-associated miRNAs and computational prediction of their targets in the green alga *Chlamydomonas reinhardtii*. Genetics 200(1):105–121. doi:10.1534/genetics.115.174797
6. Yamasaki T, Kim EJ, Cerutti H, Ohama T (2016) Argonaute3 is a key player in miRNA-mediated target cleavage and translational repression in Chlamydomonas. Plant J 85(2):258–268. doi:10.1111/tpj.13107
7. Gonzalez-Ballester D, Pootakham W, Mus F, Yang W, Catalanotti C, Magneschi L, de Montaigu A, Higuera JJ, Prior M, Galvan A, Fernandez E, Grossman AR (2011) Reverse genetics in Chlamydomonas: a platform for isolating insertional mutants. Plant Methods 7:24. doi:10.1186/1746-4811-7-24
8. Berthold P, Schmitt R, Mages W (2002) An engineered Streptomyces hygroscopicus aph7" gene mediates dominant resistance against hygromycin B in *Chlamydomonas reinhardtii* Protist 153 (4):401–412. doi:10.1078/14344610260450136
9. Yamano T, Iguchi H, Fukuzawa H (2013) Rapid transformation of *Chlamydomonas reinhardtii* without cell-wall removal. J Biosci Bioeng 115(6):691–694. doi:10.1016/j.jbiosc.2012.12.020

Chapter 12

ARGONAUTE Genes in *Salvia miltiorrhiza*: Identification, Characterization, and Genetic Transformation

Meizhen Wang, Yuxing Deng, Fenjuan Shao, Miaomiao Liu, Yongqi Pang, Caili Li, and Shanfa Lu

Abstract

Small RNA-mediated gene silencing is a vital regulatory mechanism in eukaryotes that requires ARGONAUTE (AGO) proteins. *Salvia miltiorrhiza* is a well-known traditional Chinese medicinal plant. Therefore, it is important to characterize *S. miltiorrhiza AGO* family genes as they may be involved in multiple metabolic pathways. This chapter introduces the detailed protocol for *SmAGO* gene prediction and molecular cloning. In addition, an *Agrobacterium*-mediated genetic transformation method for *S. miltiorrhiza* is presented. These methodologies can be used to functionally study *SmAGO* genes as well as other genes of interest in *S. miltiorrhiza*.

Key words *Salvia miltiorrhiza*, Small RNA, *ARGONAUTE*, Gene identification, Genetic transformation

1 Introduction

Small RNAs (sRNAs) are 20–24 nucleotides short noncoding RNAs. They play significant regulatory roles in plant development, multiple physiological processes, and stress responses [1–6]. Plant sRNAs can be divided into two major groups, miRNAs (miRNAs) and small interfering RNAs (siRNAs), based on different modes of biogenesis. MiRNAs originate from primary transcripts (pri-miRNAs) with internal stem-loop structures [7]. SiRNAs are produced from double-stranded RNAs (dsRNAs) derived from miRNA-cleaved transcripts that are converted into dsRNAs by RNA-dependent RNA polymerases (RDRs) [8] or from natural *cis*-antisense gene pairs [9] or heterochromatin and DNA repeats [10].

Authors "Meizhen Wang and Yuxing Deng" are contributed equally to this work.

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_12, © Springer Science+Business Media LLC 2017

Plant miRNA and siRNA biogenesis is a particularly complex process involving various protein members like Dicer-likes (DCLs), RDRs, ARGONAUTEs (AGOs), DAWDLE (DDL), HYPONASTIC LEAVES1 (HYL1), SERRATE (SE), HUA ENHANCER 1 (HEN1), and others [11, 12]. To regulate gene expression, miRNA or siRNA duplexes are incorporated into the RNA-induced silencing complex (RISC), the effector complex containing AGO proteins as the core component [13–15]. Subsequently, the star strands (miRNA* or siRNA*) are removed, and the functional miRNA or siRNA strands are selected to guide mRNA cleavage [10, 16] or translational repression [17], or, in some cases, regulate gene expression at the transcriptional level through DNA methylation [18].

AGO proteins are ribonucleases, consisting of three conserved domains (PAZ, MID, and PIWI) in the C-terminus and one variable domain in the N-terminus [19, 20]. The number of AGO proteins greatly varies in different plant species, from 10 in *Arabidopsis thaliana* (Arabidopsis) to 22 in soybean [21]. Phylogenetic analyses group them into three clades: AGO1/5/10, AGO2/3/7, and AGO4/6/8/9 [21, 22]. Different AGO members have specialized sRNA binding capacity. For example, AtAGO1 is mainly associated with the function of miRNAs and prefers sRNAs harboring 5′-terminal uridine [23, 24]. AtAGO10 regulates shoot apical meristems by specifically binding to miR165/166 [25–27]. AtAGO5, preferentially binding 5′-C sRNAs, is associated with megagametogenesis [23, 28]. AtAGO2 is involved in antiviral defense [29, 30] and has a strong bias for sRNAs beginning with adenosine [23]. AtAGO7 primarily binds miR390, which guides the cleavage of *TAS3* precursor and then triggers the production of ta-siRNAs [31]. AtAGO4 preferentially binds 24-nt hc-siRNAs, which mainly harbors 5′-A, and functions in RNA-directed DNA methylation [18, 32]. AtAGO9 controls female gamete formation by interacting with 24-nt sRNAs, while AtAGO8 is a pseudogene [33, 34].

S. miltiorrhiza is a perennial plant in the genus *Salvia*, with high values in its roots. There are two major types of bioactive compounds in *S. miltiorrhiza*, lipophilic tanshinones and hydrophilic phenolic acids. They have been widely used for treating cardiovascular and cerebrovascular diseases for thousands of years [35, 36]. Recently, the genome sequence of *S. miltiorrhiza* was decoded [37]. Due to its small genome size, short life cycle, and simple growth requirements, *S. miltiorrhiza* is considered to be a model medicinal plant [38]. In *Arabidopsis*, miR858 is reported to regulate the flavonoid biogenesis [39, 40]. Therefore, characterization of sRNAs associated with tanshinones or phenolic acids biogenesis is of great interest. We have performed genome-wide identification of *DCL*, *RDR*, and *AGO* family genes in *S. miltiorrhiza* [41–43]. Here, we describe the procedure of *S. miltiorrhiza*

AGO gene prediction and molecular cloning, together with *Agrobacterium*-mediated *S. miltiorrhiza* transformation protocol for further *SmAGO* gene functional analysis.

2 Materials

2.1 SmAGO Prediction

1. Stand-alone BLAST package.
2. *S. miltiorrhiza* genome sequence [37].
3. Arabidopsis and rice AGO amino acid sequences [21, 44].

2.2 Plant Material and RNA Extraction

1. *S. miltiorrhiza* Bunge (Line 993) seeds.
2. TRIzol reagent (Thermo Fisher Scientific).
3. Chloroform.
4. Isopropyl alcohol.
5. 75% ethanol [in diethyl pyrocarbonate (DEPC)-treated water].
6. RNase-free water.

2.3 DNA Digestion

1. Recombinant DNase I (Takara).
2. Recombinant RNase inhibitor (40 U/μL) (Takara).
3. Phenol/chloroform/isoamyl alcohol (25:24:1) (Amresco).
4. 3 M sodium acetate (NaAc): weigh 40.824 g of $NaAc{\cdot}3H_2O$, and make up to 100 mL with DEPC-treated water.
5. Chloroform.
6. Ethanol.
7. 75% ethanol.
8. RNase-free water.

2.4 RACE and PCR Reaction

1. SMARTer™ RACE cDNA amplification kit (Clontech).
2. Premix Taq™ (Ex Taq™ Version 2.0) (Takara).
3. Gene-specific primers (GSP) and nested gene-specific primers (NGSP) (Fig. 1).
4. 5′- and 3′-SMRT sequence-specific universal primer mix (UPM) and nested universal primers (NUP) (Fig. 1).
5. Tricine-EDTA buffer: 10 mM Tricine-KOH (pH 8.5) and 1 mM EDTA.
6. pMD-18 T vector (Takara).
7. Sterile distilled water.

2.5 Reverse Transcription and Full-Length Gene Cloning

1. Superscript III reverse transcriptase (Thermo Fisher Scientific).
2. Gene forward and reverse primers (Fig. 1).
3. $Oligo(dT)_{20}$ or random primers (Takara).

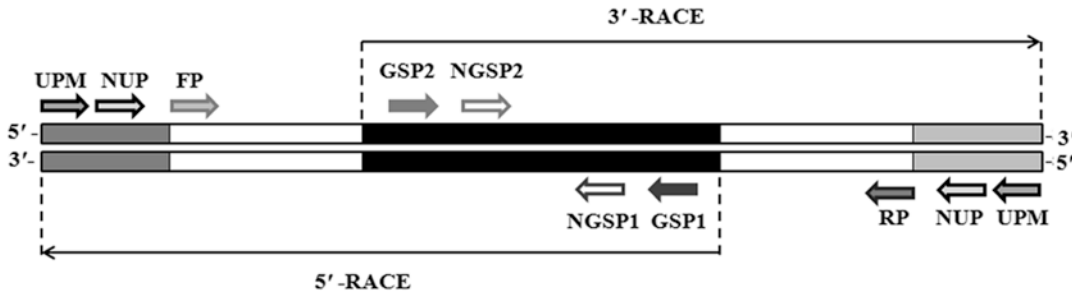


Fig. 1 The locations of primers used in RACE and full-length cDNA cloning. It shows a RNA/DNA hybrid as template. Known sequence is indicated in *black*, while unknown sequence in *white*. 5′-end oligo is marked in *dark gray*, while 3′-end in *light gray*. Primer pairs GSP1/NGSP1 and UPM/NUP are used for 5′-end unknown sequence amplification, and GSP2/NGSP2 and UPM/NUP are used for 3′-end unknown sequence amplification. Forward primer (FP) and reverse primer (RP) are used for full-length cDNA cloning

4. dNTP Mix (10 mM each) (Takara).
5. Recombinant RNase inhibitor (40 U/μL) (Takara).
6. Premix Taq™ (Ex Taq™ Version 2.0) (Takara).
7. pMD-18 T vector (Takara).
8. Sterile distilled water.

2.6 *S. miltiorrhiza* Transformation

1. Murashige and Skoog basal medium. We use the commercial MS powder including salts and vitamins (Phytotechnology Lab).
2. *Agrobacterium tumefaciens* strain GV3101.
3. Vectors derived from conventional pCAMBIA series.
4. YEB medium: 5 g/L beef extract, 1 g/L of yeast extract, 5 g/L of peptone, 0.5 g/L magnesium sulfate heptahydrate ($MgSO_4 \cdot 7H_2O$), and 5% of sucrose, pH 7.2. For solid plates, add 15 g/L Bacto agar before autoclaving.
5. Media for transformation is listed in Table 1. All media contain 2% sucrose and 0.7% agar and have a pH of 5.8 adjusted with 2 M sodium hydroxide (NaOH) before being autoclaved at 118 °C for 20 min.
6. 1 mg/mL 6-benzylaminopurine (6-BA, Sigma-Aldrich): first dissolve in a small volume of 2 M NaOH and then adjust with water.
7. 1 mg/mL 1-naphthaleneacetic acid (NAA, Sigma-Aldrich): first dissolve in a small volume of ethanol, and then adjust with water.
8. Cefotaxime (300 mg/mL, Amresco).
9. Hygromycin (50 mg/mL, Roche).
10. Polyethylene pots.
11. Nonwoven fabrics.

Table 1
Media used in *S. miltiorrhiza* transformation

	Medium	Growth regulators	Antibiotics
Cocultivation medium (CM)	MS	1.0 mg/L 6-BA 0.1 mg/L NAA	–
Shoot induction (SI)	MS	1.0 mg/L 6-BA 0.1 mg/L NAA	50 mg/L hygromycin 300 mg/L cefotaxime
Plantlet growth (PG)	MS	–	50 mg/L hygromycin 300 mg/L cefotaxime
Plantlet rooting (PR)	1/2 MS	0.2 mg/L IBA	50 mg/L hygromycin 300 mg/L cefotaxime

2.7 Equipment

1. NanoDrop 2000C Spectrophotometer (Thermo Fisher Scientific).
2. Bio-Rad CFX system (Bio-Rad).
3. Microfuge (Eppendorf).
4. DNA Engine Tetrad 2 thermal cycler (Bio-Rad).
5. Shaking incubator (Taicang).
6. Laminar flow bench (Suzhou Purification Equipment Co.).
7. Controlled growth chamber.

3 Methods

3.1 SmAGO Gene Prediction

The presented procedure is used for the genome-wide prediction of *SmAGO* family genes. Since we will identify gene sequences with high homology to known AGO proteins in *S. miltiorrhiza* genome, stand-alone BLAST should be used.

1. Download the stand-alone BLAST program package from ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/. Place it in a desired directory, named *blast*. Note that there are different packages for Linux, Mac, or Windows. Choose the appropriate file for your computer platform.
2. Install the BLAST program in the *blast* directory following the instructions (https://www.ncbi.nlm.nih.gov/books/NBK279690/). For Linux, simply extract the downloaded package typing the following tar command: tar zxvpf ncbi–blast + 2.2.17–x64–linux.tar.gz (*see* **Note 1**).
3. Download 10 Arabidopsis and 19 rice AGO protein sequences from NCBI protein database. Put them in one text document file in FASTA format, named AGO.fasta. They will be used as query sequences to perform BLAST.

4. Download a draft genome sequence of *S. miltiorrhiza* [37]. Right-click on the Salvia.scaffold.fasta file and save it in your local disk. Make sure to download the latest release if there is any.
5. Place *S. miltiorrhiza* genome database file (Salvia.scaffold.fasta) and the file containing the query sequences (AGO.fasta) in a new folder (named SmAGO_prediction).
6. Open the terminal and access the folder (in our case, /home/Lu305/SmAGO_prediction). Note that Lu305 is the username for the computer used in this analysis. Type the following command to format the genome database into a searchable database before using BLAST: formatdb -i Salvia.scaffold.fasta -p F. -i indicates the input file. -p asks if the input data is protein sequence. Use F (false) for nucleotide sequences, otherwise use T (true). There will be three new index files (Salvia.scaffold.fasta.nhr, Salvia.scaffold.fasta.nin, Salvia.scaffold.fasta.nsq) and one formatdb.log file generated (*see* **Note 2**).
7. Run a BLAST search by typing the following command: blastall -p tblastn -d Salvia.scaffold.fasta -i AGO.fasta -o SmAGO.m8 -e 0.0000000001 -a 4 (*see* **Note 3**). Here, the program tblastn should be used, since we will search translated nucleotide database using protein sequences as query. -d specifies the database (Salvia.scaffold.fasta), -i indicates the query file (AGO.fasta), and -o is the desired name of the output file. An e-value cutoff of 10^{-10} is applied as specified by -e. The blast process usually takes a while, so you can specify several processors by using -a.
8. Put all identified scaffold IDs in one text document, named ID.txt, with one ID in a row. Extract all the aligned scaffold sequences through a perl script by typing a command: perl extract_seq.pl ID.txt Salvia.scaffold.fasta. Thus, a file named result_seq.fa is produced.
9. Get access to BLASTX online website: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab. Enter scaffold sequences and search for nonredundant protein sequences (nr). By comparing with AGO amino acids from other plant species, predict the positions of each exon and intron in the scaffold (*see* **Note 4**).
10. For those scaffolds without full-length *AGO* gene, we can use RNA-seq data to predict their 5′ or 3′ sequences further. Get access to NCBI website (https://www.ncbi.nlm.nih.gov/); choose SRA database in the left dropdown menu. Input *Salvia miltiorrhiza* in the blank box and click search. Then, all *S. miltiorrhiza* high-throughput sequencing data is output. Record the accession numbers of RNA-seq data sequenced by 454

GS FLX Titanium or Illumina. Afterward, enter the blastn online website (https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome), and input partial sequences (about 100 nt) as query to search the *S. miltiorrhiza* RNA-seq data. Here, choose SRA in the database dropdown menu, and input the accession numbers recorded above. The deduced sequence can be used as query to search the *S. miltiorrhiza* RNA-seq data again. Retrieve the longest sequences if possible (*see* **Note 5**).

3.2 Molecular Cloning of SmAGO Genes

All centrifugations use a microfuge.

3.2.1 Plant Growth and RNA Extraction

1. Grow *S. miltiorrhiza* plants under natural growth conditions (e.g., in a field nursery at the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China).
2. Collect roots, stems, leaves, and flowers, and store in liquid nitrogen until use.
3. Plant tissues (50–100 mg) are ground to powder under liquid nitrogen, and add 1 mL of TRIzol (*see* **Note 6**).
4. Incubate the homogenized samples at room temperature for 5 min.
5. Add 0.2 mL of chloroform per 1 mL of TRIzol. Shake tubes vigorously by hand for 15 s, and incubate them at room temperature for 3 min. Centrifuge at 10,625 × *g* for 15 min at 4 °C.
6. Transfer the aqueous phase to a fresh tube (~600 μL).
7. Add 500 μL of isopropyl alcohol. Incubate at room temperature for 10 min and then centrifuge at 10,625 × *g* for 10 min at 4 °C.
8. Remove the supernatant. Wash the RNA pellet with 1 mL of 75% ethanol. Mix the sample by vortexing and centrifuge at 10,625 × *g* for 5 min at 4 °C.
9. Briefly dry the RNA pellet in air for 5–10 min. Dissolve RNA in RNase-free water by gently pipetting. Place at −80 °C for longtime storage (*see* **Note 7**).

3.2.2 DNA Digestion

Remove possible genomic DNA contamination by treating the isolated RNA with DNase I. All centrifugations use a microfuge.

1. Add the following components to a RNase-free microcentrifuge tube:

Total RNA	<50 μg
10× DNase I buffer	5 μL
Recombinant DNase I (RNase-free, 5 U/μL)	2 μL
Recombinant RNase inhibitor (40 U/μL)	0.5 μL
RNase-free water	up to 50 μL

2. Incubate at 37 °C for 30 min.
3. Add 50 μL of RNase-free water.
4. Add 100 μL phenol/chloroform/isoamyl alcohol (25:24:1). Shake tubes vigorously by hand. Centrifuge at 10,625 × *g* for 5 min.
5. Transfer the aqueous phase to a fresh tube. Add the same volume of chloroform. Centrifuge at 10,625 × *g* for 5 min.
6. Transfer the aqueous phase to a fresh tube. Add 10 μL of 3 M NaAc and 250 μL of cold ethanol. Shake the tube by hand and place at −80 °C for 20 min.
7. Centrifuge at 10,625 × *g* for 10 min.
8. Remove the supernatant. Wash the RNA pellet with 1 mL of 75% ethanol, and dissolve in RNase-free water as above.

3.2.3 Preparation of 5′- and 3′-RACE-Ready cDNA

1. Prepare enough of the following buffer mix for all of the 3′- and 5′-RACE-ready cDNA synthesis reactions:

5× first-strand buffer	4 μL
DTT (100 mM)	0.5 μL
dNTP mix (20 mM)	1 μL

2. Combine the following in separate 0.5 mL microcentrifuge tubes (100 ng–1 μg of total RNA):

   For 5′-RACE-ready cDNA:

RNA sample	<10 μL
5′-CDS primer (12 μM)	1 μL
RNase-free water	up to 11 μL

   For 3′-RACE-ready cDNA:

RNA sample	<11 μL
3′-CDS primer (12 μM)	1 μL
RNase-free water	up to 12 μL

3. Mix contents by pipetting; spin the tubes briefly.
4. Incubate the tubes at 72 °C for 3 min, and then at 42 °C for 2 min. Spin the tubes briefly to collect the contents at the bottom.

5. Add 1 μL SMARTer II A oligo to the 5′-RACE cDNA synthesis reaction.
6. Prepare enough of the following master mix for all of the 5′- and 3′-RACE cDNA synthesis reactions:

Buffer mix from **step 1**	5.5 μL
RNase inhibitor (40 U/μL)	0.5 μL
SMARTScribeTM reverse transcriptase (100 U)	2 μL

7. Add 8 μL of the master mix from **step 6** to the 5′-RACE (**step 5**) and 3′-RACE (**step 2**) cDNA synthesis reactions, for a total volume of 20 μL.
8. Mix contents by gently pipetting, and spin the tubes briefly to collect the contents at the bottom.
9. Incubate the tubes at 42 °C for 90 min, and then at 70 °C for 10 min in a thermocycler (*see* **Note 8**).
10. Dilute the first-strand reaction product with Tricine-EDTA buffer: add 20 μL if you started with <200 ng of total RNA, or add 100 μL if you started with >200 ng of total RNA.
11. Store samples at −20 °C.

3.2.4 PCR Reaction for 5′- and 3′-Unknown Sequence Cloning

1. Add the following contents to a PCR reaction tube:

Premix Taq™ (Ex Taq™ Version 2.0)	25 μL
cDNA (from first-strand reaction)	2.5 μL
GSP primer (10 μM) (*see* **Note 9**)	1 μL
10× UPM primer	5 μL
Double distilled water	16.5 μL

2. Perform PCR using the following touchdown PCR program:

   94 °C 2 min.

   94 °C 30 s and 72 °C 3 min for 5 cycles.

   94 °C 30 s, 70 °C 30 s, and 72 °C 3 min for 5 cycles.

   94 °C 30 s, 68 C 30 s, and 72 °C 3 min for 25 cycles.

   72°C 10 min (*see* **Note 10**)
3. If there is a smear or no PCR bands, perform nested PCR reaction using NUP and NGSP primers. Dilute 5 μL:

   l of the primary PCR product into 245 μL.

   l of sterile distilled water.

Add the following to a PCR reaction tube:

Premix Taq™ (Ex Taq™ Version 2.0)	25 μL
Diluted primary PCR product	5 μL
NGSP primer (10 μM)	1 μL
NUP primer (10 μM)	1 μL
Sterile distilled water	18 μL

4. Perform PCR Using the Following Program:

   94 °C 2 min.

   94 °C 30 s, 58 °C 30 s, and 72 °C 2 min for 30 cycles.

   72°C 10 min.

5. PCR products are purified, cloned into pMD-18 T vector, and sequenced.

3.2.5 Reverse Transcription

Total RNA is extracted using TRIzol as described. Then, ~5 μg total RNA is used to synthesize first-strand cDNA for full-length gene cloning (*see* **Note 11**).

1. Add the following components to a RNase-free microcentrifuge tube:

Oligo$(dT)_{20}$ (50 μM) or random primers	1 μL
RNA sample	~5 μg
dNTP Mix (10 mM each)	1 μL
Sterile distilled water	up to 13 μL

2. Heat at 65 °C for 5 min and incubate on ice for at least 1 min.
3. Spin the tubes briefly to collect the contents at the bottom, and add:

5 × first-strand buffer	4 μL
0.1 M DTT	1 μL
Recombinant RNase inhibitor (40 U/μL)	1 μL
SuperScript™ III RT (200 U/μL)	1 μL

4. Mix the contents by gently pipetting.
5. Incubate at 50 °C for 60 min and then 70 °C for 15 min in a thermocycler if using oligo(dT) primers. For random primers, first incubate at 25 °C for 5 min and then at 50 °C for 60 min and 70 °C for 15 min.
6. Store samples at −20 °C.

3.2.6 PCR Reaction for Full-Length Sequence Cloning

1. Add the following components to a RNase-free microcentrifuge tube:

Premix Taq™ (Ex Taq™ Version 2.0)	25 μL
cDNA (from first-strand reaction)	1–2 μL
Forward primer (10 μM)	1 μL
Reverse primer (10 μM)	1 μL
Double distilled water	up to 50 μL

2. Perform PCR using as follows:

   94 °C 2 min.

   94 °C 30 s, 56 °C 30 s, and 72 °C 2 min for 30 cycles.

   72°C 10 min.

3. Purify PCR products, clone them into pMD-18 T vector, and sequence.

3.3 Establishment of an Agrobacterium-Mediated Transformation System

To perform gene functional studies in *S. miltiorrhiza*, it is first necessary to establish a reliable and efficient genetic transformation protocol [45–48]. Next is the protocol for genetically transforming *S. miltiorrhiza* using *A. tumefaciens*.

3.3.1 Preparation of Agrobacterium Culture

1. Inoculate a freshly grown single colony of *Agrobacterium* GV3101 containing the binary vector pCAMBIA series in 2 mL of YEB liquid medium supplemented with 50 mg/L of kanamycin and 25 mg/L of rifampicin. The culture is incubated at 28 °C for 24 h in a gyratory shaker (180 rpm) (*see* **Note 12**).
2. Inoculate 100 μL of *Agrobacterium* culture into 10 mL of YEB liquid medium containing the same antibiotics as above. The culture tube is incubated at 28 °C in a gyratory shaker (180 rpm) until OD_{600} = 0.8.
3. Transfer 2 mL of *Agrobacterium* culture into Eppendorf tubes, and centrifuge at 1,844 × *g* for 10 min in a microfuge.
4. Resuspend the *Agrobacterium* cells in 30–40 mL of sterile CM liquid medium for transformation.

3.3.2 Cocultivation (48–72 h)

1. Cut leaves of sterile *S. miltiorrhiza* plants into 0.5 × 0.5 cm disks and pre-culture for 1 day on CM medium.
2. Dip for 10–15 min (with shaking) the pre-cultured leaf explants into the bacterial liquid medium obtained above.
3. Remove the bacterial solution with a pipette. Blot dry the disks with sterilized filter paper to remove the bacteria excess.

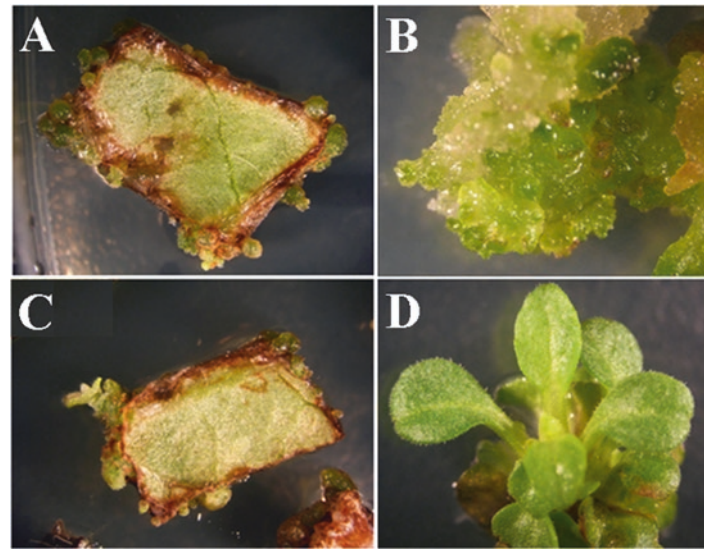


Fig. 2 Production of transgenic *S. miltiorrhiza* plants. (**a**, **b**) Resistant callus on SI medium with 50 mg/L hygromycin. (**c**, **d**) Regenerated shoots

4. Transfer the explants to CM medium (*see* **Note 13**).
5. Seal the plates with Parafilm, and incubate in the dark at 24 °C for 2 days. Gene transfer into plant genome occurs in this step (*see* **Note 14**).

3.3.3 Shoot Induction (30–45 days)

1. After cocultivation, transfer explants to SI medium and grow at 24 °C with a 16/8 h light/dark photoperiod. SI medium contains 300 mg/L cefotaxime to eliminate the bacteria and 50 mg/L hygromycin to select the transformed shoots (Fig. 2a, b) (*see* **Note 15**).
2. Subculture explants to new SI medium every 2 weeks until no shoots appear. Generally, early shoots occur between 10 and 14 days after the first transfer to SI medium (Fig. 2c).

3.3.4 Plantlet Growth and Rooting (30–60 days)

1. Place regenerated shoots on PG medium, and grow at 24 °C with a 16/8 h light/dark photoperiod for hardening (Fig. 2d). PG medium contains no phytohormones but 300 mg/L of cefotaxime to avoid the bacteria grown again. It also contains 50 mg/L of hygromycin for continuous selection.
2. Transfer strong shoots to PR medium, and grow under the same conditions for root formation (*see* **Note 16**).

3.3.5 Plant Transfer to Soil

1. Transfer plantlets with well-developed leaves and roots to non-woven fabrics or polyethylene pots containing perlite and nutrient soil mixture (1:1 v/v). Water thoroughly (*see* **Note 17**).

2. Cover them with a transparent lid, and grow in a controlled chamber with a 16/8 h light/dark photoperiod at 20–22 °C. Spray water once a day during the first 3 days.
3. Open the lid after 1 week to see newly grown roots. If using nonwoven fabrics, plants should be transferred to bigger ones or polyethylene pots because new roots may go out.
4. Water plants regularly. Approximately 100% of the transferred plants should survive.

4 Notes

1. For BLAST installation on MAC or Windows platform, please refer to the BLAST Command Line Applications User Manual (https://www.ncbi.nlm.nih.gov/books/NBK279690/). In this case, we use an early BLAST version. However, we recommend using the most recent available version.
2. For different BLAST versions in Linux or on MAC and Windows platforms, format change command might be changed. Please refer to the BLAST manual (https://www.ncbi.nlm.nih.gov/books/NBK279690/) carefully.
3. Similarly, BLAST command may be changed for different versions or on MAC and Windows platforms.
4. It is requisite for manual manipulation in this step, since scaffolds with very short sequences aligned will be output. Actually, they are not the authentic *SmAGO* genes.
5. In some cases, we cannot predict the full-length sequences of all genes due to incomplete genome assembly. Generally, 454 GS FLX Titanium sequencing generates the long read length (~500 bp), while Illumina sequencing reads are short (~100 bp). Therefore, it is better to use 454 data first and then Illumina data to predict the unknown sequences. Also, it should be noticed that there might be prediction errors, since short Illumina reads might match to more than one position. Hence, it is necessary to clone the sequences to confirm the predicted sequences.
6. We strongly recommend examining the gene tissue-specific expression by RT-qPCR before conducting RACE. Using the tissue with high expression level of the desired gene may lead to better results.
7. Do not dry the RNA completely, since this will greatly decrease its solubility.
8. The volume of the reaction mixture may be reduced due to evaporation during incubation, which will affect the first strand synthesis efficiency. Using a thermocycler with a hot lid may prevent evaporation to some extent.

9. GSP or NGSP primers should be 23–28 nt, with a 50–70% GC content, a Tm >65 °C and not complementary to the UPM primers. However, the best Tm is >70 °C, when using touchdown PCR grogram. For the gene-specific primer relationship with cDNA template, please see Fig. 1.
10. Set extension time based on the expected fragment length. If fragments are 0.2–2 kb, 2 min is recommended. Add 1 min for each additional 1 kb.
11. Similarly, using the tissue with high expression level of the desired gene will lead to better results.
12. *S. miltiorrhiza* was successfully transformed using *Agrobacterium* EHA105 harboring vector pCAMBIA series in two reports [45–47] or pPAL1 [48]. We also obtained transgenic *S. miltiorrhiza* plants using *Agrobacterium* GV3101 and a pCAMBIA1391-derived vector. Therefore, there is no specific requirement for *Agrobacterium* strains and binary plasmids.
13. Successful transformation can be achieved without addition of acetosyringone during cocultivation stage.
14. Check the plates carefully to avoid bacteria overgrown on the explants.
15. Selection is a vital step for successful transformation. Transgenic *S. miltiorrhiza* plants have been successfully obtained on selection medium supplemented with 30 or 60 mg/L of kanamycin [46, 48] or 0.6 mg/L phosphinothricin [47]. However, we find that hygromycin selection is much more effective than kanamycin treatment to select the transgenic *S. miltiorrhiza* (Line 993) plants. In this case, we use a higher concentration of hygromycin (50 mg/L), and almost 100% of the regenerated plants are positive. According to our experience, selection with 30 mg/L of hygromycin could also lead the desired transformation efficiency.
16. Plant roots might be impaired with 300 mg/L of cefotaxime in medium. In this case, use lower concentration of cefotaxime.
17. Plantlet accommodation is better when using the breathable nonwoven fabrics.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (31370327, 31570667), the Beijing Natural Science Foundation (5152021), the CAMS Innovation Fund for Medical Sciences (CIFMS) (2016-I2M-3-016), and the Fund for Postgraduate Innovation in Peking Union Medical College (2015-1007-13).

References

1. Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. Science 301(5631):336–338. doi:10.1126/science.1085242
2. Hake S (2003) MicroRNAs: a role in plant development. Curr Biol 13(21):R851–R852. doi:10.1016/j.cub.2003.10.021
3. Kidner CA, Martienssen RA (2005) The developmental role of microRNA in plants. Curr Opin Plant Biol 8(1):38–44. doi:10.1016/j.pbi.2004.11.008
4. Lu S, Sun YH, Chiang VL (2008) Stress-responsive microRNAs in Populus. Plant J 55(1):131–151. doi:10.1111/j.1365-313X.2008.03497.x
5. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science 312(5772):436–439. doi:10.1126/science.1126088
6. Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell 16(8):2001–2019. doi:10.1105/tpc.104.022830
7. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297. doi:10.1016/S0092-8674(04)00045-5
8. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC (2010) Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat Struct Mol Biol 17(8):997–1111. doi:10.1038/nsmb.1866
9. Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell 123(7):1279–1291. doi:10.1016/j.cell.2005.11.035
10. Mallory AC, Vaucheret H (2006) Functions of microRNAs and related small RNAs in plants. Nat Genet 38(Suppl):S31–S36. doi:10.1038/ng1791
11. Chen X (2010) Small RNAs - secrets and surprises of the genome. Plant J 61(6):941–958. doi:10.1111/j.1365-313X.2009.04089.x
12. Chen X (2009) Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol 25:21–44. doi:10.1146/annurev.cellbio.042308.113417
13. 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
14. Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4):669–687. doi:10.1016/j.cell.2009.01.046
15. Johnston M, Hutvagner G (2011) Posttranslational modification of Argonautes and their role in small RNA-mediated gene regulation. Silence 2:5. doi:10.1186/1758-907X-2-5
16. Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53. doi:10.1146/annurev.arplant.57.032905.105218
17. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190. doi:10.1126/science.1159151
18. Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA methylation mediated by a microRNA pathway. Mol Cell 38(3):465–475. doi:10.1016/j.molcel.2010.03.008
19. Tolia NH, Joshua-Tor L (2007) Slicer and the argonautes. Nat Chem Biol 3(1):36–43. doi:10.1038/nchembio848
20. Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9(1):22–32. doi:10.1038/nrm2321
21. Zhang H, Xia R, Meyers BC, Walbot V (2015) Evolution, functions, and mysteries of plant ARGONAUTE proteins. Curr Opin Plant Biol 27:84–90. doi:10.1016/j.pbi.2015.06.011
22. Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13(7):350–358. doi:10.1016/j.tplants.2008.04.007
23. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/j.cell.2008.02.034
24. Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A 102(33):11928–11933. doi:10.1073/pnas.0505461102
25. Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X (2011) Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell 145(2):242–256. doi:10.1016/j.cell.2011.03.024

26. Zhang Z, Zhang X (2012) Argonautes compete for miR165/166 to regulate shoot apical meristem development. Curr Opin Plant Biol 15(6):652–658. doi:10.1016/j.pbi.2012.05.007
27. Zhou Y, Honda M, Zhu H, Zhang Z, Guo X, Li T, Li Z, Peng X, Nakajima K, Duan L, Zhang X (2015) Spatiotemporal sequestration of miR165/166 by Arabidopsis Argonaute10 promotes shoot apical meristem maintenance. Cell Rep 10(11):1819–1827. doi:10.1016/j.celrep.2015.02.047
28. Tucker MR, Okada T, Hu Y, Scholefield A, Taylor JM, Koltunow AM (2012) Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in Arabidopsis. Development 139(8):1399–1404. doi:10.1242/dev.075390
29. Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC (2012) Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. Plant Cell 24(9):3613–3629. doi:10.1105/tpc.112.099945
30. Carbonell A, Carrington JC (2015) Antiviral roles of plant ARGONAUTES. Curr Opin Plant Biol 27:111–117. doi:10.1016/j.pbi.2015.06.013
31. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133(1):128–141. doi:10.1016/j.cell.2008.02.033
32. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC (2010) The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22(2):321–334. doi:10.1105/tpc.109.072199
33. Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature 464(7288):628–632. doi:10.1038/nature08828
34. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol 49(4):493–500. doi:10.1093/pcp/pcn043
35. Cheng TO (2006) Danshen: a popular chinese cardiac herbal drug. J Am Coll Cardiol 47(7):1498.; author reply 1499-1500. doi:10.1016/j.jacc.2006.01.001
36. He S, Yang Y, Liu X, Huang W, Zhang X, Yang S, Zhang X (2012) Compound Astragalus and salvia miltiorrhiza extract inhibits cell proliferation, invasion and collagen synthesis in keloid fibroblasts by mediating transforming growth factor-beta/Smad pathway. Br J Dermatol 166(3):564–574. doi:10.1111/j.1365-2133.2011.10674.x
37. Xu H, Song J, Luo H, Zhang Y, Li Q, Zhu Y, Xu J, Li Y, Song C, Wang B, Sun W, Shen G, Zhang X, Qian J, Ji A, Xu Z, Luo X, He L, Li C, Sun C, Yan H, Cui G, Li X, Li X, Wei J, Liu J, Wang Y, Hayward A, Nelson D, Ning Z, Peters RJ, Qi X, Chen S (2016) Analysis of the genome sequence of the medicinal plant salvia miltiorrhiza. Mol Plant 9(6):949–952. doi:10.1016/j.molp.2016.03.010
38. Chen L, Song J, Sun C, Xu J, Zhu Y, Verpoorte R, Fan T-P (2015) Herbal genomics: examining the biology of traditional medicines. Science 347:S27–S29
39. Sharma D, Tiwari M, Pandey A, Bhatia C, Sharma A, Trivedi PK (2016) MicroRNA858 is a potential regulator of Phenylpropanoid pathway and plant development. Plant Physiol 171(2):944–959. doi:10.1104/pp.15.01831
40. Wang Y, Wang Y, Song Z, Zhang H (2016) Repression of MYBL2 by both microRNA858a and HY5 leads to the activation of anthocyanin biosynthetic pathway in Arabidopsis. Mol Plant 9(10):1395–1405. doi:10.1016/j.molp.2016.07.003
41. Shao F, Qiu D, Lu S (2015) Comparative analysis of the Dicer-like gene family reveals loss of miR162 target site in SmDCL1 from salvia miltiorrhiza. Sci Rep 5:9891. doi:10.1038/srep09891
42. Shao F, Lu S (2014) Identification, molecular cloning and expression analysis of five RNA-dependent RNA polymerase genes in salvia miltiorrhiza. PLoS One 9(4):e95117. doi:10.1371/journal.pone.0095117
43. Shao F, Lu S (2013) Genome-wide identification, molecular cloning, expression profiling and posttranscriptional regulation analysis of the Argonaute gene family in salvia miltiorrhiza, an emerging model medicinal plant. BMC Genomics 14:512. doi:10.1186/1471-2164-14-512
44. Kapoor M, Arora R, Lama T, Nijhawan A, Khurana JP, Tyagi AK, Kapoor S (2008) Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene

families and their expression analysis during reproductive development and stress in rice. BMC Genomics 9:451. doi:10.1186/1471-2164-9-451

45. Lee CY, Agrawal DC, Wang CS, SM Y, Chen JJ, Tsay HS (2008) T-DNA activation tagging as a tool to isolate salvia miltiorrhiza transgenic lines for higher yields of tanshinones. Planta Med 74(7):780–786. doi:10.1055/s-2008-1074527
46. Yan Y, Wang Z (2007) Genetic transformation of the medicinal plant salvia miltiorrhiza by Agrobacterium tumefaciens-mediated method. Plant Cell Tiss Org Cult 88(2):175–184. doi:10.1007/s11240-006-9187-y
47. Liu Y, Yang SX, Cheng Y, Liu DQ, Zhang Y, Deng KJ, Zheng XL (2015) Production of herbicide-resistant medicinal plant salvia miltiorrhiza transformed with the bar gene. Appl Biochem Biotechnol 177(7):1456–1465. doi:10.1007/s12010-015-1826-5
48. Song J, Wang Z (2011) RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in salvia miltiorrhiza causes abnormal phenotypes and a reduction in rosmarinic acid biosynthesis. J Plant Res 124(1):183–192. doi:10.1007/s10265-010-0350-5

Chapter 13

Cloning and Characterization of *Argonaute* Genes in Tomato

Zhiqiang Xian, Fang Yan, and Zhengguo Li

Abstract

Argonaute (AGO) proteins are core elements in plant posttranscriptional RNA silencing pathways. The identification and functional characterization of tomato (*Solanum lycopersicum*) AGOs will help to better understand RNA silencing-based posttranscriptional regulation in fleshy fruits. Here we describe how to identify and clone *SlAGO* genes, as well as the methodology for their functional characterization.

Key words Argonaute, *Solanum lycopersicum*

1 Introduction

Small RNAs regulate multiple plant developmental and physiological processes including organ polarity determination and leaf and floral development, frequently controlled by microRNAs (miRNAs) [1–5]. In eukaryotes, miRNAs, small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs), scanRNAs, or 21 U-RNAs are produced. These types of small RNAs have been shown to associate with different members of the AGO family including AGO, PIWI, and group 3 proteins [6–9].

Solanum lycopersicum (tomato) is the model plant used to study fleshy fruit development. The molecular cloning [10] and the identification of the biological function of SlAGOs have been reported [11–13]. In addition, it was described that the overexpression of the AGO1-binding *Polerovirus* P0 silencing suppressor protein in tomato inhibits the expression of both *SlAGO1A* and *SlAGO1B* resulting in a dramatic modification of the radicalization of leaflets, petals, and anthers [14]. In any case, the specific mechanisms by which small RNAs function in tomato in processes such as fruit formation and development still remain largely unknown. In this chapter, we present protocols for the molecular cloning and

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_13,

for the expression and subcellular localization analyses of tomato *AGO* genes. These types of analyses should help to better understand the roles of plant AGOs in fruit development.

2 Materials

2.1 Plants, Bacteria, and Vectors

1. Mature seeds of the miniature *Lycopersicon esculentum* cultivar, Micro-Tom (Micro tomato) [15].
2. Tobacco Bright Yellow 2 Cell Line, BY2 [16].
3. *Agrobacterium tumefaciens* strain GV3101 or C58 (with pSoup helper plasmid) [17].
4. pLP100 base plasmids and pGreenII–eGFP [18] carrying the genes of interest.

2.2 Transient Expression in BY2 Protoplasts

1. MES buffer: 1% (w/v) bovine serum albumin (BSA), 20 mM potassium chloride (KCl), 20 mM 2-(N-morpholino)ethanesulfonic acid (MES), 10 mM calcium chloride ($CaCl_2$), pH 5.7, and filter the solution through a 0.45 μM syringe filter.
2. Enzyme solution: 1% (W/V) caylase, 0.2% (w/v) macerozyme R-10, 1% (w/v) BSA, 20 mM potassium chloride (KCl), 20 mM MES, 10 mM $CaCl_2$, pH 5.7.
3. Polyethylene glycol (PEG) solution: 4 g of PEG 4000, 200 mM mannitol, 100 mM of $CaCl_2$, add H_2O to 10 mL.
4. Washing and incubation solution (WI): 0.5 M mannitol, 20 mM KCl, 4 mM MES (pH 5.7).
5. W5 buffer: 154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES (pH 5.7).
6. MMg solution: 0.4 M mannitol, 15 mM $MgCl_2$, 4 mM MES (pH 5.7).

2.3 Tomato Transformation

1. 70% ethanol.
2. 5% potassium chlorate ($KClO_3$).
3. Media for seed germination (1/2 MS): 1/2 MS (20 g MS powder to 1 L, Sigma-Aldrich), R3 vitamins (thiamine, 1 g/L; nicotinic acid, 0.5 g/L; pyridoxine, 0.5 g/L; 50 μL to 100 mL 1/2 MS), agar 8 g/L, pH 5.9.
4. Pre-cultivation and cocultivation medium (KCMS solid): 4.4 g/L MS, 20 g/L sucrose, 200 mg/L of monopotassium phosphate (KH_2PO_4), 8 g/L agar, 0.9 mg/L thiamine, 200 μM acetosyringone, 0.2 mg/L 2,4-D, 0.1 mg/L kinetin, pH 5.8.
5. *Agrobacterium* resuspension and infection medium (KCMS liquid): 4.4 g/L MS, 20 g/L sucrose, 200 mg/L KH_2PO4,

8 g/L agar, 0.9 mg/L thiamine, 200 μM acetosyringone, pH 5.7.

6. Selective medium (2Z or 1Z): 4.4 g/L MS, 30 g/L sucrose, 8 g agar, 1 mL Vit Nitsch (1000×), 2 mg/L or 1 mg/L zeatin riboside, 100 mg/L kanamycin, 100 mg/L Augmentin and 100 mg/L Timentin, pH 5.8.
7. Rooting medium (ENR): 2.2 g/L MS, 10 g/L sucrose, 8 g agar, 1 mL Vit Nitsch (1000×), 75 mg/L kanamycin, 100 mg/L Augmentin and 100 mg/L Timentin, pH 5.8.
8. Solid LB with antibiotics: 25 mg/L rifampicin, 50 mg/L gentamicin, 100 mg/L kanamycin (50 mg/L tetracycline should be added if using C58).
9. Liquid LB with antibiotics: 25 mg/L rifampicin, 50 mg/L gentamicin, 100 mg/L kanamycin (50 mg/L Tet should be added if using C58).

2.4 Other Materials or Equipment

1. Parafilm.
2. Nylon mesh.
3. Laser scanning confocal microscope.

2.5 PCR and Cloning

1. TRIzol reagent (Thermo Fisher Scientific).
2. Reverse transcription kit (Thermo Fisher Scientific).
3. Pfu DNA polymerase (Transgene).
4. Agarose gel and equipment (Bio-Rad).
5. T4 DNA ligase (NEB).
6. Primers.
7. Full-length amplification kit (Takara).
8. pEASY-Blunt cloning kit (Transgene).
9. Restriction enzymes (Thermo Fisher Scientific).
10. DNA extraction kit (Omega).
11. *E. coli* competent cells T1 (Transgene).

3 Methods

3.1 Cloning of Full-Length AGO Genes

1. Obtain raw sequences of each *SlAGO* gene from https://solgenomics.net/ by blasting sequences from *Arabidopsis* AGO genes.
2. Further analyze raw sequences of each *SlAGO* gene with online tools such as Blastx (NCBI) or Pfam to check the conserved domains (CDs) and open reading frames.
3. Design gene-specific primers to amplify the coding domains of each gene.

4. Extract total RNA from tissues of whole flowering tomato and fruiting tomato using TRIzol reagent according to manufacturer's instructions.
5. Follow the reverse transcription kit instructions to produce cDNA using mixed RNA as template.
6. Set up a standard PCR reaction by mixing the following: 1 μL of pfu (5 U), 10 μL of 5× reaction buffer, 4 μL of dNTP (2.5 mM), 1 μL of cDNA (about 500 ng), 0.5 μM of each primer. Add H_2O to 50 μL.
7. Extract the DNA fragments from agarose gel after electrophoretic analysis.
8. One microliter of gel-purified product was mixed with pEASY-Blunt vector for 5 min at room temperature (16–25 °C).
9. Transfer the reaction products into 20 μL of *E. coli* T1 competent cells, and proceed to transformation according to manufacturer's instructions.
10. Screen colonies by PCR and send positive clones for sequencing.
11. Analyze sequencing results. Design 2–3 gene-specific primers for 5′ RACE and 3′ RACE. Specifically, use reverse primers for 5′ RACE and a forward primer for 3′ RACE (*see* **Note 1**).
12. Amplify the 5′ untranslated region and 3′ untranslated region by nested PCR using cDNA and added adaptors for 5′ RACE or 3′ RACE.
13. Clone the products into pEASY-Blunt as per manufacturer's instructions, and sequence the positive clones. Use multiple sequence alignments to obtain the full length of each gene.

3.2 Cloning of AGO Genes for Cellular Localization Analysis or Stable Transformation

1. For each *AGO* gene, amplify the coding region without stop codon for cellular localization analysis or the full length with stop codon for stable transformation.
2. Digest pLP100–35S and pGreenII–eGFP with appropriate restriction enzymes, and digest the fragments with the same restriction enzymes (*see* **Note 2**).
3. Prepare the ligation reaction by mixing the following components: 1 μL of T4 DNA ligase, 1 μL of 10× T4 DNA ligase buffer, vector/fragment as 5:1 (mol:mol, *see* **Note 3**), and add H_2O to 10 μL and then incubate at 16 °C for 3 h or overnight at 4 °C.
4. Transfer ligated products to *E. coli*, and send plasmids extracted from PCR-positive colonies for sequencing.

3.3 Transformation of BY2 Protoplasts

1. Centrifuge 2 g of 8-day-old BY2 cells at 3500 × *g* for 15 min, and then wash twice with MES buffer.

2. Digest using the enzyme solution for 1 h at 37 °C.
3. Centrifuge the digested protoplasts at 100 × *g* for 10 min, wash twice in W5 buffer, then filter using two layers of nylon mesh, and then wash the nylon mesh gently with W5 buffer.
4. Adjust the final protoplast concentration to 10^6/mL with MMg solution.
5. Mix 30 μg of plasmid with 0.2 mL of protoplasts and 0.2 mL PEG solution, and then incubate at room temperate for 1 h.
6. Block reaction with 500 μL of WI buffer.
7. Centrifuge at 100 × *g* for 8 min in a microfuge, then suspend with W5 buffer, and incubate at 16 °C for more than 16 h, and then the cellular localization could be observed under confocal microscope.

3.4 Tomato Transformation

1. Sterilize tomato seeds by submerging them during 10 s in 70% ethanol and during 30 min in 5% $KClO_3$. Grow seeds on 1/2 MS for about 8 days.
2. Slice each cotyledon into four pieces with a sharp blade. Discard the upper and lower sides. Grow the two middle pieces on solid KCMS for 1–2 days at 16 °C in the dark for pre-culture. All plates were sealed with Parafilm.
3. Mix plasmid and *Agrobacterium* (GV3101 or C58) competent cells on ice for 20 min, then incubate in liquid nitrogen for 5 min, and then in 38 °C water bath for another 5 min. Grow *Agrobacterium* in solid LB with antibiotics for 1–2 days at 28 °C.
4. Grow positive clones into 20 mL liquid LB with antibiotics with shaking (200–250 rpm) for 24 h at 28 °C (*see* **Note 4**).
5. Add 1 mL of *Agrobacterium* into 5 mL LB with 50 μM of acetosyringone, and shake overnight.
6. Dilute the *Agrobacterium* (OD600 = 0.05) in liquid KCMS in a 200 mL sterile bottle. Submerge the explant in the *Agrobacterium* suspension, and shake it gently by hand for 30 min in clean bench.
7. Dry the explant with sterilized bibulous paper and then put the explants on fresh KCMS for 1–2 days.
8. Transfer the explant to 2Z medium for 2 weeks, and then transfer explants to fresh 1Z medium every 2 weeks.
9. Cut seedlings into ENR medium when the stems of the explants are about 0.5–2 cm long (*see* **Note 5**).
10. Transplant the rooting seedlings into soil.
11. Extract DNA to verify the insertion of the DNA (*see* **Note 6**).

12. Collect seeds from T0, T1, and T2. Selection is on 1/2 MS with 100 mg/L of kanamycin. The seeds from homozygous T2 can be used to perform experiment (*see* **Note 7**).

4 Notes

1. The 5′ RACE primer should locate at the known 5′ terminus, and the 3′ RACE primer should locate at the known 3′ terminus sequence as closer as possible.
2. It is highly recommended to digest with *Sma* I or *Eco*R V which produces blunt ends; the vector can be treated with FastAP during the digestion, and interesting fragments can be treated with polynucleotide kinase before extraction.
3. The ratio should be 10:1 if both vector and insert are blunt-ended. Addition of 1 μL of PEG 6000 is recommended.
4. For long-term *Agrobacterium* storage, add 15% glycerol and immediately freeze in liquid nitrogen. Transfer to −80 °C freezer.
5. For positive seedlings, rooting should be observed within 1–2 weeks after transfer into ENR medium.
6. Primers to investigate the insertion should be designed taking into consideration the sequences of the binary vector. For example, design forward primer covering the promotor region and the reverse primer covering the insert or AGO fragment. This will assure that the AGO DNA fragment that is amplified corresponds to the transgene and not to the endogenous AGO gene. If the analysis is done in T0, plasmid carried by *Agrobacterium* could also interfere with the result. In this case, purify DNA by gel electrophoresis, and the control experiment should be set using at least one primer flanking the T-DNA region.
7. Theoretically, homozygous plants can be obtained in T1. However, due to the interference of abortive seedlings, it's strongly suggested to select homozygous seeds in T2. If the plasmid carries tags such as *GUS* or *GFP*, it is suggested to count the positive pollen from T1 seedlings to identify the homozygous plants, and then the T1 seeds can be used for experiments.

Acknowledgments

This work is supported by the National Research and Development Program (2016YFD0400101) and the National Natural Science Foundation of China (31501770).

References

1. Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 15(11):2730–2741. doi:10.1105/tpc.016238
2. Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Curr Biol 13(20):1768–1774. doi:10.1016/j.cub.2003.09.035
3. Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. Nature 425(6955):257–263. doi:10.1038/nature01958
4. Xie Z, Kasschau KD, Carrington JC (2003) Negative feedback regulation of dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. Curr Biol 13(9):784–789. doi:10.1016/S0960-9822(03)00281-1
5. Vaucheret H, Vazquez F, Crete P, Bartel DP (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev 18(10):1187–1197. doi:10.1101/gad.1201404
6. Tolia NH, Joshua-Tor L (2007) Slicer and the argonautes. Nat Chem Biol 3(1):36–43. doi:10.1038/nchembio848
7. Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9(1):22–32. doi:10.1038/nrm2321
8. Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC (2006) Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127(4):747–757. doi:10.1016/j.cell.2006.09.033
9. Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13(7):350–358. doi:10.1016/j.tplants.2008.04.007
10. Xian Z, Yang Y, Huang W, Tang N, Wang X, Li Z (2013) Molecular cloning and characterisation of SlAGO family in tomato. BMC Plant Biol 13:126. doi:10.1186/1471-2229-13-126
11. Xian Z, Huang W, Yang Y, Tang N, Zhang C, Ren M, Li Z (2014) miR168 influences phase transition, leaf epinasty, and fruit development via SlAGO1s in tomato. J Exp Bot 65(22):6655–6666. doi:10.1093/jxb/eru387
12. Zhang C, Xian Z, Huang W, Li Z (2015) Evidence for the biological function of miR403 in tomato development. Sci Hortic 197:619–626. doi:10.1016/j.scienta.2015.10.027
13. Huang W, Xian Z, Hu G, Li Z (2016) SlAGO4A, a core factor of RNA-directed DNA methylation (RdDM) pathway, plays an important role under salt and drought stress in tomato. Mol Breed 36(3):1–13. doi:10.1007/s11032-016-0439-1
14. Hendelman A, Kravchik M, Stav R, Zik M, Lugassi N, Arazi T (2013) The developmental outcomes of P0-mediated ARGONAUTE destabilization in tomato. Planta 237(1):363–377. doi:10.1007/s00425-012-1778-8
15. Meissner R, Jacobson Y, Melamed S, Levyatuv S, Shalev G, Ashri A, Elkind Y, Levy A (1997) A new model system for tomato genetics. Plant J 12(6):1465–1472. doi:10.1046/j.1365-313x.1997.12061465.x
16. Kato K, Matsumoto T, Koiwai S, Mizusaki S, Nishida K, Nogushi M, Tamaki E (1972) Liquid suspension culture of tobaco cells. In: Terui G (ed) Ferment technology today. Society of Fermentation Technology, Osaka, pp 689–695
17. Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol 42(6):819–832. doi:10.1023/A:1006496308160
18. Wang H, Jones B, Li Z, Frasse P, Delalande C, Regad F, Chaabouni S, Latche A, Pech JC, Bouzayen M (2005) The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. Plant Cell 17(10):2676–2692. doi:10.1105/tpc.105.033415

Chapter 14

Nonradioactive Detection of Small RNAs Using Digoxigenin-Labeled Probes

Ariel H. Tomassi, Delfina Gagliardi, Damian A. Cambiagno, and Pablo A. Manavella

Abstract

Small RNAs have been traditionally detected and quantified using small RNA blots, a modified Northern blot technique. The small RNAs are size-fractionated from the rest of the cellular RNA molecules by polyacrylamide gel electrophoresis and transferred by blotting onto a positively charged membrane. A radiolabeled probe was then traditionally used to detect a specific small RNA in the cellular pool. Small RNA blotting is a relatively simple, inexpensive approach to visualize small RNAs without artifacts. However, the radioactive labeling of the probe is sometimes an impediment, especially due to the requirement of specialized facilities. Here we describe a sensitive and simple method to detect and quantify small RNAs using digoxigenin-based nonradioactive RNA blots.

Key words Polyacrylamide electrophoresis, Northern blot, Small RNA, MicroRNA, Digoxigenin labeling

1 Introduction

Since the discovery of small RNAs over two decades ago, a modified version of the traditional Northern blot approach has been widely used to detect, quantify, and determine the size of small RNA species. Even when the detection principle is the same as in Northern blots, the approach used to detect small RNAs is commonly known as small RNA blots to highlight the differences between both methods. This approach involves the size separation of the small RNAs from the rest of the cellular RNA pool by polyacrylamide gel electrophoresis, the electrophoretic transfer of the RNA to a positively charged nylon membrane, and the subsequent detection using labeled DNA probes [1].

In the recent years, high-throughput sequencing of small RNA libraries has accelerated the discovery of novel small RNAs at an unprecedented rate. Although alternative approaches to quantify

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_14,

small RNAs were developed [2–7], the small RNA blot is still the reference technique to validate high-throughput data. In comparison with RT-qPCR or sequencing-based approaches, the small RNA blot analysis is less sensitive, has low throughput, and fails to differentiate between small RNAs with very similar sequences (such is the case of different members of a microRNA family or slightly misprocessed miRNA species). However, it is simple and relatively artifact-free and can be used to determine the small RNA size. Additionally, it allows simultaneous detection of the small RNAs and their precursors, such the case of pri- and pre-microRNAs.

Generally the first two steps of the method, the gel electrophoresis and blotting, present little differences among alternative protocols with just a few exceptions [8]. However, several approaches have been described for the probe labeling and small RNA detection. Among them, the labeling of DNA probes by the incorporation of the radioactive isotope ^{32}P is by far the most popular probe-labeling protocol. Due to the half-life of the radioisotope, it is possible to expose the X-ray film to the membrane for several weeks if necessary, allowing a high sensibility and the detection of very low-abundant small RNAs. However, isotope labeling is troublesome, hazardous, and commonly restricted by institutions. Alternative, non-radiolabeled approaches have been developed to detect small RNAs after gel transfer [1, 9–11]. We have extensively used digoxigenin (DIG)-labeled probes to detect a wide range of small RNA types, such as microRNAs (miRNAs), small interfering RNAs (siRNAs), and *trans*-acting siRNAs (tasiRNAs) [12–16]. This approach conserves the sensitivity of isotope labeling-based methods; it is safer and presents the great advantage that the whole protocol can be done in the workbench without the need of dedicated facilities.

2 Materials

Prepare all solutions using ultrapure water (18 MΩ cm at 25 °C) and analytical grade reagents. Autoclave and store all solutions at room temperature unless indicated. Use DEPC-treated water to resuspend and dilute RNA samples.

2.1 Polyacrylamide (PAA) Gel

1. 40% (w/v) acrylamide/bis-acrylamide (37.5:1) solution (PAA). The polyacrylamide solution can be purchase ready to use, or prepare it by yourself weighting 38.96 g of acrylamide and 1.04 g bis-acrylamide, and dissolve the powders in 100 mL of water (*see* **Note 1**).
2. Urea, electrophoresis grade.
3. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED).

4. 25% ammonium persulfate (APS) solution. Prepare small, 1 mL, aliquots of solubilized APS and store at −20 °C. Prepare fresh APS solution monthly.
5. 5× Tris/borate/ethylenediaminetetraacetic acid (EDTA) buffer (TBE). Dissolve 54 g Tris, 27.5 g boric acid, and 20 mL 0.5 M EDTA pH 8.0, in 750 mL of distilled deionized water. Adjust the volume to 1 L.
6. 2× RNA loading buffer: 95% formamide, 18 mM EDTA pH 8.0, 0.025% sodium dodecyl sulfate (SDS), 0.01% bromophenol blue, 0.01% xylene cyanol.
7. Mini-PROTEAN Tetra Cell (Bio-Rad) or similar.
8. Gel staining solution: prepare 30 mL of 0.5 μg/mL ethidium bromide in 1× TBE buffer per gel to be stained. Alternatively, 1× SYBR Gold/Safe Nucleic Acid Gel Stain (Thermo Fisher Scientific) can be used to stain the gel.

2.2 Small RNA Blotting

1. Positively charged nylon membrane. We use Amersham Hybond-N+ for the best results but different brands can potentially be used.
2. Extra thick filter paper (1.5–2 mm thick).
3. Trans-Blot SD semidry transfer cell (Bio-Rad) or similar.
4. UV Crosslinker.

2.3 Probe Labeling and Purification

1. DIG Oligonucleotide 3′-End Labeling Kit, 2nd generation (Sigma-Aldrich).
2. 0.2 M EDTA pH 8.0.
3. 10× Tris-buffered saline (TBS). Dissolve 60.5 g Tris and 87.6 g sodium chloride (NaCl) in 800 mL of H_2O. Adjust pH to 7.5 with 1 M hydrochloric acid (HCl), and make volume up to 1 L with H_2O.
4. TBS-milk. Dissolve 5 g of skim milk in 30 mL of H_2O, add 5 mL of 10× TBS, and adjust the volume to 50 mL.
5. Anti-DIG antibody: anti-digoxigenin-AP, Fab fragments (Sigma-Aldrich).
6. TNM-50 Buffer: 100 mM Tris pH 9.5, 100 mM NaCl, 50 mM magnesium chloride ($MgCl_2$).
7. NBT-BCIP stock solution (Sigma-Aldrich).
8. 1× Tris EDTA (TE): 10 mM Tris pH 8.0, 1 mM EDTA.
9. Bio-Gel P-6 Micro Bio-Spin chromatography columns (Bio-Rad).

2.4 Hybridization, Detection, and Stripping

1. Blocking solution: 1% (w/v) blocking reagent (Sigma-Aldrich) in maleic acid buffer. Do not autoclave blocking solutions. Prepare always fresh before use.

2. Maleic acid buffer: 0.1 M maleic acid, 0.15 M NaCl. Adjust the pH to 7.5 using sodium hydroxide (NaOH) pellets.
3. Antibody solution: 1:20,000 anti-digoxigenin-AP, Fab fragments (Sigma-Aldrich) in blocking solution.
4. Washing buffer: 0.3% (v/v) Tween 20 in maleic acid buffer.
5. Detection buffer: 0.1 M Tris, 0.1 M NaCl, adjust pH to 9.5 with HCl.
6. Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2′-(5′-chloro) tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate (CSPD), ready to use (Sigma-Aldrich).
7. Hypersensitive X-ray film (we recommend Amersham Hyperfilm ECL or MP).
8. PerfectHyb Plus Hybridization Buffer (Sigma-Aldrich).
9. Hybridization bottles and oven.
10. Stripping solution: 50% formamide; 5% SDS; 50 mM Tris–HCl pH 7.4.
11. Church buffer: 500 mL of 0.5 M sodium phosphate buffer pH 7.2, 2 mL 0.5 M EDTA pH 8, 10 g bovine serum albumin (BSA) fraction V, 70 g SDS (dissolve the SDS separately in 400 mL of H_2O). Take to 1 L with H_2O. Filtrate, fractionate, and freeze at −20 °C.
12. Sodium citrate (SSC) 20×. Dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 mL of H_2O. Adjust pH to 7.0 by adding concentrated HCl and adjust the volume to 1 L.

3 Methods

Carry out all procedures at room temperature unless otherwise specified. This protocol requires RNase-free handling. Gloves must be worn during all the steps described.

3.1 17% Denaturing Polyacrylamide (PAA) Gel and Blotting

1. Start by assembling the Mini-PROTEAN system and pouring sterile water between the glasses to check leaking and hydrate the gaskets. Leave the system filled with water until you are ready to pour the PAA mix.
2. For two 1 mm mini gels, pour into a sterile 50 mL falcon tube: 6.375 mL of 40% acrylamide/bis-acrylamide (37.5:1), 6.3 g urea, 3 mL 5× TBE, and add water to a volume of 15 mL. Heat the mixture to 40–50 °C until the urea gets dissolved. Let the solution to cool down at room temperature. Remove the water from the gel system, and leave it upside down to remove all liquid. Just before casting the gels, use a paper towel to remove any remaining water droplets from the casting system. Add 7.5 μL of TEMED and 20 μL of 25% APS to the PAA mixture.

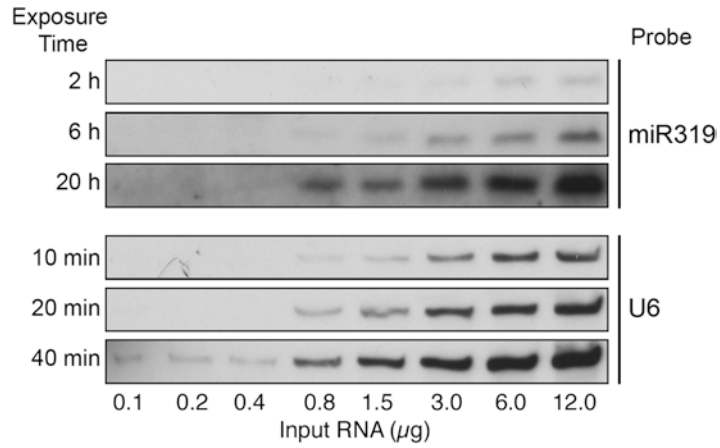


Fig. 1 Analysis of DIG nonradioactive small RNA blot sensitivity. Decreasing amount of total RNA (0.1, 0.2, 0.4, 0.8, 1.5, 3.0, 6.0, and 12.0 μg), extracted from *Arabidopsis thaliana* Col-0 wild-type 15-day-old seedlings, was loaded into a 17% denaturing polyacrylamide gel, size fractioned, blotted, and hybridized as specified in this protocol. Post-incubation with the anti-DIG antibody and CSPD solution membranes were exposed to hypersensitive X-ray film for different periods of times (2, 6, and 20 h for miR319 detection; 10, 20, and 40 min for U6 small nuclear RNA detection, normally used as loading control). Clear detection required ~80 pg of total RNA for miR319, while as little as 10 pg is enough to detected U6

Immediately pour the PAA solution to the preassembled glass-plate sandwich, and place 10–15 well combs. Allow polymerizing for 60 min (*see* **Note 2**).

3. Remove the gel comb and mount the gels in the electrophoresis unit. Fill the tank and space between gels with 1× TBE buffer. Carefully clean the wells using a 10 mL syringe with a hypodermic needle to remove any unpolymerized PAA and dissolved urea. Pre-run the gel for 60 min at 180 V.
4. Half an hour before the pre-run is over, start preparing the RNA samples. Mix 10 μL of 2× RNA loading buffer with 2 μg of total RNA (*see* technique sensibility in Fig. 1). Adjust the volume to 20 μL with DEPC-treated water. Heat the samples for 5 min at 95 °C and cool down on ice immediately. Briefly vortex the samples and spin to collect (*see* **Note 3**).
5. Repeat the wells cleaning using a 10 mL syringe with a hypodermic needle, and load the samples. Load the empty wells with 1× RNA loading buffer.
6. Run the gel at 180 V until the bromophenol blue runs out of the gel (~90 min). Disassemble the gel cassette, and cut off the wells and the lower-right corner of the gel to mark the order of the samples (*see* **Note 4**).

7. Transfer the gels to individual plastic boxes containing gel staining solution for 5 min. Visualize the stained gels using an UV transilluminator avoiding long exposures. Take a picture, and mark the ladder by cutting small triangles on the side of the gel. Ribosomal RNA will be visible in the gels and can be used to check the RNA integrity and sample loading evenness. If the samples are unevenly loaded or degraded, start all over again remeasuring RNA concentration and quality.
8. Cut out a piece of positively charged nylon membrane and two pieces of extra thick filter paper to match the gel size. Cut the lower-right corner of the nylon membrane to mark the samples orders after the transfer. You can use a pencil to label the membranes, especially to identify the side where the RNA will be blotted.
9. Soak the membrane and the extra thick filter papers in 1× TBE. Place a presoaked paper onto the platinum anode of a semidry transfer cell. Place the blotting membrane on top of the filter paper, and carefully put the gel on top of the membrane matching the marked lower-right corners. Using a pencil, mark in the membrane the position of the xylene cyanol and the ladder. Cover the gel with a second sheet of filter paper avoiding trapping air bubbles. Four mini gels can be blotted simultaneously.
10. Transfer the RNA for 60 min at 10 V and 400 mA (*see* **Note 5**).
11. Disassemble the sandwich, and allow the excess of buffer to drain by laying the membrane on top of a tissue paper (*see* **Notes 6** and **7**).
12. Fix the RNA covalently to the membrane in an UV Crosslinker at 120,000 μJ/cm^2 (*see* **Note 8**).

3.2 Probe Labeling

1. Dilute a LNA oligonucleotide to 10 pmol in 10 μL of DEPC-treated water (*see* **Note 9**).
2. Use the DIG Oligonucleotide 3′-End Labeling Kit, 2nd generation, to label the oligonucleotide as follows: add 4 μL of 5× reaction buffer, 4 μL of the cobalt chloride ($CoCl_2$) solution, 1 μL of DIG-ddUTP, and 1 μL of terminal transferase to the diluted oligonucleotide. Keep all the reagents and reaction tubes on ice all the time. Mix and spin briefly. Incubate for 15 min at 37 °C and immediately cool down on ice (*see* **Note 10**).
3. Stop the labeling reaction by adding 2 μL of 0.2 M EDTA pH 8.0.
4. Purify the DIG-labeled probes using a Bio-Gel P-6 Micro Bio-Spin chromatography column following the manufacturer's instructions (*see* **Note 11**).

5. Dilute the labeled oligonucleotide to 2.5 pmol/μL by adding 18 μL of DEPC-treated water. The labeled probe will be enough for 20 hybridizations.
6. Spot 1 μL of serial dilutions (100, 30, 10, and 3 fmol/μL) of the labeled probe, and the positive control included in the kit, into a small piece of positively charged nylon membrane to test labeling efficiency.
7. Cross-link the membrane at 120,000 μJ/cm^2.
8. Block the membrane for half an hour in TBS-milk with gentle shaking.
9. Dilute the anti-DIG antibody 1:5000 in fresh TBS-milk, and incubate the membrane with the antibody solution for 60 min at room temperature with agitation.
10. Wash the membrane five times during 1 min with 1× TBS and two times with TNM-50.
11. Dilute 100 μL of NBT-BCIP solution in 5 mL of TNM-50. Transfer the membrane to a Petri dish containing the NBT solution, and incubate 10 min protecting it from the light.
12. Briefly wash the membrane with water and stop the reaction using 1× TE.
13. Visually compare the spot intensity between the positive control and the labeled probe. Normally, a well-labeled probe has to be visible up to the 10 fmol/μL dilution or equally to the positive control. If you want to keep the blot for future references, scan it immediately and conserve the digital file, due to the fact that, after the incubation with the NBT-BCIP solution, the membrane will turn dark during the following days.
14. Labeled probe will remain stable at −20 °C for years.

3.3 Small RNA Hybridization (Continue from Subheading 3.1)

1. Use the xylene cyanol mark to cut the membrane horizontally in halves. The upper half will be used to hybridize and detect the U6 small nuclear RNA as loading reference. The lower half contains the small RNAs.
2. Put the membranes into hybridization bottles, and add 10 mL of PerfectHyb Plus Hybridization Buffer (*see* **Note 12**). Prehybridize the membranes for at least 60 min at 38 °C in rotation.
3. Add 5 pmol of DIG-labeled probe to 100 μL of PerfectHyb buffer, and heat for 10 min at 65 °C. Cool down on ice and spin briefly.
4. Add the probe to the bottles containing the membranes and PerfectHyb solution, and hybridize overnight at 38 °C.

5. Remove the hybridization solution from the bottle and briefly rinse in 4× SSC. Wash the membrane two times for 20 min each at 38 °C using at least 25 mL of 4× SSC.

3.4 Nonradioactive Small RNA Detection

1. Transfer the membranes to a suitable small plastic box or sterile Petri dish containing 15 mL of blocking solution, and incubate for 30 min at room temperature with gentle agitation. Every 10–15 min turn the membrane upside down.
2. Replace the blocking solution with 15 mL of antibody solution, and incubate at room temperature for 60 min with gentle agitation. Every 15 min turn the membrane upside down.
3. Transfer the membranes to a clean plastic container, and wash them four times with 50 mL of washing buffer, for 15 min each, at room temperature with vigorous shaking.
4. Equilibrate the membrane in detection buffer for 5 min.
5. Place the membranes with RNA side facing up on a plastic foil, and apply dropwise CSPD ready-to-use solution. Immediately cover the membrane with a second sheet of plastic foil, and carefully spread the substrate evenly and without air bubbles over the membrane (*see* **Note 13**).
6. Incubate the membrane in darkness for 10 min at 37 °C to fully activate the CSPD solution.
7. Squeeze out the excess of CSPD solution and seal the plastic foils.
8. Place the membranes into X-ray film cassette, and expose to hypersensitive X-ray film at 15–25 °C (*see* **Note 14**). Split the membranes into different cassettes depending on the expected exposure times (*see* **Note 15**).
9. Develop the X-ray film using high-quality, freshly prepared, developing, and fixing solutions or a developer machine (*see* **Note 16**).

3.5 Membrane Stripping and Re-probing

1. Rinse the membranes twice with water to remove the CSPD solution (*see* **Note 13**).
2. Incubate the membranes with 25 mL of stripping solution prewarmed to 80 °C in a plastic box for 60 min with constant shaking. Repeat this step one more time (*see* **Note 17**).
3. Wash the membranes three times for 15 min each in 4× SSC (*see* **Note 18**).
4. Before re-probing, rinse the membrane with water, and then proceed from Subheading 3.3, **step 2** onward (*see* **Note 19**) (*see* expected results in Fig. 2).
5. Stripping and re-probing can be carried on several times using the same membrane (*see* **Note 20**).

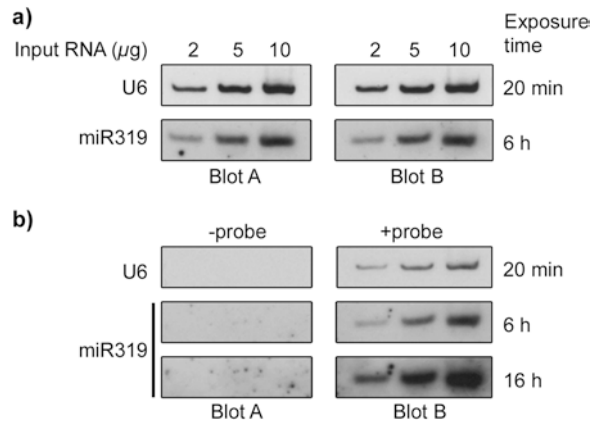


Fig. 2 Stripping and re-probing small RNA blots. (**a**) Increasing amounts of total RNA were size fractioned by duplicate and transferred to positively charged nylon membranes (blots A and B). Membranes were hybridized with U6 and miRNA319 DIG-labeled probes. Small RNA detection was done exposing the membranes for different periods of time to hypersensitive X-ray films. (**b**) The same membranes were subsequently stripped and re-hybridize with or without the same probes used in (**a**). Membranes were exposed to X-ray films to detect U6 and miR319. No signal was detected in the stripped membranes even after long exposure times when no fresh probe is applied. A minimal reduction in the detected RNA, particularly visible for the U6 when using 2 μg of total RNA, was observed after stripping

4 Notes

1. Wear a mask when weighing or manipulating acrylamide. Ideally purchase ready-to-use acrylamide/bis-acrylamide solution to avoid exposing coworkers to acrylamide. Cast the gels and work under a fume hood until the polyacrylamide gel is fully polymerized. Unpolymerized acrylamide is a strong neurotoxin, and care should be exercised to avoid skin contact. Any unpolymerized PAA should be disposed as hazardous waste. The acrylamide solution can be stored at 4 °C in dark. Do not autoclave this solution.
2. If a semidry system is used for the blotting, up to four gels can easily be casted and run simultaneously.
3. If necessary, the final volume can be scaled down or up depending of the size of the comb used to cast the gels. Total RNA concentration can also be increased until reaching half the maximum volume of the well (the other half is loading buffer) in those cases where low-abundant or rare small RNAs want to be detected (*see* Fig. 1). Column-based RNA purification kits do not retain small RNAs and therefore cannot be used to obtain the input RNA for small RNA blots; TRIzol-purified

total RNA is recommended instead. Alternatively, purified small RNAs (e.g., using PAA gel purification or small RNAs purification kits) can be loaded instead of total RNA to achieve higher sensitivity.

4. In case you need to determine small RNA sizes, you can load one well with 0.5 nmol of a RNA oligo of any specific size (a mix of 21 nt and 24 nt is recommended). Treat this RNAs in the same way as the RNA samples. In order to determine the small RNA, size is important to take a picture of the stained gel and mark the membrane.
5. Alternatively, a Wet/Tank Blotting Systems can be used to transfer the small RNAs for 90 min at 10 V and 400 mA. In this case only two mini gels can be blotted simultaneously per tank used.
6. Avoid excessive drying of the membrane. Normally less than a minute is enough to remove the excess of buffer.
7. Normally, the RNA remains stained after blotting and can be visualized using an UV transilluminator. Mark the position of the size markers in the membrane using a pencil.
8. At this point, the membrane can be stored at 4 °C by wrapping them in Saran Wrap to avoid excessive drying. Alternatively, the chemical cross-linking using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) was reported to enhance sensitivity on neutral charged nylon membrane [17].
9. Design each specific oligonucleotide complementary to the small RNA to be detected. Alternatively, regular oligonucleotides can be used only losing power to differentiate closely related small RNAs (e.g., different microRNA members of the same family).
10. Modified oligonucleotides containing a 3′ DIG-ddUTP can also be directly purchased from most commercial providers. Even when this option is more convenient and straightforward, we have experienced a lower signal from commercially labeled oligos. The DIG Oligonucleotide Tailing Kit, 2nd generation (Sigma-Aldrich) can be used if a stronger signal is necessary.
11. In case of using the DIG Oligonucleotide Tailing Kit, 2nd generation, you may need to use the Bio-Gel P-30 Micro Bio-Spin chromatography columns (Bio-Rad) for purification.
12. Church buffer can be used as an economic alternative losing a portion of the signal. Alternative hybridization buffers have been previously tested and compared [9].
13. Avoid the membranes to get dry at any point.
14. Luminescence continues for at least 48 h. The signal increases in the first few hours after initiation of the detection reaction

until it will reach a plateau where signal intensity remains almost constant during the next 24–48 h.

15. Exposure time recommendations (based on 2 μg of initial total RNA input): U6 for 15–20 min; highly expressed miRNAs for 3–4 h; low-expressed miRNAs, tasiRNA, and endogenous siRNA for 6–16 h. Keep in mind that increasing exposure times produces a raise in the background (*see* Fig. 1).
16. Poor cleaning of unincorporated nucleotides during probe labeling and non-dissolved particles in the hybridization or blocking buffer can cause the appearance of speckles on the membrane area of the developed X-ray films. Ensure that these reagents are in solution, and consider spinning in a centrifuge or filtering the solutions.
17. If a thermostatic shaker is not available, you can perform this step by putting the membranes in any incubator at 80 °C, and manually shake them every 10 min.
18. Membranes can be stored in 4× SCC for a couple of days before re-probing them.
19. Even when the stripping protocol is very efficient, it is possible that some probe remains bound to the membrane. This is generally not a problem since it will mean a substantial increase in the exposure time required to detect such un-stripped probes and therefore non-detectable at the exposure time required for the next hybridization round. However, to reduce the risk of cross detection, we recommend always starting by detecting the scarcer small RNAs first. Alternatively, complete stripping can be achieved by using probes labeled with chemically sensitive modified nucleotides, such as alkali labile DIG-ddUTP that can be degraded by specific treatments.
20. Successive membrane stripping may cause a partial loss of the membrane-bounded small RNAs losing sensitivity. We recommend performing the first hybridizations using the probes that recognize the scarcer small RNAs and progressively continue to detect the more abundant ones.

Acknowledgments

The laboratory of P.A.M. is funded by the Human Frontier Science Program, the Max Planck Society, the International Centre for Genetic Engineering and Biotechnology, and the Agencia Nacional de Promoción Científica y Tecnológica, Argentina. P.A.M. is a member of Consejo Nacional de Investigaciones Científicas y Técnicas. A.H.T. and D.G are fellows of the same institution.

References

1. Blevins T (2010) Northern blotting techniques for small RNAs. Methods Mol Biol 631:87–107. doi:10.1007/978-1-60761-646-7_9
2. Varkonyi-Gasic E, Hellens RP (2011) Quantitative stem-loop RT-PCR for detection of microRNAs. Methods Mol Biol 744:145–157. doi:10.1007/978-1-61779-123-9_10
3. Salone V, Rederstorff M (2015) Stem-loop RT-PCR based quantification of small non-coding RNAs. Methods Mol Biol 1296:103–108. doi:10.1007/978-1-4939-2547-6_10
4. Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods 3:12. doi:10.1186/1746-4811-3-12
5. Yang SW, Vosch T (2011) Rapid detection of microRNA by a silver nanocluster DNA probe. Anal Chem 83(18):6935–6939. doi:10.1021/ac201903n
6. Yang LH, Wang SL, Tang LL, Liu B, Ye WL, Wang LL, Wang ZY, Zhou MT, Chen BC (2014) Universal stem-loop primer method for screening and quantification of microRNA. PLoS One 9(12):e115293. doi:10.1371/journal.pone.0115293
7. 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
8. Harris CJ, Molnar A, Muller SY, Baulcombe DC (2015) FDF-PAGE: a powerful technique revealing previously undetected small RNAs sequestered by complementary transcripts. Nucleic Acids Res 43(15):7590–7599. doi:10.1093/nar/gkv604
9. Kim SW, Li Z, Moore PS, Monaghan AP, Chang Y, Nichols M, John B (2010) A sensitive non-radioactive northern blot method to detect small RNAs. Nucleic Acids Res 38(7):e98. doi:10.1093/nar/gkp1235
10. Huang Q, Mao Z, Li S, Hu J, Zhu Y (2014) A non-radioactive method for small RNA detection by northern blotting. Rice (N Y) 7(1):26. doi:10.1186/s12284-014-0026-1
11. Schwarzkopf M, Pierce NA (2016) Multiplexed miRNA northern blots via hybridization chain reaction. Nucleic Acids Res 44(15):e129. doi:10.1093/nar/gkw503
12. Francisco-Mangilet AG, Karlsson P, Kim MH, Eo HJ, SA O, Kim JH, Kulcheski FR, Park SK, Manavella PA (2015) THO2, a core member of the THO/TREX complex, is required for microRNA production in Arabidopsis. Plant J 82(6):1018–1029. doi:10.1111/tpj.12874
13. Karlsson P, Christie MD, Seymour DK, Wang H, Wang X, Hagmann J, Kulcheski F, Manavella PA (2015) KH domain protein RCF3 is a tissue-biased regulator of the plant miRNA biogenesis cofactor HYL1. Proc Natl Acad Sci U S A 112(45):14096–14101. doi:10.1073/pnas.1512865112
14. Manavella PA, Hagmann J, Ott F, Laubinger S, Franz M, Macek B, Weigel D (2012) Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell 151(4):859–870. doi:10.1016/j.cell.2012.09.039
15. Manavella PA, Koenig D, Rubio-Somoza I, Burbano HA, Becker C, Weigel D (2013) Tissue-specific silencing of Arabidopsis SU(VAR)3-9 HOMOLOG8 by miR171a. Plant Physiol 161(2):805–812. doi:10.1104/pp.112.207068
16. Manavella PA, Koenig D, Weigel D (2012) Plant secondary siRNA production determined by microRNA-duplex structure. Proc Natl Acad Sci U S A 109(7):2461–2466. doi:10.1073/pnas.1200169109
17. Pall GS, Hamilton AJ (2008) Improved northern blot method for enhanced detection of small RNA. Nat Protoc 3(6):1077–1084. doi:10.1038/nprot.2008.67

Chapter 15

Nonradioactive Plant Small RNA Detection Using Biotin-Labeled Probes

Jun Hu and Yingguo Zhu

Abstract

Small noncoding RNAs are essential for gene expression at transcriptional and posttranscriptional levels. Northern blot is the most used method for small RNA detection in tissues. Here we present an improved protocol for the Northern blot-based small RNA detection from plant tissues by using biotin-labeled probes. MicroRNAs and small interfering RNAs derived from *Arabidopsis* and *Oryza sativa*, respectively, have been detected with this methodology. Results suggest that this method is sensitive and efficient enough to detect small RNAs from plant tissues by using as low as 5 μg of total RNA. Furthermore, biotin-labeled probes are safer and easier to store for long term than radiolabeled probes.

Key words Small RNA, Biotin-labeled probe, Nonradioactive, Northern blot, Epigenetic

1 Introduction

In eukaryotic cells, small noncoding RNAs have been shown to play fundamental roles in gene expression modification during development [1]. Small noncoding RNAs are 20–30 nt long and function as sequence-specific negative regulators of gene expression at the transcriptional and/or posttranscriptional levels [2–4]. Recently, high-throughput sequencing technology has facilitated the exploration of small noncoding RNAs [5, 6]. To date, thousands of small noncoding RNAs and their target mRNAs or genes have been computationally predicted. To gain further insight, several methods have been developed to investigate the expression of small noncoding RNAs such as Northern blotting, quantitative reverse-transcription PCR (RT-PCR), and in situ hybridization.

The low abundance of certain small RNAs in tissues can be problematic for their detection via Northern blotting. Therefore, researchers have developed new methods to improve Northern blot-based small RNA detection. These include treating samples with LiCl to enrich for small RNAs [7], using non-isotopic-labeling

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_15,

methods using digoxigenin (DIG) which represents a safe alternative for small RNA detection [8, 9] or modifying probes with locked nucleic acid (LNA) to improve sensitivity [10–12].

Although isotope labeling is often inconvenient, hazardous, and restricted by many institutions, this classic method is still the most popular method for investigating the expression of small RNAs. Here, we present a protocol for the generation of biotin-labeled oligonucleotide probes to investigate the expression of small RNAs by Northern blotting [13]. Results suggest that our protocol is sensitive and efficient, as it is capable of detecting small RNAs from as little as 1–5 μg of total RNA from *Arabidopsis thaliana* and *Oryza sativa*. Importantly, the probes are safe, and storage of biotin-labeled probes is stable and convenient.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water, to attain a sensitivity of 18 MΩ cm at 25 °C) and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Tips and tubes should be DNase and RNase free.

2.1 Probe Preparation

The probes modified with biotin on 5′ terminus and/or 3′ terminus were synthesized and purified via HPLC by GenScript (*see* **Note 1**).

2.2 RNA Extraction

1. Plants were grown on plates or soil under proper management.
2. TRIzol reagent (Thermo Fisher Scientific) and RNAiso (TaKaRa) for total RNA or small RNA purification.
3. Chloroform.
4. Isopropanol.
5. Ethanol.
6. NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific).
7. Diethypyrocarbonate (DEPC)-treated H_2O (*see* **Note 2**).

2.3 7 M Urea 15% Denaturing Polyacrylamide Gel

1. 10× Tris borate ethylenediaminetetraacetic acid (EDTA) (TBE): 108 g Tris base, 54 g boric acid, 40 mL 0.5 mM EDTA, and H_2O up to 1 L, pH 8.0 (*see* **Note 3**).
2. 15% acrylamide/bis-acrylamide containing 7 M urea solution: weigh 14.25 g of acrylamide monomer and 0.75 g bis-acrylamide, and transfer to a 100 mL graduated cylinder containing about 30 mL of deionized water and mix for about 30 min. Add 42 g of urea and 10 mL of 10× TBE. Warm the

solution at 60 °C until urea completely dissolves. Make up to 100 mL with water and filter through Whatman filter paper (*see* **Note 4**). Store at 4 °C in a bottle wrapped with aluminum foil (*see* **Note 5**).

3. Ammonium persulfate (APS): 10% solution in water (*see* **Note 6**).
4. *N*,*N*,*N*′,*N*′-tetramethyl-ethylenediamine (TEMED): store at 4 °C (*see* **Note 7**).
5. 2× loading buffer: 5 mM EDTA, 0.1% bromophenol blue, 0.1% Xylene cyanol, and 95% deionized formamide.

2.4 RNA Transfer and UV Cross-Linking

1. Transfer buffer: 5.8 g Tris base, 12.1 g glycine, 0.37 g sodium dodecyl sulfate (SDS), 200 mL methanol, and H_2O up to 1 L.
2. Whatman filter paper.
3. Hybond-N+ nylon membranes positively charged (GE healthcare).
4. UV cross-linking oven.

2.5 Hybridization

1. Pre-hybridization buffer: 7% SDS, 200 mM sodium phosphate dibasic (Na_2HPO_4) pH 7.0, 5 μg/mL salmon sperm DNA (SSDNA).
2. Hybridization buffer: pre-hybridization buffer containing 50 pmol/mL labeled probes (*see* **Note 8**).
3. Washing buffer: 1× SSC, 0.1% SDS.

2.6 ECL Lighting

1. Chemiluminescent Nucleic Acid Detection Module kit (Thermo Fisher Scientific).
2. X-film.

3 Methods

3.1 RNA Extraction

1. Grind approximately 50 mg of leaf tissue to fine powder with liquid nitrogen, and add 1 mL of TRIzol or RNAiso.
2. Shake the tube vigorously for 30 s, and incubate at room temperature for 5 min.
3. Add 300 μL of chloroform and shake vigorously for 2 min, then incubate at room temperature for 5 min.
4. Centrifuge at 12,000 × *g* for 15 min at 4 °C.
5. Carefully transfer the aqueous (upper) phase of the sample to a new RNase-free tube (s*ee* **Note 9**).
6. Add 400 μL chloroform, vortex for 2 min, and centrifuge the mixture at 12,000 × *g* for 5 min at 4 °C (s*ee* **Note 10**).
7. Precipitate total RNA or small RNA by adding equal volume of isopropanol and incubating at −20 °C for 1 h.

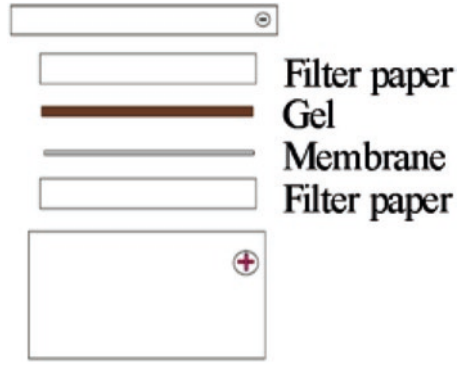


Fig. 1 Gel-membrane assembly for the transfer of PAGE gel to membrane for Northern blot analysis. PAGE gel is sandwiched between membrane and filter papers in transfer apparatus in the indicated order

8. Centrifuge at 12,000 × *g* for 15 min at 4 °C.
9. Wash RNA pellets with 200 μL 75% ethanol, and resuspend them in DEPC-treated H_2O.
10. Determine the concentration of total RNA or small RNA with a NanoDrop 2000c spectrophotometer.

3.2 7 M Urea 15% Denaturing Polyacrylamide Gel

1. Wash glass plates thoroughly with warm water and liquid detergent. Rinse them thoroughly with deionized water to remove detergent residues and wipe with tissue paper soaked in 70% alcohol. Dry the glass plates by laying them on a Whatman filter paper (*see* **Note 11**).
2. Add 75 μL APS and 7.5 μL of TEMED to 15 mL 15% PAGE containing 7 M urea solution, and cast gel into the gel cassette.
3. Insert the comb in gel to the edge of the plate. Clamp with clips and keep in appropriate position until the gel has polymerized (*see* **Note 12**).
4. Pre-run the gel with 1× TBE at 40 mA (600 V) for 30 min.
5. Wash the wells by pipetting for several times (*see* **Note 13**).
6. Mix RNAs with equal volume of 2× loading buffer. RNAs samples are denatured at 70 °C for 5 min (*see* **Note 14**).
7. Run the PAGE gel at 40 mA (600 V) for about 2 h until bromophenol blue reaches about 1 cm above the bottom of the gel.

3.3 RNA Transfer and UV Cross-Linking

1. Place the gel following the order shown in Fig. 1.
2. Transfer the RNAs onto nylon membrane at 200 mA (9–10 V) for 2–3 h. The membrane is cross-linked at 1200 μJ for 20 min and dried at 50 °C for 30 min to improve sensitivity (*see* **Note 15**).
3. Membranes can be stored at 4 °C for several months (*see* **Note 16**).

4. Pre-hybridize for at least 30 min at 40 °C in pre-hybridization buffer (*see* **Note 17**).
5. Remove the pre-hybridization buffer and add hybridization buffer (*see* **Note 18**).

3.4 Hybridization

1. Hybridize the membrane for 12–16 h at 40 °C.
2. Rinse the membrane with washing buffer for 3 times about 15 min each time at room temperature (*see* **Note 19**).

3.5 ECL Lighting

Follow the manufacturer's instructions. Next are some specifications. Solutions, buffers, and enzymes are included in the ECL lightning kit.

1. Block membranes again with blocking buffer for 15 min with gentle shaking at room temperature.
2. Hybridize the membrane with hybridization buffer containing stabilized streptavidin-horseradish peroxidase conjugate for 15 min.
3. After washing for 3 times, equilibrate the membranes in substrate equilibration buffer for 5 min.
4. Completely cover the membrane with working solution, and incubate for 5 min in the dark (*see* **Note 20**).
5. Place the membrane in a cassette and expose to X-ray films (*see* **Note 21**).

4 Notes

1. The oligonucleotide can be modified on its 5′ or 3′ end, or at both ends. We recommend to biotinylate the oligonucleotide at both ends.
2. Drop 1 mL of DEPC to 1 L of deionized H_2O, and shake with magnetic stir bar overnight, followed autoclave at 121 °C for 20 min. Since RNA is unstable, please use DEPC-treated H_2O as much as possible.
3. 10× TBE precipitates easily. Thus, we find it is better to use it up within 3 months.
4. Acrylamide is a biohazardous chemical material; to avoid exposure of acrylamide to coworkers, please wear a mask and gloves and weigh it carefully in the hood. Filtering the prepared acrylamide is necessary before refrigerator storage. The used mixed resin should be disposed as biohazardous waste.
5. Store the acrylamide solution at 4 °C, and mark the date on the bottle, we recommend preparing the acrylamide solution every month.

6. Store the APS solution at 4 °C and mark the date on the tube. We recommend to prepare the APS solution with DEPC-treated H_2O.
7. To avoid the pungent smell, it is better to operate in the hood.
8. Store the probes at −20 °C. To make the hybridization buffer, denature the probe at 70 °C for 5 min before adding into pre-hybridization buffer.
9. Transfer the aqueous (upper) phase of the sample to a new RNase-free tube carefully by angling the tube at 45 °C. Be sure to not transfer any contaminant such as proteins or phenol.
10. The chloroform extraction helps removing the phenol and other contaminants.
11. Detergent and dust left on the glass plate may result in a high background.
12. Keep the gel at room temperature for about 1 h until it has polymerized. Once polymerized, the gel can be sealed with plastic wrap and stored at 4 °C for a week.
13. After pre-running the gel, we wash the wells by pipetting to remove the urea.
14. RNA are denatured at 70 °C and transferred to ice immediately. Another option is to denature the RNA samples at 65 °C for 15 min.
15. Some researchers use chemical cross-linking such as 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) to improve Northern blotting sensitivity.
16. Regarding the membrane for Northern blot, we tested 0.45 μM nitrocellulose and nylon membranes positively charged. Results suggest that nylon membranes are better than nitrocellulose. We can store the membrane sealed with plastic wrap in the refrigerator for at least 1 month.
17. 30 min to 3 h is recommended for decreasing the background noise. Increasing the pre-hybridization time is not helpful for increasing the intensity of signals. SSDNA is important for blocking the membrane to decrease the background noise.
18. Hybridization buffer should be pre-warmed at 40 °C.
19. To decrease the background noise, membranes should be rinsed for at least 3 times completely.
20. Place the membranes in a clean container, do not incubate the membranes for too long. Use facial tissues to absorb the solutions carefully and do not wipe.
21. The exposure time is dependent on the intensity of signals, since the amount of miRNA or siRNA is variable.

Acknowledgements

This work was supported by funds from National Natural Science Foundation of China (31371698 and 31670310).

References

1. Zamore PD, Haley B (2005) Ribo-gnome: the big world of small RNAs. Science 309(5740): 1519–1524. doi:10.1126/science.1111444
2. Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4):669–687. doi:10.1016/j.cell.2009.01.046
3. He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5(7):522–531. doi:10.1038/nrg1379
4. Chen X (2012) Small RNAs in development - insights from plants. Curr Opin Genet Dev 22(4):361–367. doi:10.1016/j.gde.2012.04.004
5. Studholme DJ (2012) Deep sequencing of small RNAs in plants: applied bioinformatics. Brief Funct Genomics 11(1):71–85. doi:10.1093/bfgp/elr039
6. Creighton CJ, Reid JG, Gunaratne PH (2009) Expression profiling of microRNAs by deep sequencing. Brief Bioinform 10(5):490–497. doi:10.1093/bib/bbp019
7. Song X, Wang D, Ma L, Chen Z, Li P, Cui X, Liu C, Cao S, Chu C, Tao Y, Cao X (2012) Rice RNA-dependent RNA polymerase 6 acts in small RNA biogenesis and spikelet development. Plant J 71(3):378–389. doi:10.1111/j.1365-313X.2012.05001.x
8. Ramkissoon SH, Mainwaring LA, Sloand EM, Young NS, Kajigaya S (2006) Nonisotopic detection of microRNA using digoxigenin labeled RNA probes. Mol Cell Probes 20(1):1–4. doi:10.1016/j.mcp.2005.07.004
9. Kim SW, Li Z, Moore PS, Monaghan AP, Chang Y, Nichols M, John B (2010) A sensitive non-radioactive northern blot method to detect small RNAs. Nucleic Acids Res 38(7):e98. doi:10.1093/nar/gkp1235
10. Gao Z, Peng Y (2011) A highly sensitive and specific biosensor for ligation- and PCR-free detection of microRNAs. Biosens Bioelectron 26(9):3768–3773. doi:10.1016/j.bios.2011.02.029
11. Lopez-Gomollon S (2011) Detecting sRNAs by northern blotting. Methods Mol Biol 732: 25–38. doi:10.1007/978-1-61779-083-6_3
12. Valoczi A, Hornyik C, Varga N, Burgyan J, Kauppinen S, Havelda Z (2004) Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. Nucleic Acids Res 32(22):e175. doi:10.1093/nar/gnh171
13. Huang Q, Mao Z, Li S, Hu J, Zhu Y (2014) A non-radioactive method for small RNA detection by northern blotting. Rice (N Y) 7(1):26. doi:10.1186/s12284-014-0026-1

Chapter 16

Computational Analysis of Genome-Wide ARGONAUTE-Dependent DNA Methylation in Plants

Kai Tang, Cheng-Guo Duan, Huiming Zhang, and Jian-Kang Zhu

Abstract

Whole-genome bisulfite sequencing (WGBS) has become a powerful tool to dissect genome-wide methylation profiles at single-base resolution. In this chapter we describe in detail the bioinformatics pipeline used for the analysis of ARGONAUTE-dependent DNA methylation in *Arabidopsis thaliana*. We provide tools and command lines used for mapping bisulfite sequencing reads, for estimating methylation levels at individual cytosine sites, for identifying differentially methylated regions (DMRs), and for calculating methylation levels of DMRs.

Key words ARGONAUTE, Methylation, Whole-genome bisulfite sequencing

1 Introduction

In 2008, whole-genome bisulfite sequencing (WGBS) was used by two independent groups to describe the DNA methylome of *Arabidopsis thaliana* (*Arabidopsis*) at a single-base resolution [1, 2]. Since then, the door of methylome sequencing has been opened. Multiple studies using WGBS were published for different species, including human [3, 4], mouse [5, 6], chimpanzee [7, 8], zebrafish [9], rice [10], tomato [11], maize [12], soybean [13], and several more.

ARGONAUTE (AGO) proteins are essential components of small RNA-induced silencing pathways. In *Arabidopsis*, AGO6 mutation partially suppresses transcriptional gene silencing in the *ros1* mutant and can reduce CHG and CHH methylation levels at several endogenous loci [14]. It was generally assumed that AGO6 was redundant with AGO4 in mediating RNA-dependent DNA methylation activities. By using WGBS and the bioinformatics method described herein, we found that the redundancy between

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_16,

AGO6 and AGO4 is unexpectedly negligible on a genome-wide scale [15]. In this chapter we will describe in detail the bioinformatics pipeline we used for analyzing WGBS data in the AGO study [15]. In particular, we provide tools and command lines used for mapping bisulfite sequencing reads, for estimating methylation levels at individual cytosine sites, for identifying differentially methylated regions (DMRs), and for calculating methylation levels of DMRs.

2 Materials

2.1 Hardware

Computer with Unix-based operating system (e.g., Linux, Mac OS X).

2.2 Publicly Available Software

1. NCBI SRA Toolkit version 2.7.0 (https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software).
2. BRAT-nova (http://compbio.cs.ucr.edu/brat/) [16].

2.3 In-House Tools

A set of utility Perl scripts (Table 1) were written to perform various tasks associated with data processing. These scripts are open source and freely available upon request.

2.4 Files

1. The FASTA-formatted files with complete sequences of the *Arabidopsis* chromosomes are TAIR10_chr1.fas, TAIR10_chr2.fas, TAIR10_chr3.fas, TAIR10_chr4.fas, TAIR10_chr5.fas, TAIR10_ChrC.fas, and TAIR10_ChrM.fas. These files can be downloaded at ftp://ftp.arabidopsis.org/home/tair/Sequences/whole_chromosomes/.
2. WGBS data set [15] can be downloaded from NCBI GEO with the accession number GSE56388. The accession number for wild-type (WT) Col-0 is GSM1080803.

Table 1
List of utility Perl scripts

Perl script	Utility
step1_retain.pl	Retaining cytosines with depth ≥4 and transforming the BRAT–nova output format to input format of step2_DMR.pl
step2_DMR.pl	Identifying DMRs
step3_cal_meth.pl	Calculating methylation levels for DMRs in different genotypes

3 Methods

3.1 Preparation

1. Download BRAT-nova by typing the following commands (*see* **Note 1**):

```
$ wget http://compbio.cs.ucr.edu/brat/downloads/
brat_nova.tar.gz
$ tar -xvzf brat_nova.tar.gz
$ cd brat_nova
$ make
```

   Five executable files are generated: acgt-count, brat_bw, build_bw, remove-dupl, and trim.

2. Download the 7 FASTA-formatted chromosome files online, and put the names (with full path) of FASTA files in a file named "TAIR10_ref.txt."

3. Build the index with the following command:

```
$ brat_nova/build_bw -P tair10_index -r TAIR10_ref.txt
```

   Thirteen files are generated in the directory "tair10_index":

```
CT.bwt, CT.bwt_marked, CT.chr, CT.no, CT.nosqr,
CT.ns, CT.pos_ind, CT.pos_num_occ, CT.pos_num_sqr,
CT.seeds, cg.cg, pos_strand.txt, and ta.ta.
```

4. Download fastq data from NCBI (e.g.,taking GSM1360162 double_MethylC-Seq).

5. Download the sra files from NCBI with commands (*see* **Note 2**):

```
$ wget ftp://ftp-trace.ncbi.nlm.nih.gov/sra/sra-
instant/reads/ByExp/sra/SRX/SRX504/SRX504864/
SRR1210378/SRR1210378.sra
$ wget ftp://ftp-trace.ncbi.nlm.nih.gov/sra/sra-
instant/reads/ByExp/sra/SRX/SRX504/SRX504864/
SRR1210379/SRR1210379.sra
```

6. Convert sra file to fastq files with commands:

```
$ sratoolkit.2.7.0-mac64/bin/fastq-dump.2.7.0
--split-3 -Q 33 --defline-seq "@\$sn" --defline-qual
"+" -E SRR1210378.sra
$ sratoolkit.2.7.0-mac64/bin/fastq-dump.2.7.0
--split-3 -Q 33 --defline-seq "@\$sn" --defline-qual
"+" -E SRR1210379.sra
```

   Four .fastq files are generated in the current directory:

```
"SRR1210378_1.fastq," "SRR1210378_2.fastq," "SRR12
10379_1.fastq," and "SRR1210379_2.fastq."
```

3.2 BRAT-Nova Analysis

1. To trim low-quality bases type the following commands:

```
$ brat_nova/trim -1 SRR1210378_1.fastq -2 SRR1210378_2.
fastq -P SRR1210378_double -q 20 -L 33 -m 0
$ brat_nova/trim -1 SRR1210379_1.fastq -2 SRR1210379_2.
fastq -P SRR1210379_double -q 20 -L 33 -m 0
```

2. Run brat_bw by typing the following commands:

```
$ brat_nova/brat_bw -P tair10_index -1 SRR1210378_
double_pair1.fastq -2 SRR1210378_double_pair2.fastq
-pe -o SRR1210378_double.sam -i 0 -a 1000 -m 2
$ brat_nova/brat_bw -P tair10_index -1 SRR1210379_
double_pair1.fastq -2 SRR1210379_double_pair2.fastq
-pe -o SRR1210379_double.sam -i 0 -a 1000 -m 2
```

3. Before running remove-dupl, we need a file containing the names of the output sam files for the double mutant. We can create a text file named "double_mapped_reads.txt" and put "SRR1210378_double.sam" and "SRR1210379_double.sam" in it. Each name should be in its own line (*see* **Note 3**).
4. Then run the following command:

```
$ brat_nova/remove-dupl -r TAIR10_ref.txt -s double_
mapped_reads.txt
```

   This command will generate two outputs: "SRR1210378_double.sam.nodupl" and "SRR1210379_double.sam.nodupl."
5. Before running acgt-count, we need a file containing the names of the output from step 3. We can create a text file named "double_mapped_reads_nodupl.txt" and put "SRR1210378_double.sam.nodupl" and "SRR1210379_double.sam.nodupl" in it. Each name should be in its own line.
6. Then run the following command:

```
$ brat_nova/acgt-count -r TAIR10_ref.txt -P double_
methylome.txt -s double_mapped_reads_nodupl.txt
```

   Then we will get the output "double_methylome.txt."
7. Repeat steps 1–6 for WT, *ago4*, and *ago6* data:

   Three files are generated: "WT_methylome.txt," "ago4_methylome.txt," and "ago6_methylome.txt."

3.3 Retaining the Cytosines that Have Depth ≥4 in All Four Libraries

1. Make an output directory named "out_dep4" by typing the following command:

```
$ mkdir out_dep4
```

2. Run the following command:

```
$ step1_retain.pl out_dep4 4 4 WT_methylome.txt WT
ago4_methylome.txt ago4 ago6_methylome.txt ago6
double_methylome.txt double
```

   This will generate four files in the out_dep4 directory: "WT_dep4_Meth.txt," "ago4_dep4_Meth.txt," "ago6_dep4_Meth.txt," and "double_dep4_Meth.txt."

3.4 Identifying DMRs for the ago4ago6 Double Mutant

Type the following commands:

```
$ mkdir DMR_output
$ step2_DMR.pl out_dep4/WT_dep4_Meth.txt  out_dep4/
double_dep4_Meth.txt DMR_output double_vs_WT
```

Two files corresponding to the hyper- and hypo-DMR lists are generated in the "DMR_output" directory: "double_vs_WT_hyper_list.txt" and "double_vs_WT_hypo_list.txt."

3.5 Calculation of the Methylation Level for Hypo-DMR

Calculate the methylation level for hypo-DMR in all the four genotypes using the step3_cal_meth_level.pl script (*see* **Note 4**). Type the following commands:

```
$ mkdir out_meth
$ step3_cal_meth.pl out_dep4 DMR_output/double_vs_
WT_hypo_list.txt out_meth double_vs_WT_hypo_meth_
level.txt
```

The file "double_vs_WT_hypo_meth_level.txt" is generated in the "out_meth" directory.

3.6 Categorizing Hypo-DMRs in the ago4ago6 Double Mutant

Categorize hypo-DMRs into four groups based on the DNA methylation levels (mC) in the different genotypes. Use the artificial cutoff "25%." If a DMR shows methylation reduction <25% in both single mutants, classify it as a locus that is redundantly regulated by AGO4 and AGO6. AGO4 (AGO6)-specific loci are regions that have mC reduction in *ago4* (*ago6*) ≥25%, while mC reduction in *ago6* (*ago4*) <25%. The third group is loci that require both AGO6 and AGO4. This group is defined by "mC reduction in the *ago4ago6* double mutant ≤1.25 × mC reduction in either *ago4* or *ago6*" and "the ratio of mC reduction in *ago4* to mC reduction in *ago6* is between 0.75 and 1.25." The remaining group is loci where DNA hypomethylation is observed in the single mutants, while the *ago4ago6* double mutant shows more severe DNA hypomethylation, indicating more complex genetic interactions between the two AGO proteins (*see* **Note 5**).

4 Notes

1. Download files either using "wget" in the command line or by clicking the link in web browser.
2. To download SRA file (for WT) from NCBI GEO, search "GSM1080803" in Google. From search results, click the first result: "GSM1080803 – GEO Accession viewer." In the GEO webpage, scroll down and click "(ftp)" in the row starting with "SRX/SRX234/SRX234882." Then click "SRR707458." After that, you will see the "SRR707458.sra" file; right-click on it and choose "Copy Link Address" to get the link of the SRA file.

3. In the BRAT-nova analysis, when it is required to write file names into a text file, it is important to put each name in a line. It is better to use the full absolute path of the file. If we want to use relative path, make sure it is accessible within the working directory.
4. The value output from step3_cal_meth_level.pl is a percentage. For example, 0.543 means the methylation level is 0.543%. The largest value is 100, meaning fully methylated.
5. Categorizing the hypo-DMRs of the *ago4ago6* double mutant can be achieved through calculation in Excel when we get the methylation level file "double_vs_WT_hypo_meth_level.txt."

References

1. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452(7184):215–219. doi:10.1038/nature06745
2. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single–base resolution maps of the epigenome in Arabidopsis. Cell 133(3):523–536. doi:10.1016/j.cell.2008.03.029
3. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462(7271):315–322. doi:10.1038/nature08514
4. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S, Downes M, Yu R, Stewart R, Ren B, Thomson JA, Evans RM, Ecker JR (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471(7336):68–73. doi:10.1038/nature09798
5. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W (2010) Genome–wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature 463(7284):1101–1105. doi:10.1038/nature08829
6. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schubeler D (2011) DNA–binding factors shape the mouse methylome at distal regulatory regions. Nature 480(7378):490–495. doi:10.1038/nature10716
7. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD (2011) Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell 146(6):1029–1041. doi:10.1016/j.cell.2011.08.016
8. Zeng J, Konopka G, Hunt BG, Preuss TM, Geschwind D, Yi SV (2012) Divergent whole–genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. Am J Hum Genet 91(3):455–465. doi:10.1016/j.ajhg.2012.07.024
9. Jiang L, Zhang J, Wang JJ, Wang L, Zhang L, Li G, Yang X, Ma X, Sun X, Cai J, Zhang J, Huang X, Yu M, Wang X, Liu F, CI W, He C, Zhang B, Ci W, Liu J (2013) Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. Cell 153(4):773–784. doi:10.1016/j.cell.2013.04.041
10. Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, Wang GL, Meyers BC, Jacobsen SE (2013) Plants regenerated from tissue culture contain stable epigenome changes in rice. elife 2:e00354. doi:10.7554/eLife.00354
11. Zhong S, Fei Z, Chen YR, Zheng Y, Huang M, Vrebalov J, McQuinn R, Gapper N, Liu B, Xiang J, Shao Y, Giovannoni JJ (2013) Single–base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nat Biotechnol 31(2):154–159. doi:10.1038/nbt.2462
12. Gent JI, Ellis NA, Guo L, Harkess AE, Yao Y, Zhang X, Dawe RK (2013) CHH islands:

de novo DNA methylation in near–gene chromatin regulation in maize. Genome Res 23(4):628–637. doi:10.1101/gr.146985.112

13. Schmitz RJ, He Y, Valdes-Lopez O, Khan SM, Joshi T, Urich MA, Nery JR, Diers B, Xu D, Stacey G, Ecker JR (2013) Epigenome–wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res 23(10):1663–1674. doi:10.1101/gr.152538.112
14. Zheng X, Zhu J, Kapoor A, Zhu JK (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J 26(6):1691–1701. doi:10.1038/sj.emboj.7601603
15. Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, Wang P, Zhou J, Liang G, Liu N, Wang C, Zhu JK (2015) Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA–directed DNA methylation. EMBO J 34(5):581–592. doi:10.15252/embj.201489453
16. Harris EY, Ounit R, Lonardi S (2016) BRAT–nova: fast and accurate mapping of bisulfite–treated reads. Bioinformatics 32(17):2696–2698. doi:10.1093/bioinformatics/btw226

Chapter 17

Structural and Functional Characterization of Plant ARGONAUTE MID Domains

Filipp Frank and Bhushan Nagar

Abstract

The interaction of small silencing RNA 5′ nucleotides with the MID domain of ARGONAUTE (AGO) proteins provides an anchor point that contributes to strong binding between RNA and protein. The following protocols describe the necessary procedures to characterize the structure of AGO MID domains using X-ray crystallography as well as their interaction with nucleotides that mimic the 5′ end of small silencing RNAs using two-dimensional NMR spectroscopy.

Key words ARGONAUTE, MID domain, Small silencing RNAs, Protein expression, Protein purification, X-ray crystallography, NMR

1 Introduction

Small silencing RNAs bind to ARGONAUTE (AGO) proteins to carry out their gene regulatory effects. The 5′ end of a small RNA (typically 20–25 nucleotides in length) is anchored to the AGO protein through specific interactions between the MID domain and the 5′ nucleotide, which is usually in the form of a 5′ monophosphate. The interaction occurs in a binding pocket that is lined with positively charged amino acids for coordination of the phosphate group. Specific recognition of the nucleobase is achieved via a rigid loop in the MID domain, the *nucleotide specificity loop* [1, 2]. Selective binding of the nucleobase is particularly important in plants. *Arabidopsis thaliana* (*Arabidopsis*) encodes ten AGO (AtAGO) proteins and a diverse set of small RNA classes that are distributed between them. This small RNA sorting in *Arabidopsis* is mainly directed by the 5′ nucleotide [3]. Using a combination of X-ray crystallography and nuclear magnetic resonance (NMR) with recombinant MID domain proteins and nucleoside monophosphates, we have characterized these interactions in vitro.

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_17,

We solved crystal structures of MID domains from different AGOs and their complexes with nucleotides and dinucleotides [1, 2, 4]. In particular, we showed that the MID domains of AtAGO1, AtAGO2, and AtAGO5 can discriminate 5′ nucleotide identity: AtAGO1 preferentially binds to U, AtAGO2 is selective for A, and AtAGO5 has a bias toward C [1]. These studies confirmed the selectivity of AtAGOs for different 5′ nucleotides observed in vivo [3]. Here, we describe the methods to characterize AGO MID domain structures using X-ray crystallography and determine their affinities to ligands that mimic the structure of a small RNA 5′ end (e.g., nucleoside monophosphates). The detailed procedures include overexpression and purification of native, ^{15}N-labeled, or selenomethionine-labeled MID domain proteins, as well as determination of dissociation constants using ^{1}H–^{15} N–HSQC NMR spectroscopy and preparation of MID domain crystals/ligand complexes for structure determination by X-ray crystallography.

2 Materials

2.1 Purification of ARGONAUTE MID Domains

2.1.1 Protein Expression

1. Plasmid for bacterial expression of SUMO-fusion proteins (e.g., pSmt3 [5]).
2. cDNA for the AGO of interest.
3. *E. coli* strain BL21 DE3 for protein expression.
4. Autoinducing medium: 10 g/L tryptone, 5 g/L yeast extract, 20 mL/L 50× M solution (1.25 M disodium phosphate (Na_2HPO_4, 177.5 g/L), 1.25 M monopotassium phosphate (KH_2PO_4, 170.3 g/L), 2 M ammonium chloride (NH_4Cl, 107 g/L), 0.25 M ammonium sulfate [$(NH_4)_2SO_4$, 33 g/L], 0.5 M sodium chloride (NaCl, 29.3 g/L)), 20 mL/L 50 × 5052 solution [25% (v/v) glycerol, 2.5% (w/v) glucose, 10% (w/v) alpha-lactose monohydrate], 2 mM $MgSO_4$.
5. 10× M9 salts, autoclaved: 260 mM sodium phosphate dibasic heptahydrate ($Na_2HPO_4{\cdot}7H_2O$, 70 g/L), 220 mM KH_2PO_4 (30 g/L), 86 mM NaCl (5 g/L), 93 mM NH_4Cl (5 g/L), and use $^{15}NH_4Cl$ for production of ^{15}N-labeled protein to be used in ^{1}H–^{15}N–HSQC NMR experiments.
6. Luria–Bertani (LB) medium: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl. Autoclave.
7. Defined medium for production of Se–Met-labeled proteins. For 1 L of medium, mix the following: 1× M9 salts, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.1 mM iron (III) chloride ($FeCl_3$), 0.5% glucose, 1 mg/mL biotin, 1 mg/mL thiamin, and 50 μg/mL kanamycin.
8. Prepare the following compounds for addition to the medium when OD_{600} of ~0.8–1.0 has been reached before induction

with isopropyl β–D–1–thiogalactopyranoside (IPTG): 100 mg/L of culture each lysine, phenylalanine, and threonine, 50 mg/L of culture each isoleucine, leucine, and valine, 60 mg/L of culture selenomethionine.

9. Defined medium for production of ^{15}N-labeled proteins: 1× M9 salts (prepared with $^{15}NH_4Cl$), 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.1 mM $FeCl_3$, 0.5% glucose, 1 mg/mL biotin, 1 mg/mL thiamin, 50 μg/mL kanamycin.

2.1.2 Protein Purification

1. Ulp1 protease.
2. Buffer NiA: 25 mM Tris pH 8.0, 500 mM NaCl, 10 mM imidazole.
3. Buffer NiB: 25 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazole.
4. Buffer SPA: 25 mM Tris pH 8.0, 50 mM NaCl.
5. Buffer SPB: 25 mM Tris pH 8.0, 1 M NaCl.
6. Buffer GF: 25 mM MES pH 6.5, 200 mM NaCl, 3 mM dithiothreitol (DTT).

2.2 Equipment and Other Materials

1. Ni–NTA column (e.g., HisTrap HP 5 mL, GE Healthcare Life Sciences).
2. Cation exchange column (e.g., HiTrap SP HP 5 mL, GE Healthcare Life Sciences).
3. HiLoad 16/600 Superdex 75 pg (GE Healthcare Life Sciences).
4. NMR spectrometer (e.g., 600 MHz spectrometer, Bruker).
5. Liquid handling system (e.g., Phoenix, Art Robbins Instruments).
6. Magnetic susceptibility-matched NMR tube (Shigemi Inc.).
7. Sonicator or emulsifier.

3 Methods

3.1 Preparation of Large Amounts of Protein for Structural and Functional Studies

3.1.1 Construct Design

AGO MID domains express and purify well with an N-terminal SUMO-fusion tag. In order to identify a suitable expression construct of an AGO MID domain of interest, the domain boundaries should be determined by preparing amino acid sequence alignments with available MID domain structures found in the Protein Data Bank (e.g., 4G0O, 4G0M, 4G0X, or 3LUC). The construct is then cloned into the pSmt3 vector [5] using the *BamH*I restriction site at the 5′ end and any other available restriction site for the 3′ end. Using *BamH*I ensures that cleavage of the fusion protein by the SUMO protease Ulp1 only leaves a single residue at the

N-terminus of the MID domain. If a *BamH*I site is present in the coding sequence, an enzyme with a non-palindromic restriction site such as *Bsa*I can be utilized to generate an appropriate insert that matches the *BamH*I overhang of the vector.

3.1.2 Protein Expression: Native Protein

Expression using autoinducing medium [6] ensures maximum protein yield per culture volume. A convenient procedure is to inoculate four flasks of 1 L each of autoinducing medium with ~10 mL healthy bacterial culture (growing in log phase) in the afternoon (~3 pm) and letting the culture grow at 30 °C overnight. Autoinducing medium is well buffered so that cultures are not very sensitive to the time of harvest. It is critical, however, to let the culture grow to high density, so that strong induction is achieved.

1. Inoculate 5 mL of LB medium with a single colony of BL21 DE3 carrying the expression plasmid (e.g., pSmt3 containing the MID domain coding sequence) and grow overnight at 37 °C.
2. Transfer 1–5 mL of overnight culture into 50 mL of LB and grow at 37 °C during the course of the day.
3. Inoculate four flasks of 1 L each of autoinducing medium and grow overnight at 30 °C.
4. In the morning of the next day, harvest cells by spinning at 4000 × *g* for 10 min.
5. Resuspend pellets in ~50 mL of buffer NiA per liter of culture.

3.1.3 Protein Expression: Se–Methionine Labeled

Incorporation of selenomethionine (Se–Met) into AGO MID domain proteins has been successfully used to solve their crystal structures when the phasing technique of molecular replacement failed. The procedure involves protein overexpression using IPTG and in the presence of high concentrations of amino acids known to suppress the bacterial methionine biosynthesis pathway and in the absence of methionine [7]. Methionine is replaced by selenomethionine, which is then incorporated into newly produced proteins.

1. Prepare four times 1 L of defined medium for production of SeMet-labeled protein.
2. Inoculate 50 mL of LB medium with a single colony of BL21 DE3 carrying the expression plasmid (pSmt3 containing the MID domain coding sequence) and grow overnight at 37 °C.
3. Centrifuge the 50 mL starter culture at 4000 × *g* for 10 min, and resuspend in 40 mL of defined medium for production of SeMet-labeled protein (taken from the previously prepared 4 L).
4. Transfer 10 mL of the resuspended cells into each the four 1 L samples of defined medium for production of SeMet-labeled protein.

5. Grow at 37 °C to an OD_{600} of ~0.8–1.0, and then add 100 mg/L each lysine, phenylalanine, and threonine, 50 mg/L each isoleucine, leucine, valine, and 60 mg/L selenomethionine.
6. After 30 min induce overexpression by addition of 0.5 mM IPTG.
7. 4 h after induction, harvest cells by spinning at 4000 × *g* for 10 min.
8. Resuspend cell pellets in ~25 mL of buffer NiA per liter of culture.

3.1.4 Protein Expression: ^{15}N Labeled

^{15}N–^{1}H–HSQC NMR experiments require protein, which is isotopically labeled with ^{15}N. For preparation of ^{15}N-labeled protein, a defined minimal medium is used, in which the only source of nitrogen is ammonium chloride containing the ^{15}N isotope. Overexpression is then achieved by IPTG induction (*see* **Note 1**).

1. Prepare 4 L of defined medium containing $^{15}NH_4Cl$.
2. Inoculate 50 mL of LB medium with a single colony of BL21 DE3 carrying the expression plasmid (pSmt3 containing the MID domain coding sequence) and grow overnight at 37 °C.
3. Centrifuge the 50 mL starter culture at 4000 × *g* for 10 min, and resuspend in 40 mL of defined medium containing $^{15}NH_4Cl$ (taken from the previously prepared 4 L).
4. Transfer 10 mL of the resuspended cells into each of the four 1 L samples of defined medium containing $^{15}NH_4Cl$.
5. Grow at 37 °C to an OD_{600} of ~0.8–1.0, and then induce overexpression by addition of 0.5 mM IPTG.
6. 4 h after induction, harvest cells by spinning at 4000 × *g* for 10 min.
7. Resuspend cell pellets in ~25 mL of buffer NiA per liter of culture.

3.1.5 Protein Purification

Methods are the same for purification of native or labeled proteins.

1. Lyse cells using a sonicator or emulsifier.
2. Spin down cell debris at ~20,000 × *g* for at least 30 min.
3. Load the cleared lysate onto a Ni–NTA column equilibrated in buffer NiA. Wash unbound sample generously with buffer NiA, and then elute with a gradient of 0–100% buffer NiB over 10 column volumes.
4. Pool the peak fractions, add ~0.5–1 mg of Ulp1 protease, and incubate at 4 °C for >1 h (*see* **Note 2**).
5. Dilute the cleaved protein at least fivefold with 25 mM Tris pH 8.0, and immediately load onto a cation exchange column equilibrated with buffer SPA (*see* **Note 3**). Wash unbound

sample with buffer SPA, and elute with a gradient of 0–50% buffer SPB over 25 column volumes.

6. Pool peak fractions and concentrate protein to <2 mL.
7. Purify the sample over a Superdex 75 gel filtration column using buffer GF (*see* **Note 4**).
8. Pool peak fractions and concentrate the pure protein: ~15 mg/mL (~1 mM) for crystallization or ~3 mg/mL for ^{15}N-labeled samples (equiv. to ~200 μM) to be used in NMR experiments.

3.2 NMR Titration Experiments

NMR titrations are a useful tool to determine the dissociation constants of protein–ligand interactions, especially for low-affinity interactions where other techniques might fail. Specific interaction of a ligand with protein results in changes in chemical shifts Δδ of the protein's backbone amide peaks. The addition of increasing amounts of ligand to a sample will result in chemical shift changes that are proportional to the fractional occupancy (θ) of the protein:

$$\Delta\delta \propto \theta;$$

This signal can then be used to determine the dissociation constant for the interaction.

3.2.1 Preparation of Nucleoside Monophosphate Solutions

Nucleoside monophosphates are mimics of the 5′ ends of small RNAs, which are the natural substrates of most AGO MID domains. The protocol below describes how to prepare solutions of nucleoside monophosphates that can be used in NMR titration as well as crystallization experiments. However, other molecules may also be tested for their interaction with AGO MID domains. We have, for example, determined dissociation constants of MID domains with nucleoside triphosphates, analogs of the mRNA 5′ cap structure [4], and chemically modified nucleotide analogs as well as dinucleotides [8].

1. Weigh out enough nucleoside monophosphate to make ~1 mL of a 1 M solution (approximately 350 mg). Dissolve in water at a concentration of ~1.2 M (>1 M in order to adjust to 1 M after determination of the concentration). Adjust to pH 7 using 1 M sodium hydroxide (NaOH). Measure pH using pH paper.
2. Measure concentration by diluting 1000-fold and using correct molar extinction coefficients [9]. Adjust concentration to 1 M.
3. Store aliquots at −20 °C.

3.2.2 Titration Experiments

1. Prepare 300–350 μL of a solution of ^{15}N-labeled protein at 0.1–0.25 mM.
2. Add the protein solution to a magnetic susceptibility-matched NMR tube, and measure a ^{15}N–^{1}H–HSQC NMR spectrum.

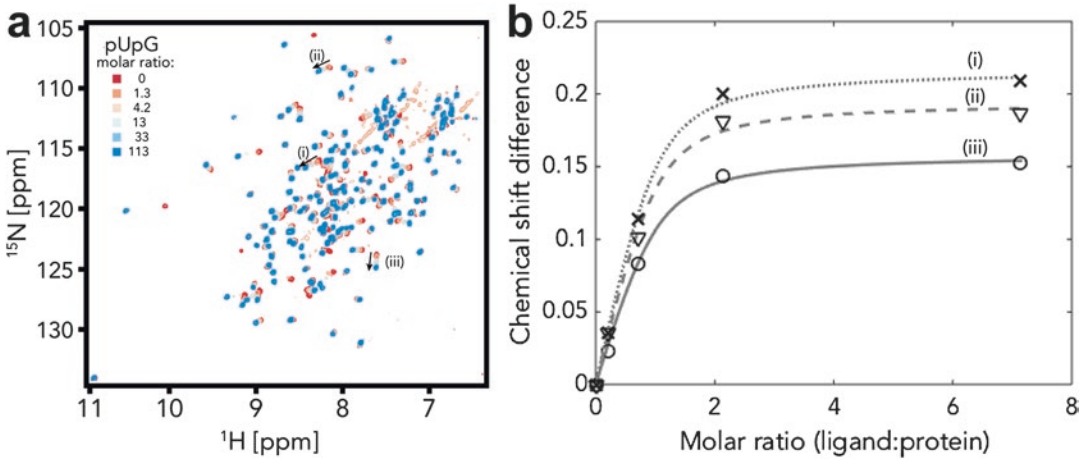


Fig. 1 Example ^{1}H–^{15}N–HSQC NMR spectra and fitting procedure. (**a**) ^{1}H–^{15}N–HSQC NMR spectra of *Arabidopsis thaliana* AGO1 MID domain with increasing amounts of the dinucleotide pUpG. (**b**) Binding curve of data extracted from the titration spectra shown in (**a**). The initial protein concentration was 140 μM, and the dissociation constant of the interaction determined by least-squares fitting is 21.0 ± 1.5 μM

3. Remove the protein solution from the NMR tube. Since there is loss of sample during handling, measure the remaining volume of protein using a pipette. Add an appropriate amount of ligand to the protein. This amount will depend on the affinity of the reaction. A good starting point is ~100 μM (*see* **Note 5**).
4. Add the protein–ligand complex solution back to the NMR tube, and measure a ^{15}N–^{1}H–HSQC NMR spectrum. Plot the ^{15}N–^{1}H–HSQC spectrum, and check for peaks that have changed position after addition of ligand. The amount of ligand to add in the subsequent titration steps will depend on the extent to which peaks have shifted during this first step of titration: if there are peaks that have shifted considerably, the amount of ligand added in the second titration step should not exceed a threefold increase in ligand to protein ratio. If chemical shift changes were minimal, this may indicate low binding affinity, and the ratio may be increased more drastically. A typical spectrum for a titration experiment is shown in Fig. 1a. The extent of chemical shift changes here should serve as a guide for the kinds of chemical shift changes expected in such an experiment.
5. Repeat **steps 3** and **4** with addition of more ligand. Approximately threefold increases of the molar ratio of ligand to protein will generally yield suitable coverage of fractional occupancy for a reliable determination of a dissociation constant for the interaction. A total number of at least five titration points (including the zero ligand condition) should be collected.

3.2.3 Data Analysis

In order to determine the dissociation constant of the measured interaction, the data will be plotted using NMR visualization software. The peak positions of moving peaks will then be determined, and chemical shift differences will be analyzed to calculate the dissociation constant of the interaction.

For the interaction of a protein, P, with ligand L:

$$P_f + L_f \leftrightarrow PL;$$

where P_f and L_f are the free protein and free ligand concentrations, respectively, and PL is the concentration of protein–ligand complex. The dissociation constant for the interaction is then (Eq. 1):

$$K_D = \frac{[P_f][L_f]}{[PL]} \tag{1}$$

Free protein or ligand concentrations are not measured in the NMR experiments. Thus, we rewrite the above equation in terms of total protein and ligand concentrations, $[P_f] = [P_T] - [PL]$ and $[L_f] = [L_T] - [PL]$. This gives (Eq. 2):

$$K_D = \frac{([P_T] - [PL])([L_T] - [PL])}{[PL]} \tag{2}$$

The NMR experiments measures the partial occupancy of protein, $\theta = \frac{[PL]}{[P_T]}$. The solution of Eq. 2 in terms of fractional occupancy, θ, is a quadratic equation with the following solution (Eq. 3):

$$\theta = \frac{[PL]}{[P_T]} = \frac{[P_T] + [L_T] + [K_D] - \sqrt{([P_T] + [L_T] + [K_D])^2 + 4[P_T][L_T]}}{2 \times P_T} \tag{3}$$

Fractional occupancy varies between 0 and 1, whereas the experimentally determined chemical shift difference, $\Delta\delta$, varies between 0 and the maximal chemical shift difference, $\Delta\delta_{max}$. Since these terms are proportional to each other, we can write (Eq. 4):

$$\Delta\delta = \Delta\delta_{max} \times \theta \tag{4}$$

A more detailed description of the equations derived above can be found in [10].

In order to use Eqs. 3 and 4 for the determination of the dissociation constant, use the following procedure:

1. For each titration point, i, determine the total protein and ligand concentrations, $[P_{T,i}]$ and $[L_{T,i}]$, respectively, in the

sample. Take into account the volume of sample recovered from the NMR tube and the volume of ligand added in each step and how this affects these concentrations.

2. Plot the ^{15}N–^{1}H–HSQC spectra using an appropriate NMR visualization software (e.g., *NMRView J*).
3. Identify at least three peaks whose ^{1}H and ^{15}N chemical shifts, δ_H and/or δ_N, respectively, change significantly during the titration experiment so that a chemical shift difference at each ligand to protein ratio can be determined.
4. Determine the chemical shifts of the peak centers for each selected peak, j, and each titration step, i: $\delta_{H,i,j}$ and $\delta_{N,i,j}$.
5. The total chemical shift difference $\Delta\delta$ of peak j in titration step i is the distance of peak j between the spectra collected at titration step i compared to the zero ligand condition. This distance can be calculated as $\Delta\delta_{i,j} = [(\Delta\delta_{H,i,j})^2 + (0.2 \times \Delta\delta_{N,i,j})^2]^{1/2}$ (*see* **Note 6**).
6. Perform a direct least squares fitting procedure on the calculated data ($\Delta\delta_{i,j}$, $[P_{T,i}]$, and $[L_{T,i}]$) using (Eq. 5).

$$\Delta\delta_{i,j} = \Delta\delta_{i,j,\max} \times \theta = \Delta\delta_{i,j,\max} \times \frac{[P_T] + [L_T] + [K_D] - \sqrt{([P_T] + [L_T] + [K_D])^2 + 4[P_T][L_T]}}{2 \times P_T} \quad (5)$$

and the maximal chemical shift difference, $\Delta\delta_{i,j,\max}$, and the dissociation constant, K_D, as adjustable parameters. A typical plot of a titration curve is shown in Fig. 1b.

3.3 Crystallization and Structure Determination

The structures of various eukaryotic AGO MID domains have been solved by X-ray crystallography [1, 2, 11, 12]. Standard procedures for crystallization and structure determination will be used: (1) large-scale screening of crystallization conditions using commercially available crystallization screens, (2) optimization of hit conditions to yield high-quality crystals suitable for structure determination, (3) data collection, and (4) structure determination.

3.3.1 Identification of Crystallization Conditions

Once pure protein has been generated, screens are carried out to identify crystallization conditions. The procedure described here was carried out using sitting drops prepared by a liquid handling system with 0.2 μL drop sizes.

1. Prepare ~25 μL of protein solution per crystallization screen (96 conditions) to be tested. Suitable starting conditions are 15 mg/mL, 10 mg/mL, and 5 mg/mL in buffer GF. In past studies, most hits were achieved at 15 mg/mL. Additionally, conditions in the presence of ligand should be screened (*see* **Note 7**).

2. Set up 96-well plates by depositing sitting drops of 0.2 μL protein solution and equal amounts of crystallization screen solutions. Incubate at an appropriate temperature, typically room temperature and 4 °C (*see* **Note 8**).
3. Monitor for crystal growth daily.

3.3.2 Optimization of Initial Hit Conditions

The following protocol describes minimal standard procedures that yielded high-quality crystals of AGO MID domains in the past. More elaborate procedures (crystal seeding, additive screens, etc.) may be necessary if these procedures fail to produce crystals suitable for structure determination.

1. Set up small-scale screens around initial hit conditions in hanging drop 24-well format with 500 μL of the crystallization solution in the well and 0.5–2 μL of protein and equal amounts of screen solution on siliconized cover slides. Pipetting of these screens is done by hand.
2. Initially, screen conditions around the original hit solution by varying pH, salt concentration, precipitant concentration, addition of glycerol (up to ~7.5%), and/or temperature.
3. Monitor crystal growth daily, and set up new 24-well screens by hand to refine conditions, if necessary.

3.3.3 Soaking of Crystals with Ligands for Structure Determination of Protein–Ligand Complexes

Crystal structures of MID domains in complex with nucleoside monophosphates or other ligands have been solved by soaking crystals of apo–MID domains in solutions containing high concentrations of the ligand (millimolar range).

1. Once a condition has been identified that reliably produces high-quality crystals, set up a 24-well plate of that condition to generate large numbers of crystals for optimization of soaking conditions.
2. Previously, MID domains have been crystallized in conditions that contained high concentrations of either phosphate or sulfate. These ions specifically bind to the MID domain's natural binding site of the phosphate group from the small silencing RNA's 5′-nucleotide. Thus, they interfere with binding of nucleotides and other ligands to this site and need to be removed. In order to achieve this, crystals grown in these conditions should be transferred to solutions that do not contain phosphate or sulfate. The ions may be replaced by the ligand of choice; however, screening of large numbers of conditions that do not contain these ions may be necessary to identify solutions that do not dissolve the crystals or crack the crystals (*see* **Note 9**).

   In order to identify such a solution, prepare >2 μL drops of solutions from crystallization screens on standard cover slides used for hanging drop crystallization. Large drop sizes

are recommended so that evaporation does not dry out the drops and/or dramatically change the concentrations of buffer components. Transfer a single crystal from the original crystallization drop into a solution to be tested, and monitor the crystal's integrity under a microscope. Most conditions will start dissolving and/or cracking crystals over the course of a few seconds or minutes. In our laboratory, this method was successful in 50% of cases.

3. Once a condition that keeps crystals stable in the absence of phosphate or sulfate has been identified, test the addition of different cryoprotectants and monitor crystal integrity.
4. Once a cryoprotecting condition has been established, prepare ~2 μL drops of this condition with increasing amounts of MID domain ligand. In order to achieve saturation of ligand binding sites in the crystal, the concentration of ligand should be >tenfold higher than the dissociation constant for the interaction determined using NMR. Transfer single crystals from the original crystallization conditions to these drops, soak for ~30 s to 5 min, and then flash freeze crystals in liquid nitrogen.

   Alternatively, prepare a 24-well crystallization screen with cryoprotecting condition in the reservoir. Then transfer crystals to drops of this condition (as in 4(a)), and soak for longer amounts of time.
5. Screen crystals by collecting complete X-ray diffraction data sets and checking for the presence electron density representing bound ligand in different Fourier maps. This step can only be done after the initial MID domain structure has been solved.

3.3.4 Data Collection

In our lab, AGO MID domains have produced high-quality crystals, and complete data sets suitable for structure determination could be collected on a home lab X-ray source. If a home lab X-ray source is available, structure determination procedures are highly facilitated: screening of crystal quality, choice of cryoprotectant solutions, and structure determination of MID domain–ligand complexes can all be carried out efficiently without the delay of sending crystals to a synchrotron source.

3.3.5 Structure Determination

Previous structures of AtAGO MID domains were solved either using Se–methionine-labeled protein and single anomalous dispersion (SAD) phasing (AtAGO1 and AtAGO2 [1]) or by molecular replacement using the structure of hAGO2 or AtAGO1 MID domain as a search model (AtAGO1 [11] and AtAGO5 [1]). Currently there are crystal structures of MID domains from six different eukaryotic AGO proteins available (hAGO2, QDE-2, AtAGO1, AtAGO2, AtAGO5, and PIWI) and additional structures of bacterial AGOs. Consequently, standard molecular replacement methods will likely be successful in the determination of future MID domain structures.

4 Notes

1. An autoinducing strategy can be employed with specialized media preparations [6]; however, IPTG induction has proven more efficient for MID domains in our hands.
2. In most cases cleavage by Ulp1 occurs very rapidly. The time it takes to equilibrate the next column (cation exchange) is generally enough to cleave all of the sample.
3. MID domains proteins tend not be stable in low salt concentrations after cleavage of the SUMO tag. Some precipitation may occur when diluting with low-salt buffer for binding to the cation exchange column. To maximize protein yield, it is important to minimize the time the protein spends in the low-salt buffer by diluting the sample immediately before loading it onto the column.
4. This buffer worked well for previously studied AGO MID domains. If there are issues with protein stability, optimization of this buffer may be required.
5. Even at the highest previously measured affinity of ~100 μM between and AGO MID domain and a nucleotide ligand, the fractional occupancy of protein is still below 50% at this starting ligand concentration, and an appropriate binding curve can be established.
6. We use $0.2 \times \Delta\delta_N$ in order to account for the higher sensitivity in chemical shift changes observed in the δ_N measurement compared to the measurement of δ_H.
7. In our hands, co-crystallization has not yielded crystals of MID domain–ligand complexes. However, every protein behaves differently, and we recommend also screening crystallization conditions in the presence of ligand.
8. The first eukaryotic AGO MID domain (hAGO2) crystallized in our lab produced high-quality crystals at 4 °C. Subsequent screens of other MID domains were all initiated at 4 °C and yielded suitable crystals.
9. This step can be very tedious and laborious. In at least one case in our lab, we identified a condition that kept crystals intact in the absence of stabilizing sulfate or phosphate very quickly by testing all other conditions that had yielded crystals in initial 96-well screens. These conditions had failed to produce high-quality crystals during optimization steps but turned out to be useful for soaking of ligands.

Acknowledgements

We thank Matthew Hassler and Masad Damha for preparation of the synthesis of pUpG which was used to generate the NMR spectrum shown in Fig. 1.

References

1. Frank F, Hauver J, Sonenberg N, Nagar B (2012) Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. EMBO J 31(17):3588–3595. doi:10.1038/emboj.2012.204
2. Frank F, Sonenberg N, Nagar B (2010) Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature 465(7299):818–822. doi:10.1038/nature09039
3. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis Argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/j.cell.2008.02.034
4. Frank F, Fabian MR, Stepinski J, Jemielity J, Darzynkiewicz E, Sonenberg N, Nagar B (2011) Structural analysis of 5′-mRNA-cap interactions with the human AGO2 MID domain. EMBO Rep 12(5):415–420. doi:10.1038/embor.2011.48
5. Mossessova E, Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Mol Cell 5(5):865–876
6. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41(1):207–234
7. Doublie S, Carter C (1992) Preparation of selenomethionyl protein crystals. In: Ducruix A, Giegé R (eds) Crystallization of nucleic acids and proteins: a practical approach. Oxford University Press, Oxford, UK
8. Deleavey GF, Frank F, Hassler M, Wisnovsky S, Nagar B, Damha MJ (2013) The 5′ binding MID domain of human Argonaute2 tolerates chemically modified nucleotide analogues. Nucleic Acid Ther 23(1):81–87. doi:10.1089/nat.2012.0393
9. Cavaluzzi MJ, Borer PN (2004) Revised UV extinction coefficients for nucleoside-5′-monophosphates and unpaired DNA and RNA. Nucleic Acids Res 32(1):e13. doi:10.1093/nar/gnh015
10. Wilkinson KD (2004) Quantitative analysis of protein-protein interactions. Methods Mol Biol 261:15–32. doi:10.1385/1-59259-762-9:015
11. Zha X, Xia Q, Yuan YA (2012) Structural insights into small RNA sorting and mRNA target binding by Arabidopsis Argonaute Mid domains. FEBS Lett 586(19):3200–3207. doi:10.1016/j.febslet.2012.06.038
12. Boland A, Tritschler F, Heimstadt S, Izaurralde E, Weichenrieder O (2010) Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. EMBO Rep 11(7):522–527. doi:10.1038/embor.2010.81

Chapter 18

Identification and Analysis of WG/GW ARGONAUTE-Binding Domains

Andrzej Zielezinski and Wojciech M. Karlowski

Abstract

WG/GW domains recruit ARGONAUTE (AGO) proteins to distinct silencing effector complexes using combinations of just two amino acids: tryptophan (W) and glycine (G), forming a wide arsenal of highly simplified interaction surfaces. These unstructured domains exhibit very low sequence identity and excessive length polymorphism, which makes identification of new AGO-binding proteins a challenging task as they escape detection with standard sequence comparison-based methods (e.g., BLAST, HMMER).

In this chapter, we explain the use of tools for prediction of AGO-binding WG/GW domains in protein sequences. We also show how to computationally explore an up-to-date information about AGO-interacting proteins and discover new properties of WG/GW domains. Finally, we encourage readers to explore the game-like web application for in silico designing/modifying AGO-binding sequences as well as modeling mutagenesis experiments and predicting their potential effect on AGO-binding activity.

Key words WG/GW domain, AGO-binding domain, Protein domain, Sequence analysis, Web application, ARGONAUTE

1 Introduction

The glycine/tryptophan binary code (termed WG/GW)—present throughout the eukaryotic kingdom—is a critical determinant for capturing and recruitment of ARGONAUTE (AGO) proteins to key effectors of RNA silencing pathways [1]. Furthermore, some plant viruses have co-opted the polypeptides that mimic host-encoded WG/GW motifs and act as baits for the plant AGOs to hack into the host's defense systems [2–5]. Recently, it has been shown that NEF protein in HIV-1 virus binds human AGO2 through its conserved GW motifs and inhibits the slicing activity of AGO2 resulting in the suppression of miRNA-induced silencing [6].

Although the core WG/GW sequence signatures are clearly evolutionarily conserved, the more precise definition of AGO-binding domains is challenging. First, sequences of the WG/GW

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_18,

domains are remarkably adaptive to amino acid substitutions displaying little to no recognizable identity (<25%), even between closely related species [1, 7]. Second, AGO-binding domains vary greatly in both sequence lengths (from 22 residues in yeast protein Tas3 [8] to 750 in SPT5) and a number of WG/GW repeats (from 1 to 50 copies). Third, AGO-binding sites represent a class of intrinsically disordered domains (IDD), which do not fold into well-organized globular structures in the absence of AGO partners but exist rather as ensembles of rapidly interconverting conformations [1]. Finally, WG/GW domains have been found in many different protein families of unrelated domain organization—ranging from polymerase V subunit (NRPE1) [9] to putative oxidoreductase (WGRP1) [7] and transcription elongation factors (SPT5) [10] in plants, to trinucleotide repeat-containing proteins (GW182) [11] and prion proteins [12] in animals. To see how far evolution could push the properties of the minimalistic WG/GW binding surfaces, El-Shami and colleagues engineered a plant NRPE1 variant in which the WG/GW repeat region was swapped with the corresponding AGO-binding region from human GW182. The chimeric protein rescued most of the polymerase V function, even though the sequence similarities between the transplanted domains were undetectable [9].

Such an excessive variability of WG/GW domains makes the identification of new AGO-binding proteins very difficult basing solely on standard homology detection methods (e.g., BLAST, HMMER) [7], probably explaining why only 18 proteins in all three domains of life have been shown to interact with AGO proteins [1]. Furthermore, the WG/GW domains in these proteins have not yet been annotated in any of the central protein databases (UniProt [13], RefSeq [14]) and/or domain resources (Pfam [15], InterPro [16]). Also, going through the literature to gain a general view of WG/GW domains is impractical, since each expert research group often focuses only on a single species-specific protein.

This chapter describes how to browse, analyze, predict, and design AGO-binding proteins using computational tools that are available via a web portal (Whub) at http://www.combio.pl/whub. This knowledge base provides access to prediction tools and current information about experimentally verified AGO-interacting proteins. This chapter can be read alongside the help pages on the Whub website and publications describing the service [1].

2 Materials

1. Computer with Internet connection.
2. Web browser (e.g., Google Chrome, Internet Explorer, Mozilla Firefox, Safari).
3. Python Software (optional).

3 Methods

All tools used in this chapter are available through the Whub portal (http://www.combio.pl/whub). The methods will be broken down into three parts to describe the browsing, analyzing, and predicting of AGO-binding proteins.

3.1 Accessing Information on Experimentally Confirmed AGO-Binding Proteins

A catalog of known AGO-binding proteins in eukaryotes and viruses is available from the Whub homepage, under the "GW proteins" menu. The pages for protein record are the center point for viewing current knowledge about a given protein. This information comes from literature mining and includes protein domain organization, experimentally verified functional sequence regions, mutagenesis studies, and cross-links to other databases. This section describes how to access the main information that Whub stores for each protein family;

1. Click on the "GW proteins" from the menu to access the list of organisms.
2. Choose "Arabidopsis" from the drop-down menu.
3. Click on the "NRPE1" protein to go to the protein record page (Fig. 1).

3.1.1 Viewing Domain Organization

Beside basic information about the protein (UniProt accession, description, length, and number of WG/GW motifs) (Fig. 1a), a graphical view of the domain organization (Fig. 1b) shows a representation of the location of all tryptophan occurrences (vertical lines) in the context of other domains and the entire sequence. In this example, NRPE1 protein contains three N-terminal domains related to RNA polymerase, AGO-binding domain, and C-terminal glutamine-rich sequence. Mouse over a domain to see a tooltip providing the domain's description and to highlight its positions in the full-length protein sequence (Fig. 1e).

3.1.2 Inspecting Experimentally Verified, Functional Sequence Regions, and Mutagenesis Studies

The "Regions" section in the protein record page (Fig. 1c) shows a list of confirmed functional sequence regions along with information about their corresponding sequence positions and literature references. For example, the protein region at position 1280–1710 was shown by El-Shami and others to interact with AGO4 [9]. Likewise, the "Mutagenesis" window (Fig. 1d) contains experimental data about the effects of sequence mutations on the protein's phenotype related to gene silencing (e.g., interaction with protein–effector complexes such as RISC or CCR4–NOT and PAN2–PAN3 deadenylase complexes). Hovering a mouse over these features underlines the corresponding region in the domain organization map as well as highlights the corresponding fragment in the full-length sequence (*see* **Note 1**).

Fig. 1 Protein record page for *Arabidopsis* NRPE1 (polymerase V subunit). (**a**) Basic information about the protein. (**b**) Domain organization showing a visual representation of the location of all Trp occurrences (vertical lines). (**c**) Information about experimentally verified protein regions. Hovering a mouse cursor over any feature will highlight its position in the full-length sequence (**e**) and mark position in protein map (**b**). (**d**) Information about reported mutagenesis experiments performed in the protein. (**e**) Full-length sequence; highlighted is the region corresponding to one of the mutagenesis experiments

3.1.3 Browsing Articles Relevant to AGO-Binding Proteins

In addition to the "References" section of the protein record page, which contains the literature citations of publications that have been used as the sources to annotate the entry sequence, Whub offers a handy way to explore the bibliographic citations on WG/GW proteins.

1. Click on the "Papers" from the menu to see a graphical list of papers referenced in Whub. Literature citations are shown as cards, each containing information about the first author, journal, and year of the publication.
2. Select "AGO binding" from the "Filters and sorting" section to view on the fly the articles concerning GW proteins.
3. Select "research" articles ("Article type" selection bar) concerning "plants" ("Taxonomy" selection bar), and sort the list of articles by "journal." This will narrow the search to scientific articles concerning AGO-binding proteins in plants sorted by the journal.
4. Click on the card related to the article by He et al. (2009) [10]. It will display summarized information about the paper.
5. Beside the article's abstract that is automatically retrieved from PubMed, the paper record page presents a list of AGO-binding proteins that were the main subject of the study and additionally provides information on how many experiments were performed for a given protein. In this study, He and others conducted five experiments on SPT5 protein and one on NRPE1.
6. Click on the icon of "SPT5" protein to go to the protein record, and see a list of all experiments (from all publications).

3.2 Discovering Properties of AGO-Binding Motifs

The Whub portal offers a web application (Domain Analyzer) for interactive exploration of various properties of AGO-binding domains (e.g., length, amino acid composition, the number of WG/GW copies) based on bioinformatics inference [1] of several thousand W-containing motifs from experimentally verified plant and animal AGO-binding domains.

1. From the sidebar menu "Domain Analyzer," choose "AGO-Plants."
2. The resulting page has three separate panels: "General composition," "PSSM," and "Motif sequences," which will be described in the following sections.

3.2.1 Defining Overall Amino Acid Composition of AGO-Binding Domains

The interactive donut chart (Fig. 2a) shows the overall amino acid content of AGO-binding domains indicating that small and charged amino acids constitute 83% of all residues. The table lists amino acid residues with their corresponding log–odd scores, which indicate the likelihood of finding a given amino acid in the analyzed domain rather than in nonfunctional WG/GW-containing

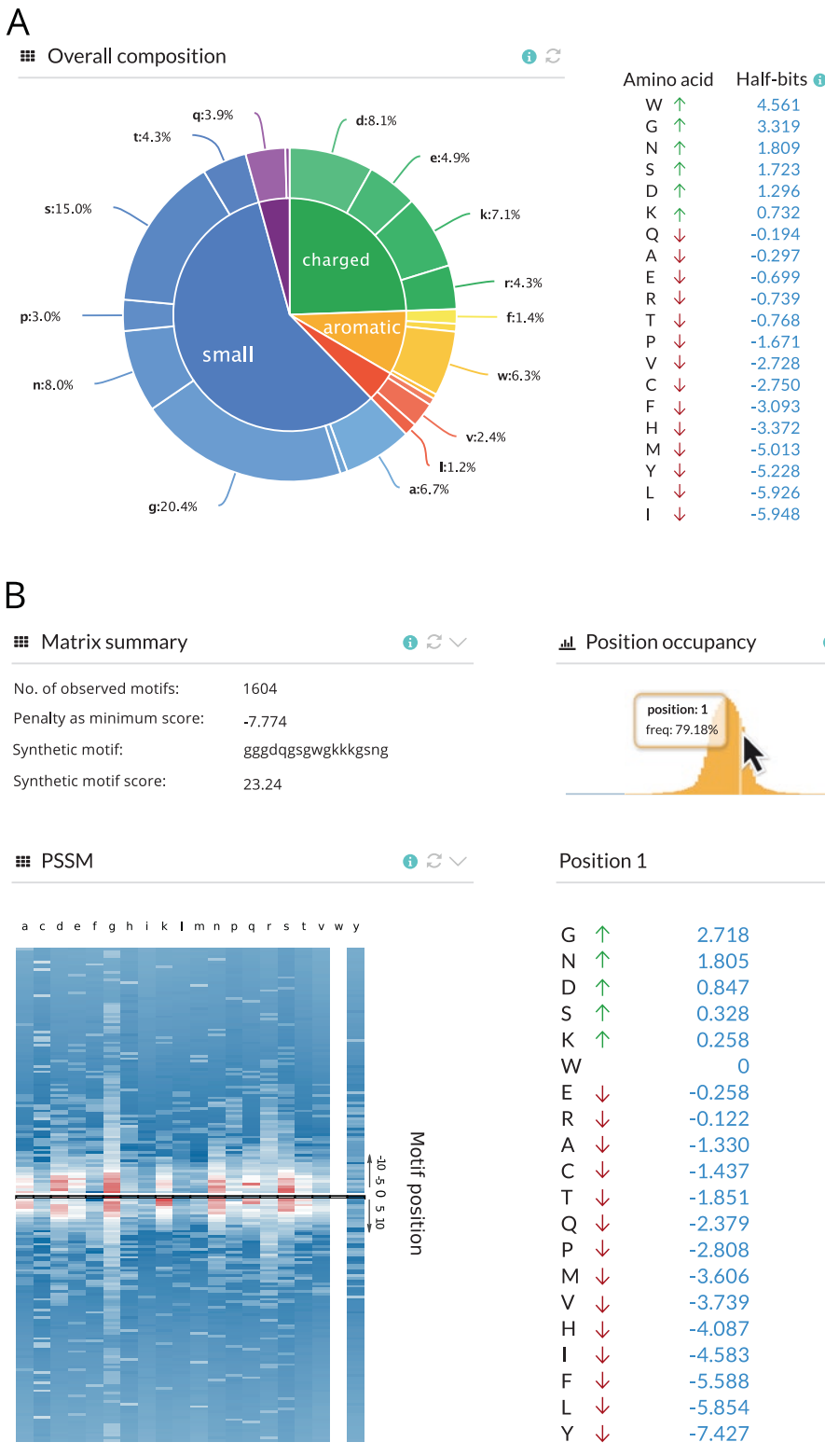


Fig. 2 Analysis of the amino acid composition and length of plant AGO-binding domains using Domain Analyzer. (**a**) Overall amino acid content of WG/GW domains. (**b**) The position-dependent amino acid composition of W-containing motifs

proteins (Fig. 2a). A positive score indicates that the given amino acid is found more often in AGO-binding domain than in other parts of protein (by chance), and a negative score indicates that the amino acid is preferably located within the domain.

As could be expected, the highest-scoring amino acids are tryptophan and glycine, followed by asparagine, serine, glutamic acid, and lysine. The large number of amino acids having negative values (14 negative versus 6 with positive scores) and higher value range for negative scores (5.9 for isoleucine versus 4.8 for tryptophan) suggests a rather stronger negative selection against the presence of some amino acids rather than a positive selection for others (*see* **Note 2**).

3.2.2 Revealing the Trp-Containing Motifs (W-Motifs) of AGO-Binding Domains

The section "PSSM" of Domain Analyzer allows examination of the amino acid context in the close neighborhood to tryptophan residues required for AGO-binding activity. This computational analysis was based on 1600 W-containing sequence motifs that were compared to the overall amino acid composition of W-including subsequences from the protein universe. The result of this comparison is a profile of AGO-binding motifs (PSSM, position-specific scoring matrix) [1]. The PSSM profile shows preferences of amino acids to be specifically present or absent at certain positions of W-containing motifs and can be used for novel AGO-binding domain detection:

1. Switch tab in Domain Analyzer to "PSSM."
2. Mouse over the heatmap which graphically represents PSSM. Red and blue cells show positions in W-containing motifs that are occupied by favorable/unfavorable amino acids, respectively (Fig. 2b).
3. Mouse over the central row of the heatmap (position 0). This position represents the tryptophan residue being a midpoint of the motif (e.g., STG**W**NTS).
4. Move the mouse cursor to position 1 (Fig. 2b), which corresponds to the amino acid preferences in the right surrounding of W (e.g., STGW**N**TS). Red cells in this position indicate a positive tendency toward small, polar, and non-hydrophobic amino acids (G, N, D, S, K).
5. Hover the mouse pointer over different positions of motifs. Note that more distant positions from the central W have lower preferences for specific amino acids; this trend completely vanishes from positions −9 to 9. The PSSM data confirms the experimental studies indicating that functional size of the AGO-binding motif does not seem to exceed the 10-amino-acid boundary [1].
6. Mouse over the "Position occupancy" chart, which gives an estimation of motif size distribution and shows how often each

position is occupied by an amino acid. In most cases (>70%), each tryptophan residue in the functional motif is symmetrically surrounded by 4–6 amino acids on each side. In other words, tryptophan residues in AGO-binding domains are in most cases separated by 8–12 amino acids.

The abovementioned protocol suggests that AGO-binding domains, though significantly variable in sequence and length, are composed of repeated sequence motifs that span from 10 to 20 amino acids, placing the W residue at the center of the hydrophilic and charged surface. This may indicate that W-based, AGO-binding domains are constrained within a narrow subset of possible sequences, which most likely are the result of the biophysical restraints of AGO interactions that are yet to be described (*see* **Note 3**).

3.3 Computational Prediction of AGO-Binding Domains: Single-Sequence Protein Searches and Annotation of New AGO-Binding Proteins

Whub offers three algorithmically different tools for in silico prediction of AGO-binding domains: Agos [17], Wsearch, and iWsearch [1]. All three web applications allow for reliable identification of the WG/GW domain sequence, determination of its boundaries, as well as statistical estimation of the predictions' quality. Agos targets long WG/GW-rich stretches of protein sequences that have biased overall amino acid composition statistically similar to that of experimentally verified AGO-binding domains. In turn, Wsearch and iWsearch look for short and single Trp-containing motifs and determine the probability that any of these motifs will constitute an AGO-binding site. While Wsearch uses PSSM scores for predictions, iWsearch is based on machine learning (RF, random forest) and classifies each W-containing motif as either AGO binding or nonfunctional.

In the early stages of WG/GW research, it was observed that protein sequence of transcription elongation factor SPT6 in *Arabidopsis* is rich in WG/GW repeats [9], but its functional annotation awaits further investigation. In this section we will use Whub's tools to scan the sequence of SPT6 protein to determine whether it may contain AGO-binding sites. The input SPT6 sequence (UniProt accession: A8MS85) can be used either in FASTA or plain sequence format.

3.3.1 Prediction of AGO-Binding Motifs Using Wsearch

1. From Whub's menu "Sequence Search" choose "Wsearch."
2. Paste the protein sequence of SPT6 into the box under "Sequence," or click on the "Example 1" button.
3. From the drop-down menu, select the PSSM matrix specific for AGO-binding proteins in plants ("AGO-Plants").
4. Click on the "Search" button.

Prediction results are presented in four separate panels (Fig. 3). First two include an interactive chart (Fig. 3a) and a query sequence map (Fig. 3b), which provide information about the likelihood

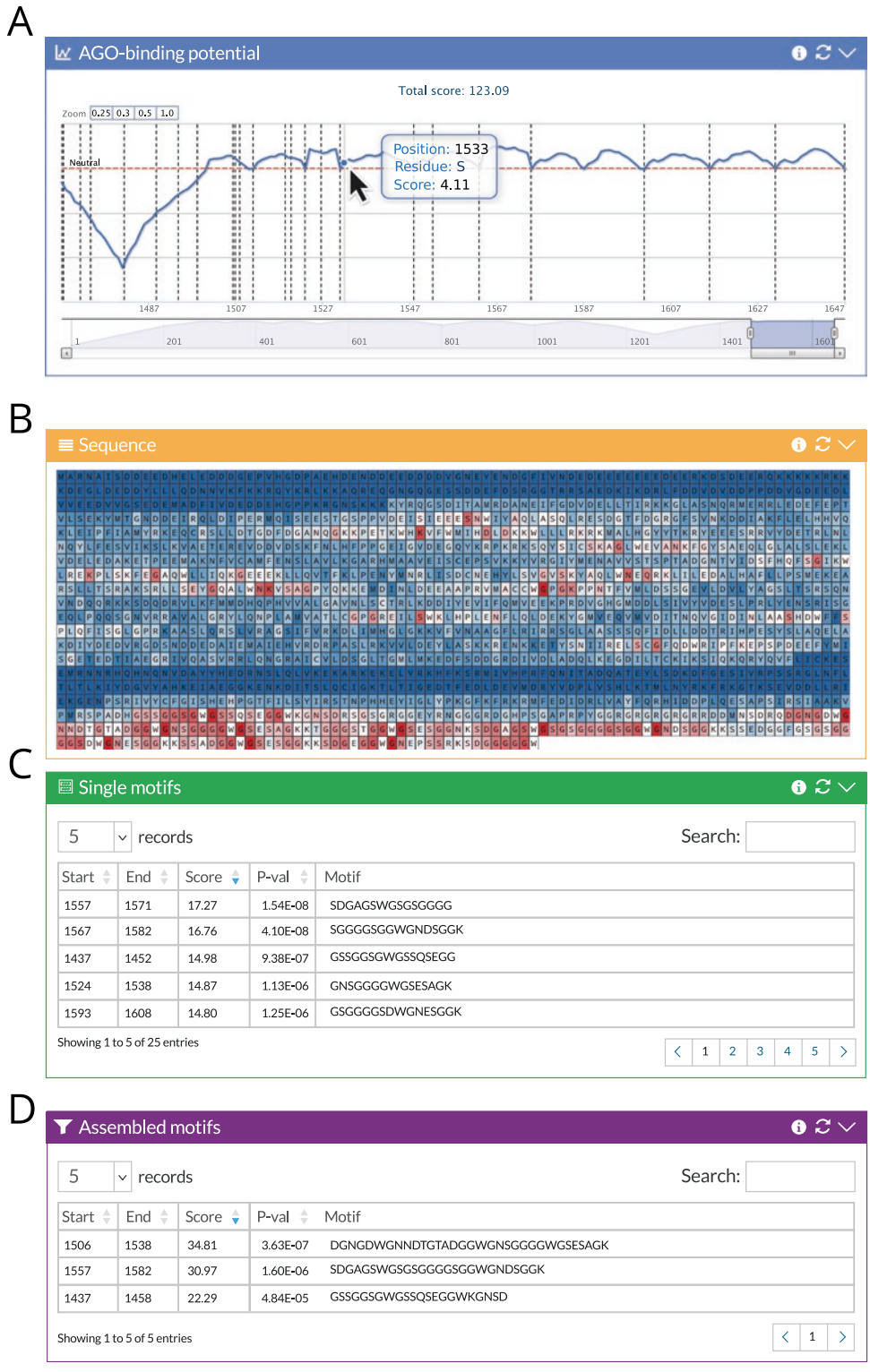


Fig. 3 Wsearch predictions of AGO-binding motifs in *Arabidopsis* SPT6 protein. (**a**) The line graph shows the distribution of scores across the full-length query sequence (tryptophan residues are marked as vertical dashed lines). The chart can be dynamically zoomed in on any individual sequence region. (**b**) A color-coded map of amino acid scores across the full-length query sequence. (**c** and **d**) Tables containing information about all identified single W-motifs and assembled W-motifs (domains), respectively

that a given amino acid will be found in a particular position of an AGO-binding site. The positions of the query sequences that are favorable in AGO-binding activity have positive values in the chart and are colored red on the sequence map. On the other hand, the lowest negative value in the chart (intense blue color) of amino acids in the sequence map indicates that it is unlikely that this position is involved in an interaction with AGO proteins. In our analysis, N-terminus of SPT6 consists of amino acid residues (highlighted in dark blue) unfavorable in AGO-binding sites, while C-terminal region shows many amino acids compatible with confirmed AGO-binding domains (Fig. 3b).

The panel "Single Motifs" shows a sortable and searchable table of single W-containing motifs with information about their localization in the query sequence, score, and p-value (Fig. 3c). For example, the highest-scoring motif (located in the query SPT6 sequence at 1557–1571) achieved a score of 17.27 half bits. The probability of finding (in random sequences) a peptide that would have at least as high a score is less than one in ten million, meaning that the identified motif is a very strong candidate to be involved in the AGO-binding activity.

In the panel "Assembled domains" (Fig. 3d), the overlapping single motifs are assembled into longer domains and statistically reevaluated. In this example, the highest-scoring motif was merged with the second top-rated motif into the domain of 26 amino acid residues. In addition, Wsearch found two other WG/GW domains located in 1506–1538 and 1437–1458 (*see* **Note 4**).

3.3.2 Prediction of AGO-Binding Motifs Using iWsearch

1. From Whub's menu "Sequence Search," choose "iWsearch."
2. Paste the protein sequence of SPT6 or click on the "SPT6" link.
3. Click the "Search" button.

iWsearch provides a list of all single W-containing motifs of 21 amino acids in length with the corresponding probabilities (from 0 to 1) indicating if the motif is involved in the AGO-binding activity (Fig. 4). In this example, iWsearch classified 10 W-motifs from the C-terminus as AGO binding. Similarly to the Wsearch's predictions, the motif with the tryptophan at position 1562 achieved the highest probability (0.999) of being involved in the AGO-binding activity.

3.3.3 Prediction of AGO-Binding Motifs Using Agos

1. From Whub's menu "Sequence Search," choose "Agos."
2. Select "Submit" from the menu.
3. Enter an email address in the appropriate box (this is required in case the search takes too long and times out).
4. Paste the protein sequence of SPT6 in FASTA format into the box under "Sequence" or click on the "Example 1" button and click "Search."

10 records Search:

Trp(W) position	Motif	AGO-binding activity	Probab. of AGO-binding
297	deisieeesnwiyaqlasqlr	no	0.001
372	anqgkkpetkwhkvfwmihdl	no	0.025
377	kpetkwhkvfwmihdldkkwl	no	0.021
386	fwmihdldkkwlllrkrkmal	no	0.004
481	qysicskaglwevankfgysa	no	0.041
587	dsfhqfsgikwlrekplskfe	no	0.006
601	kplskfegaqwlliqkgeeek	no	0.019
647	svgvskyaqlwneqrkliled	no	0.015
692	rllseygqalwnkvsagpyqk	no	0.045
722	eaaprvmaccwgpgkppntfv	no	0.119
879	lcgpgreilswklhplenflq	no	0.011
920	idinlaashdwffsplqfisg	no	0
1075	irelscgfqdwripfkepspd	no	0.024
1433	adhgssggsgwgssqseggwk	no	0.345
1452	gwgssqseggwkgnsdrsgsg	yes	0.81
1510	nsdrqdgngdwgnndtgtadg	yes	0.868
1522	nndtgtadggwgnsggggwgs	yes	0.96
1530	ggwgnsggggwgsesagkktg	yes	0.979
1547	kktgggstggwgsesggnksd	yes	0.99
1562	ggnksdgagswgsgsggggsg	yes	0.999
1574	sgsggggsggwgndsggkkss	yes	0.977
1600	gsgsggggsdwgnesggkkss	yes	0.926
1615	ggkkssadggwgsesggkksd	yes	0.993
1630	ggkksdgeggwgnepssrksd	yes	0.978

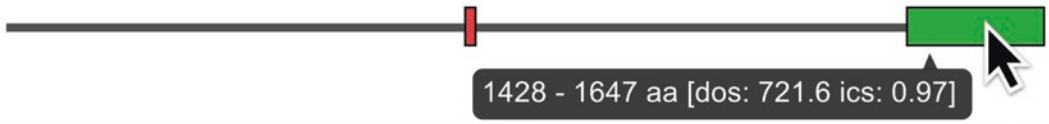
Fig. 4 iWsearch prediction of AGO-binding motifs in *Arabidopsis* SPT6 protein

Agos' output includes a graphical and textual view of compositionally biased sequence regions (Fig. 5). The quality of WG/GW domain predictions is color-coded. Green blocks indicate domains that passed statistical threshold values. In our example, Agos identifies a 220-aa-long C-terminal region comprising 19 WG/GW copies as AGO-binding domains. In the output, the red block of length 14 has a very low-compositional compatibility to the known AGO-binding domains—this is the case of the region located in 719–723 and containing single WG repeat. Similar to Wsearch, moving the cursor over a matching block in the graphical

Results

Data info:

id:	tr|A8MS85|A8MS85_ARATH
description:	Transcription elongation factor SPT6-like protein
length:	1647aa
WG/GW no:	20



WG/GW domains:

no.	start	end	length	motifs	dos	p-value	ics
1	719	732	14	1	-6.84	7.08E-01	23.39
2	1428	1647	220	19	721.6	8.35E-07	0.97

Sequence:

marnaisddeedheledddgepvhgdpaehdenddeeddddvgneyendgfivndedeeeeeeedeerkdsdeer
qkkkkkrkkkdegldeddylllqdnnvkfkkrqykrlkkaqreqgngqgessddefdsrggtrrsaedkikdrlf
ddvdvddppddvgdeedlvveedvvgsedemadfivdeddehgppkrgnskkkkyrqgsditamrdaneifgdvd
elltirkkglasnqrmerrledefeptvlsekymtgnddeirqldipermqiseestgsppvdeisieeesnwiy
aqlasqlresdgtfdgrgfsvnkddiakflelhhvqkleipfiamyrkeqcrslldtgdfdganqgkkpetkwhk
vfwmihdldkkwlllrkrkmalhgyytkryeeesrrvydetrlnlnqylfesvikslkvaeterevddvdskfnl
hfppgeigvdegqykrpkrksqysicskaglwevankfgysaeqlglalsleklvdeledaketpeemaknfvca
mfenslavlkgarhmaaveiscepsvkkyvrgiymenavvstsptadgntvidsfhqfsgikwlrekplskfega
qwlliqkgeeekllqvtfklpenymnrlisdcnehylsvgvskyaqlwneqrkliledalhafllpsmekearsl
ltsraksrllseygqalwnkvsagpyqkkemdinldeeaaprvmaccwgpgkppntfvmldssgevldvlyagsl
tsrsqnvndqqrkksdqdrvlkfmmdhqphvvalgavnlsctrlkddiyevifqmveekprdvghgmddlsivyv
deslprlyensrisgeqlpqqsgnvrravalgrylqnplamvatlcgpgreilswklhplenflqldekygmveq
vmvditnqvgidinlaashdwffsplqfisglgprkaaslqrslvragsifvrkdlimhglgkkvfvnaagflri
rrsglaasssqfidllddtrihpesyslaqelakdiydedvrgdsnddedaiemaiehvrdrpaslrkvvldeyl
askkrenkketysniirelscgfqdwripfkepspdeefymisgetedtiaegrivqasvrrlqngraicvldsg
ltgmlmkedfsddgrdivdladqlkegdiltckiksiqkqryqvflickesemrnnrhqhnqnvdayyhedrnsl
qlvkekarkekelvrkhfksrmivhprfqnitadqateylsdkdfgesivrpssrglnfltltlkiydgvyahke
iaeggkenkditslqcigktltigedtfedldevmdryvdplvshlktmlnyrkfrkgtksevddllriekgenp
srivycfgishehpgtfilsyirstnphheyiglypkgfkfrkrmfedidrlvayfqrhiddplqesapsirsia
akvpmrspadhgssggsgwgssqseggwkgnsdrsgsgrggeyrngggrdghpsgaprpyggrgrgrgrgrrddm
nsdrqdgngdwgnndtgtadggwgnsggggwgsesagkktgggstggwgsesggnksdgagswgsgsggggsggw
gndsggkkssedggfgsgsggggsdwgnesggkkssadggwgsesggkksdgeggwgnepssrksdgggggw

Fig. 5 Agos predictions of AGO-binding domains in *Arabidopsis* SPT6 protein

protein view will highlight its position in the full-length protein sequence as well as the corresponding row in the "WG/GW domains" table.

Taken together, all three prediction tools annotated SPT6 as AGO-binding protein, which constitutes this protein as a promising candidate for further experimental verification.

3.4 Computational Prediction of AGO-Binding Domains: Large-Scale Protein Searches

In order to screen large data sets of protein sequences for identification of potential AGO-binding sites, the most favorable solution is to run the prediction programs on local computers. It is a less user-friendly task than the web-based searches, requiring a download of the Wsearch software that needs a working Python 3 installation. Nevertheless, Wsearch does not require installation or any additional library and can be used on any operating system.

1. Download and install the latest version of Python 3 (https://www.python.org/downloads/).
2. Download Wsearch from Whub ("Download" > "Software"). As input, Wsearch requires a FASTA file with protein sequences and a PSSM file specific for AGO-binding activity (by default, Wsearch provides PSSM files for plant, animal, and eukaryotic AGO-binding sites). To see how Wsearch works, on the command line, type:

```
python3 wsearch.py -help
```

3. In order to scan sample input protein sequences using PSSM matrix from plants, on the command–line enter:

```
python3 wsearch.py -query example/proteins.fa -pssm
pssm/ago-plants.pssm
```

4. Wsearch generates output in two files, which contain textual information about the identified single and assembled motifs, separately (*see* **Note 5**). The output is presented in the following format:

```
sequenceID        start end   score    subsequence
XP_004308760.1    4     16    7.469    NNNSTWGVAAAEN
XP_004308760.1    18    26    11.705   DQGTGWGKS
XP_004308760.1    24    36    10.128   GKSESWGAKVGGD
XP_004308760.1    35    47    3.574    GDSNLSDTWQKAS
XP_004308760.1    47    59    8.301    SEPASSSWGVAAA
XP_004308760.1    70    82    11.395   GKSDAWGAKVGGD
XP_004308760.1    81    93    6.295    GDSTSSDTWQKAS
XP_004308760.1    93    107   8.719    SEPASSSWGVAAAAD
XP_004308760.1    110   120   9.876    DQGSVWGGNDS
XP_004308760.1    129   143   12.071   SGDRGSAWSKPAGGS
```

5. It is possible to create own PSSM profiles based on a custom collection of protein sequences. For example, to create a PSSM matrix using FASTA sequences in the *test.fa* file and background proteins from UniProt/Swiss-Prot, run the command:

```
python3 makepssm.py –in test.fa –db uniprot_sprot_
plant.fa –out ago.animals.mat
```

3.5 Game-like Design of In Silico Mutagenesis Studies on AGO-Binding Domain Properties

Whub provides an interactive puzzle-like framework enabling the user to in silico design synthetic AGO-binding domains or to modify existing ones with a single residue resolution and real-time, color-encoded prediction of binding quality. The main purpose of this application is a simulation of mutagenesis experiments and hypothesis testing. In this section, we will computationally reproduce the wet lab experiment performed by Szabó and colleagues [4], where the nonfunctional sequence of the P1 protein from *Sweet potato feathery mottle virus* (SPFMV) was transformed into a functional silencing suppressor with AGO-binding capacity by introducing two WG/GW motifs.

1. From Whub's menu click "Design your domain."
2. In the input text, paste the nonfunctional viral sequence of P1 protein,

   `DVLDGHKCDSCGHRYIRRDDNIADSMNDIARALGGYDAYYAS`. Hit "Start."
3. The input sequence is represented by a string of blocks (amino acids). From the drop-down menu, select PSSM matrix for AGO-binding function in plant viruses (*AGO-PlantVir*).
4. Since the input sequence lacks tryptophan residues, the blocks are gray, and the score of AGO-binding activity cannot be calculated. Mouse over the very first histidine's block in position 6, and substitute the residue with tryptophan. As soon as the sequence is modified, the score and block colors are automatically computed and displayed. In this case, the sequence obtained the score of −89.34, meaning that the sequence in the current form lacks any probable AGO-binding activity. The intensity of red and blue colors reflects PSSM scoring values. Red indicates that a given amino acid is favorable at given position, for example, residues surrounding the introduced tryptophan. Blue blocks indicate that residues at given positions are incompatible with AGO-binding activity.
5. Simulate the second mutation by substituting tyrosine in position 36 with tryptophan. This sequence achieves score of 22.33, and it was shown by Szabó et al. (2012) to counteract with active plant RNA-induced silencing complex (RISC) by binding the host AGO proteins.
6. Find what is the score of this sequence according to the PSSM matrix specified for eukaryotes. Modify the existing sequence

to find a configuration that maximizes score and thereby the red color intensity of all blocks. In addition to substituting amino acids, you can move residues via drag-and-drops, if necessary pushing its neighbors (emulating transposition events) as well as insert and delete amino acid blocks. The authors of this chapter ended up with the sequence of 51.37 scores.

4 Notes

1. Browse various AGO-binding protein records for different organisms (e.g., human, *Arabidopsis*, yeast). Note that these proteins have different domain organization, and AGO-binding domains show high divergence in sequence length and number of WG/GW repeats. Look at human TNRC6 (aka GW182) proteins, which are the best characterized AGO partners in animal cells and have been intensively investigated because of their primary function in miRNA-dependent post-transcriptional silencing.
2. WG/GW domains are significantly depleted in bulk hydrophobic (Ile, Leu, and Val) and aromatic amino acid residues (Tyr and Phe), which would normally form the hydrophobic core of folded globular proteins. The domains possess very low content of Cys, which is known to have a significant contribution to the protein conformation stability via the disulfide bond formation or being involved in coordination of different prosthetic groups.
3. Look at "Domain Analyzer" of AGO-binding domains in animals and eukaryotes. Although the sequence of this domain shows no sequence conservation, the biased amino acid composition is conserved in different kingdoms.
4. When scanning *Arabidopsis* SPT6 sequence on the Wsearch website, try changing the PSSM matrix to animal specific. Note that such search also identifies potential AGO-binding motifs and domains, meaning that amino acid context of Trp-containing motifs is conserved between plant and animal kingdoms.
5. Although AGO-binding proteins have not been found in bacteria, it is possible that some pathogenic microorganisms use WG/GW mimicry to disarm host's defense system. From UniProt, download proteins of *Propionibacterium acnes*, which is linked to the skin condition of acne. Use Wsearch to scan the proteins for WG/GW domains. The screening should take less than a second. Among results, single-stranded DNA-binding protein (UniProt Accession: F9NUV8) contains 50-amino acid-long C-terminal fragment that may have AGO-binding activity (p-value = 3.01E–09).

References

1. Zielezinski A, Karlowski WM (2015) Integrative data analysis indicates an intrinsic disordered domain character of Argonaute-binding motifs. Bioinformatics 31(3):332–339. doi:10.1093/bioinformatics/btu666
2. Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, Voinnet O (2010) Argonaute quenching and global changes in dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev 24(9):904–915. doi:10.1101/gad.1908710
3. Giner A, Lakatos L, Garcia-Chapa M, Lopez-Moya JJ, Burgyan J (2010) Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs. PLoS Pathog 6(7):e1000996. doi:10.1371/journal.ppat.1000996
4. Szabó EZ, Manczinger M, Goblos A, Kemeny L, Lakatos L (2012) Switching on RNA silencing suppressor activity by restoring argonaute binding to a viral protein. J Virol 86(15):8324–8327. doi:10.1128/JVI.00627-12
5. de Ronde D, Pasquier A, Ying S, Butterbach P, Lohuis D, Kormelink R (2014) Analysis of tomato spotted wilt virus NSs protein indicates the importance of the N-terminal domain for avirulence and RNA silencing suppression. Mol Plant Pathol 15(2):185–195. doi:10.1111/mpp.12082
6. Aqil M, Naqvi AR, Bano AS, Jameel S (2013) The HIV-1 Nef protein binds argonaute-2 and functions as a viral suppressor of RNA interference. PLoS One 8(9):e74472. doi:10.1371/journal.pone.0074472
7. Karlowski WM, Zielezinski A, Carrere J, Pontier D, Lagrange T, Cooke R (2010) Genome-wide computational identification of WG/GW Argonaute-binding proteins in Arabidopsis. Nucleic Acids Res 38(13):4231–4245. doi:10.1093/nar/gkq162
8. Till S, Lejeune E, Thermann R, Bortfeld M, Hothorn M, Enderle D, Heinrich C, Hentze MW, Ladurner AG (2007) A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. Nat Struct Mol Biol 14(10):897–903. doi:10.1038/nsmb1302
9. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T (2007) Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev 21(20):2539–2544. doi:10.1101/gad.451207
10. He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu HL, Wang CS, Jin H, Zhu JK (2009) An effector of RNA-directed DNA methylation in arabidopsis is an ARGONAUTE 4- and RNA-binding protein. Cell 137(3):498–508. doi:10.1016/j.cell.2009.04.028
11. Zielezinski A, Karlowski WM (2015) Early origin and adaptive evolution of the GW182 protein family, the key component of RNA silencing in animals. RNA Biol 12(7):761–770. doi:10.1080/15476286.2015.1051302
12. Gibbings D, Leblanc P, Jay F, Pontier D, Michel F, Schwab Y, Alais S, Lagrange T, Voinnet O (2012) Human prion protein binds Argonaute and promotes accumulation of microRNA effector complexes. Nat Struct Mol Biol 19(5):517–524., S511. doi:10.1038/nsmb.2273
13. UniProt C (2015) UniProt: a hub for protein information. Nucleic Acids Res 43(Database issue):D204–D212. doi:10.1093/nar/gku989
14. Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI reference sequences (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res 40(Database issue):D130–D135. doi:10.1093/nar/gkr1079
15. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44(D1):D279–D285. doi:10.1093/nar/gkv1344
16. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43(Database issue):D213–D221. doi:10.1093/nar/gku1243
17. Zielezinski A, Karlowski WM (2011) Agos—a universal web tool for GW Argonaute-binding domain prediction. Bioinformatics 27(9):1318–1319. doi:10.1093/bioinformatics/btr128

Chapter 19

Identification and Evolutionary Characterization of ARGONAUTE-Binding Platforms

Joshua T. Trujillo and Rebecca A. Mosher

Abstract

ARGONAUTE (AGO) proteins are eukaryotic RNA silencing effectors that interact with their binding partners via short peptide motifs known as AGO hooks. AGO hooks tend to cluster in one region of the protein to create an AGO-binding platform. In addition to the presence of AGO hooks, AGO-binding platforms are intrinsically disordered, contain tandem repeat arrays, and have weak sequence conservation even between close relatives. These characteristics make it difficult to identify and perform evolutionary analysis of these regions. Because of their weak sequence conservation, only a few AGO-binding platforms are characterized, and the evolution of these regions is only poorly understood. In this chapter we describe modules developed for computational identification and evolutionary analysis of AGO-binding platforms, with particular emphasis on understanding evolution of the tandem repeat arrays.

Key words ARGONAUTE, AGO hook, Tandem repeat

1 Introduction

ARGONAUTE (AGO) proteins bind small RNAs and play a central role in regulating gene expression in most eukaryotes as a component of the RNA-induced silencing complex [1]. AGOs interact with other proteins through AGO hooks, short peptide motifs consisting of tryptophan and glycine(s) in one of three contexts (WG, GW, GWG). AGO hooks are found in diverse RNA silencing proteins, including NRPE1, SPT5L, and Tas3, which mediate transcriptional gene silencing [2–4]; GW182, which promotes posttranscriptional gene silencing [5]; and Wag1p and CnjBp, which are required for genome elimination [6]. AGO hooks also occur in viral suppressors of RNA silencing, where they inhibit AGO function [7, 8]. The number of AGO hooks within these proteins varies, but they often cluster within a particular region of the protein, referred to as an AGO-binding platform.

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_19,

Although ARGONAUTE-interacting proteins might possess well-conserved and highly structured domains, such as the polymerase domains in NRPE1, AGO-binding platforms generally have poor sequence conservation between species and are intrinsically disordered [2, 6, 9–12].

In addition to the presence of AGO hooks, many AGO-binding platforms also contain repetitive sequence arrays [3, 11]. These repetitive sequence arrays vary in size, repeat unit, and number of repeats between even closely related species, contributing to the lack of sequence conservation between AGO-binding platforms [12]. The repetitive sequence units tend to contain at least one AGO hook and are enriched in disorder-promoting amino acids that contribute to the intrinsic disorder of the region [12].

The discovery of AGO-binding platforms as functionally important domains for mediating interaction with ARGONAUTE proteins is a relatively recent discovery [2, 3], but with only a few well-characterized examples, our understanding of these regions is just beginning. Identification and characterization of additional AGO-binding platforms within the wealth of available genome sequence will uncover additional aspects of ARGONAUTE function and evolution.

In this chapter, we outline modules that are designed to computationally identify and analyze AGO-binding platforms, with particular emphasis on understanding the evolution of the tandem repeat arrays.

2 Materials

The pipeline outlined here primarily utilizes publicly available tools and software, although there are additional commercial programs and software that can assist this process. Most of the public tools are available in a web-based platform, and many are additionally run in the command line after download.

2.1 Publicly Available Bioinformatic Tools and Software

1. BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) identifies and aligns homologous sequences [13].
2. MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/) and/or ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) are algorithms for alignment of multiple nucleotide and amino acid sequences [14, 15].
3. IUPred (http://iupred.enzim.hu/) identifies intrinsically disordered and unstructured regions based on the pairwise interaction energy of the amino acid sequence [16, 17].
4. RADAR (https://www.ebi.ac.uk/Tools/pfa/radar/) uses a segmentation-based approach to identify both simple and complex repeats from amino acid sequence [18].

5. T-REKS (http://bioinfo.montp.cnrs.fr/?r=t–reks/) is a k-means algorithm for identification of highly similar tandem repeats in amino acid sequences [19].
6. RAxML (http://embnet.vital–it.ch/raxml–bb/index.php) performs maximum likelihood-based phylogenetic inference [20].
7. MEGA (http://www.megasoftware.net/) provides a suite of tools for the analysis of DNA and protein sequences from an evolutionary perspective [21].
8. DnaSP (http://www.ub.edu/dnasp) is a software package for the analysis of nucleotide polymorphism and evolutionary selection forces [22, 23].

2.2 Commercial Tools and Software

Although the following modules do not require purchasing computational tools or software, molecular editing and viewing programs such as Geneious, Vector NTI, MacVector, or SnapGene are very useful for handling and analyzing nucleotide and amino acid sequences. For our analyses, we prefer and used Geneious (http://www.geneious.com/).

3 Methods

The workflow for characterizing AGO-binding platforms consists of three modules: (1) identifying AGO-binding platforms, (2) phylogenetic analysis of repetitive DNA, and (3) assessing selective evolutionary pressures/processes. These modules are described in the following subsections, but it is important to note that not all of the modules are necessary to understand an AGO-binding platform nor appropriate for comparison of all AGO-binding platforms.

3.1 Identifying AGO-Binding Platforms

Mapping AGO hooks is the first step in determining if a protein possesses an AGO-binding platform. AGO hooks can be identified with various computational programs or even a simple character search in a text viewer. However, we recommend using a molecular editing and viewing program (e.g., Geneious, UGENE) for quick annotation and analysis of AGO hooks. Begin by searching for and annotating glycine (G) and tryptophan (W) resides in the three different AGO hook contexts (GW, WG, GWG). The presence of these amino acids adjacent to each other does not necessarily indicate AGO interaction; however, enrichment of AGO hooks suggests the presence of an AGO-binding platform. Draft genome sequence, poor assembly, or incorrect annotation can lead to difficulties identifying and characterizing AGO-binding platforms (*see* **Note 1**). Verification of putative AGO-binding platforms through Sanger sequencing can improve confidence in later analyses. Comparison of orthologous sequences from close relatives can also assist annotation and analysis.

AGO-binding platforms are typically enriched with amino acids that promote disorder [12, 24]. To predict intrinsically disordered regions, submit your protein sequence to IUPred [16]. IUPred reports the disorder tendency of each residue in the peptide, which can then be used to identify regions of elevated disorder. Overlap between AGO hooks and intrinsically disordered regions is additional evidence for the presence of an AGO-binding platform. Using a sliding window approach, the average disorder along the protein sequence can be plotted to visualize regions of elevated disorder and better define the platform boundaries.

AGO-binding platforms also tend to contain at least one tandem repeat array [3, 11, 12]. Depending upon the level of conservation between the repeat units and the complexity of the repeat array, tandem repeats might be difficult to identify with a single approach; we therefore recommend using a combination of tools. The first indication that a protein contains a repeat array can be seen in a self-by-self comparison in a dot plot. Dot plots can be performed in many DNA management programs (e.g., Geneious or UGENE) or in R. In a self-by-self dot plot, repeats appear diagonal lines separated from the primary diagonal (Fig. 1). The extent of this signal is an indication of the size and conservation of the repeat array. Although a dot plot provides a quick assessment of a repeat array, it does not directly identify the repeat unit. One can manually search the sequence using the dot plot as a guide or take advantage of repeat finder programs to uncover the repeat unit. T-REKS [19] is adept at finding highly conserved and well-structured repeats, while RADAR [18] can identify larger and more complex repeat arrays. Neither program is perfect, and we recommend further assessing the repeat array manually. The "align two sequences" function in BLAST can be used to aid manual searches for repeat sequences (*see* **Note 2**). In our research, we have applied all of these methods to identify repetitive sequences.

3.2 Phylogenetic Assessment of Repeat Arrays

A key feature of repeat arrays is that they are labile and tend to expand and contract [25]. Phylogenetic analysis of repeats from a single gene elucidates the relationship among the repeats and might provide evidence for duplication of a unit composed of several repeats. If sequences from closely related species are available, changes to the repeat array can be inferred by the phylogenetic relationship among repeat units. Because expansion occurs through duplication of existing repeats, these duplicates repeat units will occur next to each other in a phylogenetic tree.

To begin, each repeat unit should be extracted from the AGO-binding array and treated as an individual sequence or "gene." Individual repeat units should be initially named based on their linear order in the gene. Although phylogenetic analysis is possible with protein sequences, it is preferable to use nucleotide sequences

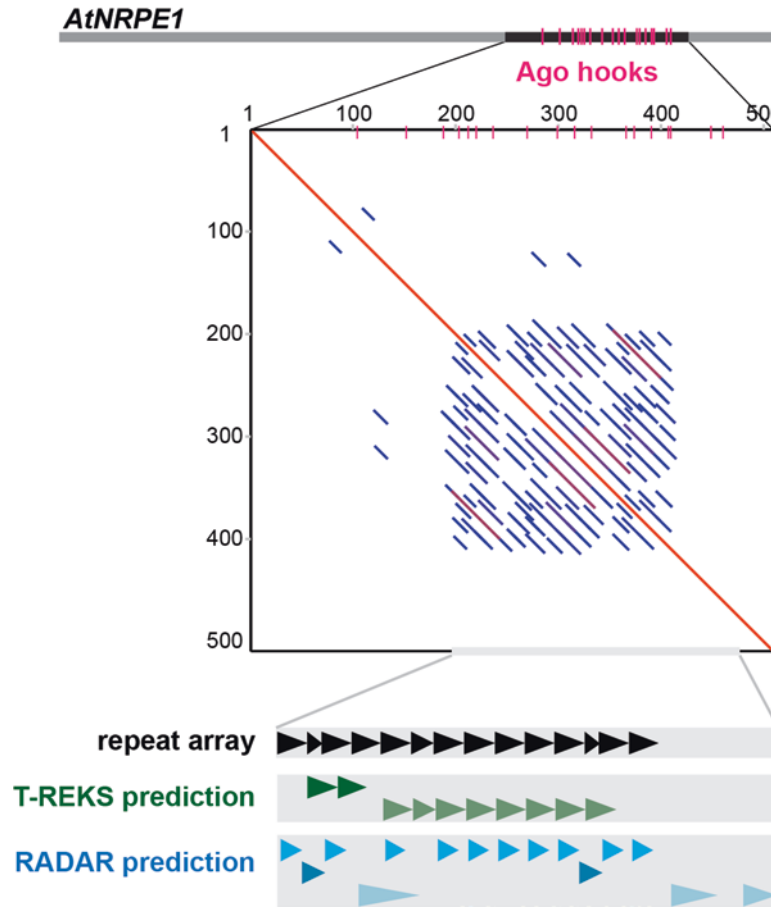


Fig. 1 Identifying repeat arrays in AGO-binding platforms. (*top*) AtNRPE1 contains an approximately 500-amino acid region that is rich in AGO hooks (GW, WG, or GWG peptides, *pink tick marks*). When aligned in a self-by-self dot plot (*middle*), repetitive sequence can be identified by signal away from the red diagonal line. (*bottom*) Repeat prediction programs, including T-REKS and RADAR, identify parts of the repeat array (*colored arrowheads*), but manual curation is necessary to define the entire array. Differently colored arrowheads designate different repeat sequences; the manually curated repeat array is in black

to access the additional information contained in codon choice. Repeat units are then aligned with MUSCLE [14], ClustalW [15], or another alignment algorithm, and the resulting alignment can be passed to a maximum likelihood-based phylogenetic analysis, such as RAxML, to determine the evolutionary relationship among repeat units (*see* **Note 3**). Geneious also performs alignments and contains a plug-in for maximum likelihood phylogenetic analysis using PhyML. Several species can be analyzed simultaneously; however, interpreting the phylogenetic relationship can be very difficult with more than three or four related species.

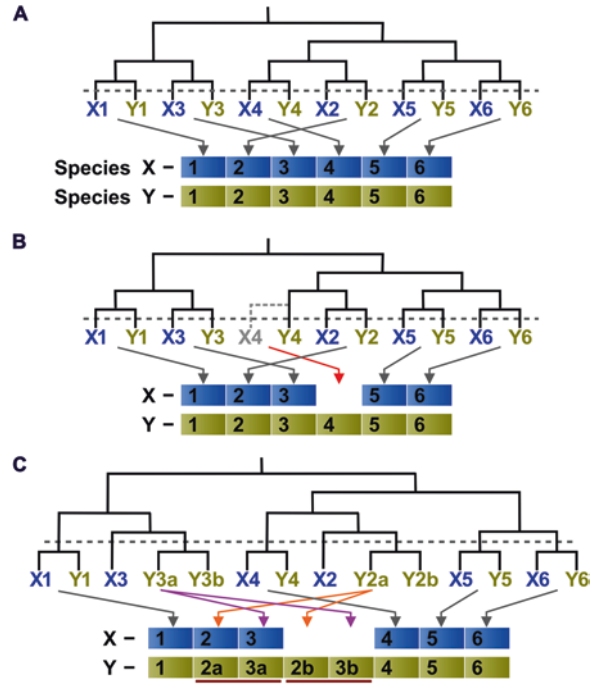


Fig. 2 Phylogenetic analysis of repeat units from related species reveals expansions and contractions within the repeat array. Sequences of individual repeat units from two homologous repeat arrays (X and Y) were aligned and used to create a maximum likelihood phylogenetic tree. The inferred time of speciation is depicted with a dashed line. Evolutionary interpretation of repeats arrays is facilitated by comparing the phylogenetic and linear relationships between repeat units (*arrows*). There can be multiple interpretations for any single set of data. We represent the most parsimonious explanation here; however, additional comparisons, especially with ancestral species, will strengthen the analysis. (**a**) Repeat arrays X and Y are colinear: each repeat unit is most similar to a single repeat unit in the homologous array, a pattern known as a "cherry." All of the cherries in this example contain repeats in the same linear position, indicating there have been no changes to the array since speciation. (**b**) Repeat array X has contracted since divergence from repeat array Y. This can be inferred because repeat Y4 is not part of a cherry, but is instead sister to a cherry. The position in the tree where the missing repeat unit should be found is marked with a dotted line and red arrow. (**c**) Repeat array Y has duplicated two units of the repeat array in a single event. This duplication is inferred because repeat units X2 and X3 are not part of cherries, but instead are sister to a pair of repeats from the homologous array. The relevant repeat units are marked with orange and purple arrows

If there have been no expansions or contractions of the repeat array since the divergence of two species, the repeats will form pairs in the phylogenetic tree, a pattern known as "cherries" (Fig. 2). Rearrangements of the repeat array are detected by changes to an all-cherries pattern (Fig. 2). It is not possible to determine if a

single event is an expansion in species A or a contraction in species B without comparison with other species to determine what the ancestral state of the array was.

3.3 Assessing Selection

Evolutionary selective pressures acting on the AGO-binding platform can be determined if there are at least two orthologous sequences that have not experienced expansion or contractions of the repeat array since divergence (the arrays are colinear). This ensures that any variation between sequences is due to substitution rather than duplication/loss of repeat units. The orthologous sequences might be from closely related species or from multiple individuals of the same species.

Tajima's D is a statistical test to determine ongoing selection within a population. It uses polymorphism data to determine if a DNA or gene sequence is evolving neutrally or experiencing evolutionary selective forces. The difference between two measures of genetic diversity, average pairwise difference in SNPs among individuals and the number of segregating sites in the population, is used to determine if the observed amount of polymorphism varies from the expectation under neutral evolution. Sequences that significantly deviate from the expectation suggest that there is directional selection acting on the sequence. As few as three individuals can be used in Tajima's D analysis; however, including more individuals will result in a more robust and accurate analysis. DnaSP can be used to quickly calculate Tajima's D and statistical significance and produce sliding window graphical outputs. Accurate interpretation of the results requires the D value to be statistically significant; however, most genes are under some level of purifying selection, and therefore the normal distribution underlying assumptions of significance might not be appropriate, complicating interpretation of the D value. One approach to account for this difficulty is to measure Tajima's D at neighboring genes to assess whether a significant value could be obtained through linkage to a highly selected gene.

Historic evolutionary selection is frequently determined by calculating the ratio between the rate of synonymous (dS or Ks) versus nonsynonymous (dN or Ka) amino acid substitutions (dN/dS or ω). Synonymous substitution rate provides a measure of the background mutation rate within the gene and serves as a reference to interpret the rate of nonsynonymous and potentially meaningful substitutions. Low ω values indicate purifying selection, while high ω values suggest positive selection. DnaSP is capable of computing ω between orthologous sequences as well as pairwise comparisons within a population of sequences. A sliding window approach can be used to plot the average ω along the alignment to visualize regions with differential evolutionary rates.

Finally, comparison of the rate of evolution both within (pN/pS) and between (dN/dS) two species provides information regarding

evolutionary selection acting on a gene since the speciation of these relatives. Known as the McDonald–Kreitman test, within and between species evolution rates are compared in order to determine evidence of adaptive evolution in either species. Rates of evolution should be similar in the absence of selection, but significance differences in rates suggest that selection has occurred since speciation. DnaSP easily allows for comparison of sets of population-level data and calculates McDonald–Kreitman values. Similar to Tajima's D, only a few individuals are needed within each population, but the analysis will improve with the inclusion of more population data.

3.4 Conclusion

This chapter has described a process to identify and characterize AGO-binding platforms, including evolutionary analysis of these regions. As the key specification determinant for RNA silencing pathways, AGO proteins are important regulators of gene expression in many eukaryotes. Identifying AGO-binding platforms and understanding their diversity and evolution will provide additional understanding of the mechanism of RNA silencing.

4 Notes

1. It is important to carefully review annotated sequences as the weak conservation and repetitive nature of AGO platforms make genome assembly and annotation difficult. In particular, the rapid evolution of AGO-binding domains has negative consequences for ortholog identification and homology-driven annotation, while the presence of a highly conserved repeat array can lead to improper genome assembly following short-read sequencing. One useful strategy is to confirm that exon number and structure are the same as the reference species: AGO-binding platforms tend to occur within a single exon, but occasionally annotated genes contain false introns due to incomplete or incorrect assembly of the repetitive regions. FGENESH, which uses a reference protein sequence to predict coding sequence from related genomes, is also useful to improve annotation of related AGO-binding proteins, but manual curation may be needed.
2. To assess repetitive sequence using BlastAlign, split the amino acid sequence into smaller pieces before alignment. If the whole repeat array is entered as both sequences (as with a self-by-self dot plot), BLAST will only report the perfect match to itself. Instead, put up to half of the repeat array as the query and other portions of the array as the subject. Once a single repeat unit has been defined, it is helpful to use this single repeat as the query against the rest of the array.

3. It is important to perform translation alignment of nucleotide sequences to maintain proper codon alignment at each amino acid position. It is advisable to include and report bootstrap support when building the maximum likelihood tree to determine. High bootstrap values (>90) between repeat units support that those repeats arose from a duplication (single-species analysis) or are homologous units between species. Low bootstrap values (<90) might reflect legitimate concerns, such as undetected changes due to duplication then loss of repeat units or partial duplications of repeats (e.g., the second half of repeat two and the first half of repeat three). Weak support for a node(s) can also arise when the repeat unit is short, and therefore the alignment matrix has too few columns for robust bootstrapping.

References

1. Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14(7):447–459. doi:10.1038/nrg3462
2. Till S, Lejeune E, Thermann R, Bortfeld M, Hothorn M, Enderle D, Heinrich C, Hentze MW, Ladurner AG (2007) A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. Nat Struct Mol Biol 14(10):897–903. doi:10.1038/nsmb1302
3. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T (2007) Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev 21(20):2539–2544. doi:10.1101/gad.451207
4. Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, Lagrange T (2009) RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. EMBO Rep 10(6):649–654. doi:10.1038/embor.2009.31
5. Takimoto K, Wakiyama M, Yokoyama S (2009) Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. RNA 15(6):1078–1089. doi:10.1261/rna.1363109
6. Bednenko J, Noto T, DeSouza LV, Siu KW, Pearlman RE, Mochizuki K, Gorovsky MA (2009) Two GW repeat proteins interact with Tetrahymena Thermophila argonaute and promote genome rearrangement. Mol Cell Biol 29(18):5020–5030. doi:10.1128/MCB.00076-09
7. Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, Voinnet O (2010) Argonaute quenching and global changes in dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev 24(9):904–915. doi:10.1101/gad.1908710
8. Giner A, Lakatos L, Garcia-Chapa M, Lopez-Moya JJ, Burgyan J (2010) Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs. PLoS Pathog 6(7):e1000996. doi:10.1371/journal.ppat.1000996
9. Partridge JF, DeBeauchamp JL, Kosinski AM, Ulrich DL, Hadler MJ, Noffsinger VJ (2007) Functional separation of the requirements for establishment and maintenance of centromeric heterochromatin. Mol Cell 26(4):593–602. doi:10.1016/j.molcel.2007.05.004
10. Lian SL, Li S, Abadal GX, Pauley BA, Fritzler MJ, Chan EK (2009) The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. RNA 15(5):804–813. doi:10.1261/rna.1229409
11. Huang Y, Kendall T, Forsythe ES, Dorantes-Acosta A, Li S, Caballero-Perez J, Chen X, Arteaga-Vazquez M, Beilstein MA, Mosher RA (2015) Ancient origin and recent innovations of RNA polymerase IV and V. Mol Biol Evol 32(7):1788–1799. doi:10.1093/molbev/msv060
12. Trujillo JT, Beilstein MA, Mosher RA (2016) The Argonaute-binding platform of NRPE1 evolves through modulation of intrinsically

disordered repeats. New Phytol. doi:10.1111/nph.14089

13. 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
14. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797. doi:10.1093/nar/gkh340
15. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
16. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21(16):3433–3434. doi:10.1093/bioinformatics/bti541
17. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J Mol Biol 347(4):827–839. doi:10.1016/j.jmb.2005.01.071
18. Heger A, Holm L (2000) Rapid automatic detection and alignment of repeats in protein sequences. Proteins 41(2):224–237
19. Jorda J, Kajava AV (2009) T-REKS: identification of tandem REpeats in sequences with a K-meanS based algorithm. Bioinformatics 25(20):2632–2638. doi:10.1093/bioinformatics/btp482
20. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313. doi:10.1093/bioinformatics/btu033
21. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874. doi:10.1093/molbev/msw054
22. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452. doi:10.1093/bioinformatics/btp187
23. Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. Methods Mol Biol 537:337–350. doi:10.1007/978-1-59745-251-9_17
24. Karlowski WM, Zielezinski A, Carrere J, Pontier D, Lagrange T, Cooke R (2010) Genome-wide computational identification of WG/GW Argonaute-binding proteins in Arabidopsis. Nucleic Acids Res 38(13):4231–4245. doi:10.1093/nar/gkq162
25. Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet 44:445–477. doi:10.1146/annurev-genet-072610-155046

Chapter 20

Phylogenetic and Evolutionary Analysis of Plant ARGONAUTES

Ravi K. Singh and Shree P. Pandey

Abstract

Comparative sequence analysis is widely used for the reconstruction of phylogeny and for understanding the evolutionary history of gene families. Here, we describe the methodologies to reconstruct the phylogenetic and evolutionary history of a gene family across genomes with a focus on the ARGONAUTE (AGO) family of proteins in plants. The method described here may easily be adapted for studying molecular evolution of a wide variety of gene families. We enlist methods as well as parameters for the collection of molecular data (nucleic acids and peptides), preparation of datasets, and selection of evolutionary models and various methods for the phylogenetic and evolutionary analysis, such as maximum likelihood and Bayesian inference.

Key words ARGONAUTE (AGO), Phylogeny, Molecular evolution, Maximum likelihood, Bayesian inference, Neighbor joining, Plants

1 Introduction

Regulatory small RNAs (smRNAs) form an important layer in modulating gene expression via the process of RNA interference (RNAi) or posttranscriptional gene silencing (PTGS) [1–3]. During this process, smRNAs are loaded on to the ARGONAUTE (AGO) family of proteins to form the RNA-induced silencing complex (RISC). Thus, the AGOs act as effectors to either slice the target mRNA or inhibit its translation [3–5]. smRNAs regulate a variety of developmental, physiological, and reproductive processes in plants in an AGO-dependent manner. smRNAs have been isolated from lineages of almost all taxonomic group of organism, suggesting that this mode of gene regulation is widespread in nature [6, 7]. Further, the presence of AGO homologs in almost all group of "green lineages" indicates an ancient origin for the pathway that modulates biogenesis and action of smRNAs in the plants [7, 8]. As AGOs are the central component of the RISC

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7_20,

machinery, they offer a perfect model to study the molecular evolution of smRNA pathways [8]. The number of AGO genes present in genomes is highly variable, for instance, *Chlamydomonas reinhardtii* harbors four genes, whereas 18 are present in *Oryza sativa* [8, 9]. It has been suggested that the AGO protein family has evolved through multiple but variable number of duplication and loss events among the green lineages [8]. The dynamic evolution of AGO family leads to a varied number of homologs among plant genomes [8, 9]. The expansion of plant AGO family may also have proceeded with the evolutionary diversification in the sequences among different AGOs [8]. These diversities may have resulted into the functional specialization of plant AGOs [8, 10]. For instance, AGO1 and AGO10 preferentially bind to smRNAs of 21-nt length with a 5′-uridine, whereas AGO4 and AGO9 bind to 24-nt endogenous smRNAs having 5′-cytosine [8, 10, 11].

The evolution of AGO family in green lineages has been a complex molecular phenomena [8, 12]. It has always been challenging to correctly unfold the evolutionary history of a gene family across different domains of life [13–15]. The use of high-throughput technologies, such as deep sequencing and mass spectrometry, provides access to data on complete genomes, transcriptomes, as well as proteomes of multiple organisms of interest. There are various databases and repositories in public domains, for instance, Phytozome and NCBI, which host the sequence information of multiple species [16–18]. The advent of availability of complete genome sequences brings many opportunities and challenges to retrieve the information from the sequences, such as the identification and annotation of genes, and the possibility of drawing inferences about the evolutionary history of gene families, pathways, and even lineages across kingdoms [19, 20]. Several computational and statistical methods have been developed for the identification of genes and their comparative analysis among various genomes [21–24]. Reconstruction of the phylogeny and evolutionary history of gene families may provide the fundamental insights into the evolution of plants and their diversity from varied ecological habitats.

Here, we shall describe the methodologies used for the phylogenetic reconstruction and evolutionary analysis of the plant AGO gene family across genomes in broader perspective. Various aspects of the methods are explored to infer the evolutionary history of a gene family. Plant AGO family of proteins is used as a model to study the phylogeny and evolutionary history of such important gene families across green lineages.

2 Materials

2.1 Computational Resources

A Linux-like operating system [25] is preferred for analyzing phylogenies, as well as to study the molecular evolution of a gene family. Most of the recommended computational resources (Table 1)

Table 1
List of software and tools used for phylogenetic and evolutionary analysis

Sl No.	Software/tools	Objective/functions	Reference
1	BLAST	Homology search method to predict homologs	[29]
2	HMMER	Profile search methods to predict homologs	[31]
3	Clustal X	Non-iterative progressive alignment of sequences	[72]
4	MAFFT	Iterative progressive based heuristic alignment method	[75]
5	TrimAl	Heuristically search and eliminate the segments of contiguous non-conserved and non-informative positions in MSA	[38]
6	ReadAl	Conversion of various file formats	[38]
7	ProtTest	To optimize the best substitution model followed by a matrix	[51]
8	RAxML	Maximum likelihood phylogenetic inferences	[26]
9	PhyloBayes	Bayesian Monte Carlo Markov Chain (MCMC) sampler for phylogenetic reconstruction using protein alignments	[27]
10	MEGA	Phylogenetic inferences using distance-based and likelihood inferences	[41]
11	NOTUNG	Reconciliation of gene family tree with the species tree	[58]
12	DIVERGE	To calculate the shift in the evolutionary rates in the gene family after duplication	[59]
13	PAML	Phylogenetic and evolutionary analysis using maximum likelihood	[61]
14	CAPS	Identifies coevolution between amino acid sites using Blosum-corrected amino acid distances	[68]
15	MISTIC	To estimate the mutual coevolutionary relationship between two residues in a protein family using corrected mutual information (MI)	[70]

for biological sequence analysis are indeed Linux compatible (*see* **Note 1**).

2.2 Software and Tools

Various computational resources and tools have been developed for the phylogenetic and evolutionary analyses of gene families. A list of such resources, which have open access to academic and research use, is shown in Table 1. Several methods have been

proposed to test the evolutionary hypothesis based on various evolutionary models of phylogenetic inferences (*see* **Note 2**) [26, 27]. Performance of these tools are often limited by the size of the data and availability of computational resources (*see* **Note 3**) [26]. The compatibility of related softwares and files must be checked prior to the start of experiment.

2.3 Input Data and Files

The full-length sequences of gene and protein may be obtained by analyzing the transcriptome (CDS) and proteome (peptide) sequences of the plant lineages that may be accessed from various public repositories, such as NCBI (https://www.ncbi.nlm.nih.gov/home/proteins.shtml) [18], Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) [16], and MIPS (http://mips.helmholtz-muenchen.de/plant/genomes.jsp) [17]. Each sequence type (CDS or peptides) must be stored in a separate text file in an adequate format (*see* **Notes 4** and **5**).

3 Methods

In this section, we shall describe the methodologies adapted for the phylogenetic and evolutionary analysis of genes by taking AGO family of proteins as an example. Figure 1 describes a workflow for each of the experiments, such as mining of homologs, dataset preparation, optimization of best substitution model for the input dataset, phylogeny reconstruction, and testing evolutionary measures with a number of methods and models. Commands at the terminal in the Linux OS, such as Ubuntu, shall also be described in "Courier New" font throughout the text, for example,

```
$perl parsefile.pl
```

The file names in the command shall be italicized.

3.1 Mining of Homologs

Homologs are the genes that share common ancestor and are often classed as paralogs, orthologs, or xenologs [28]. Paralogs are the duplicates of a gene within the same genome, while orthologs are duplicates in different genomes that arise by speciation events from a common ancestor. Xenologs are homologs that come through horizontal gene transfer (HGT) from other genome (donor). For drawing the phylogenetic and evolutionary inferences, all the input sequences are assumed to be homologous. So, the first critical step for the evolutionary analysis of a gene family is to search for the homologs. Various methods have been proposed to find the homologs of a particular gene (*see* **Note 6**). Homology search methods [such as BLAST (basic local alignment search tool) [29]; http://blast.ncbi.nlm.nih.gov/Blast.cgi] scan a sequence database against a query through local searches and report sequences with statistically significant similarities. With the appropriate parameters, it calculates

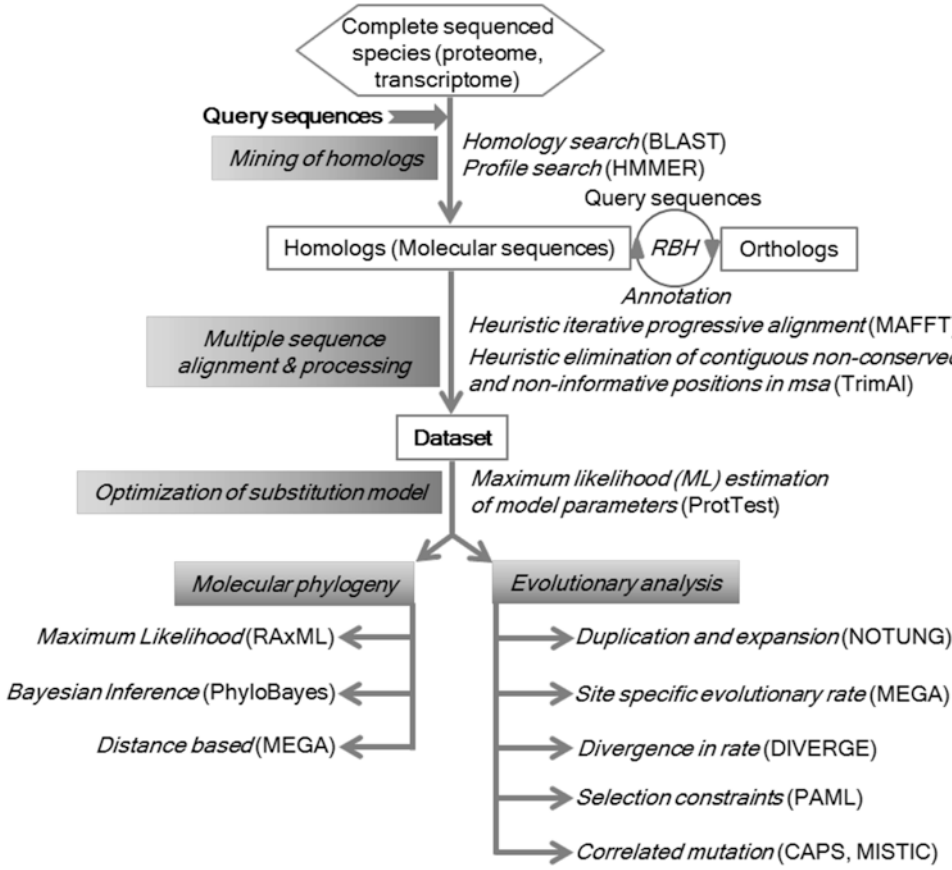


Fig. 1 An overview of steps involved in phylogenetic and evolutionary analysis of plant ARGONAUTE family of proteins

the pairwise distances with every sequence in the database for a query and quickly finds the best hit in the provided database. Profile search methods (as available, e.g., in HMMER [30, 31]; http://hmmer.org/) first construct a "profile" out of the homologs of a gene family by using approaches such as hidden Markov model (HMM); such profiles are then used to search homologs. Finally, it reports the potential homologs in a database that fits the profile. Profile searches are more sensitive than the homology searches (*see* **Note 6**).

Each homology search method needs a query (or reference) sequence. *Arabidopsis thaliana* is one of the most studied model plants for which maximum information is available for the biochemical and functional mechanism of the AGO proteins. *A. thaliana* genome harbors ten *AGO* genes (*AtAGO*) [7, 9]. In our description, these sequences are used as query to search its homologs in other plant genomes using both the homology search and profile search methods.

3.1.1 Homology Search

Sequences of each of the *AtAGO*s were used as query to search homologs in the transcriptomes and proteomes of other plants. Before executing the BLAST search, the database of the transcriptomes and proteomes of other species was constructed using the following command:

```
$formatdb -p F –i transcriptome.fasta -n transcriptome.db
```

Depending on the type of the sequences in query and the databases, an appropriate BLAST search module may be selected (*see* **Note 7**). The following command is used at the terminal to search the homolog of *AtAGOs* in the transcriptome database of other plant species:

```
$blastall -p blastn –i query.fasta -d transcriptome.db
-e 0.00001 -o blasthits
```

Here, options "-i," "-d," and "-o" specify the query (*AGO* nucleotide sequence), the subject data (transcriptome), and the output file, respectively. The option "-e" is the expected value ("-e") threshold ($\leq 10^{-5}$), which is used to limit the search of the "homologs" to the statistically significant numbers only. Based on the source of the query sequences and the subject databases, the "-e" value threshold may be adjusted (*see* **Note 8**). Finally, the homolog sequences are extracted and stored in the *homolog.fasta* text file.

3.1.2 Profile Search

Profile searches are more sensitive than the homology searches. This type of searches can detect the homolog that has largely diverged from its ancestor during the course of evolution. "HMMER" [31] is the most recommended tool package used for the mining of homologs using profile searches:

1. Generate an HMM profile of the sequences of AGO family using the HMMER software.
2. Generate a multiple sequence alignment (MSA) for the AGO homologs obtained from the blast analysis, either using ClustalX or MAFFT.
3. Convert the "fasta" format of MSA into the "Stockholm" format (".sto," with the help of "sreformat"), so that the HMMER could produce HMM profile. The following command is used in such a transformation:

   ```
   $sreformat Stockholm homolog_msa.fasta
   ```

4. Following this conversion of msa into the *homolog_msa.sto*, the HMM profile is generated with the help of "hmmbuild," using the following command:

   ```
   $hmmbuild homolog_msa.sto AGO.hmm
   ```

5. Using this profile (*AGO.hmm*), a set of subject sequences (e.g., from the proteome) of the other plant species is scanned using the following command:

```
$hmmsearch -E 1e-5 AGO.hmm subject.fasta -o hmmhits
```

Here, the expected value ("-E") threshold of $\leq 10^{-5}$ is applied to limit the search for the significant homologs; *hmmhits* is the output file.

3.1.3 Annotation of Homologs

Use the method of "reciprocal best hits" (RBH) to annotate each of the predicted homologs. This method is implemented in the "Proteinortho" perl program [32]. "Proteinortho" uses a BLAST-based approach to generalize the "reciprocal best alignment heuristic" and determines the sets of co-orthologs between the two compared datasets. For example, use "Proteinortho v2.3.0" to compare potential AGO homologs with the AtAGOs with the help of the following command:

```
$proteinortho_v2.3.0.pl -e=1e-5
At_AGOs.fasta AGO_homologs.fasta >orthologs
```

Option "-e" is the *e*-value threshold of $\leq 10^{-5}$; files *At_AGOs.fasta* and *AGO_homologs.fasta* contain the AGO peptide sequences of *A. thaliana* and the homologous peptide sequences from other plant lineages, respectively. Each of the homologous sequences from other plant genomes shall then be compared to the AGOs of *A. thaliana*. Orthologous groups are generated in the output file, *orthologs*, with "blastp" output details.

3.2 Dataset Preparation

To study the phylogenetic history and evolutionary relatedness among the identified protein homologs, the next step is to identify the homologous regions from the sequences. This helps assessing the character homology between the sequences, which is attained with the help of MSA. MSA tries to align the corresponding positions from all sequences (which may be believed to be descended from a common ancestor) [33]. Finally, it generates a matrix of having maximum positions (columns) with set of similar residues at the cost of generation of gaps, called indels (insertions and deletions). Various methods have been proposed for performing MSA for up to thousands of sequences (*see* **Note 9**).

1. Use MAFFT to globally align the AGO homologs with the help of the following command:

```
$mafft --op 1.53 --ep 0.123
--localpair AGO_homologs >AGO_msa
```

Here, option "--op" specifies the gap opening penalty, while "--ep" designates the gap extension penalty. Option "--localpair," also denoted as "L-INS-i," is used for multiple

alignment strategy. It uses iterative refinement method for the alignment using the "WSP (weighted sum-of-pairs)" and the "consistency scores" [34], which evaluates the consistency between a multiple alignment and pairwise alignment [34].

MSA generally produces poorly aligned regions, particularly when homologs from very distant lineages are aligned. In order to improve the phylogeny for deriving robust inferences about the evolutionary history of genes, elimination of these regions has been recommended as a "best practice" (*see* **Note 10**) [35, 36]. "Gblocks" [37] (http://molevol.cmima.csic.es/castresana/Gblocks.html) and "trimAl" [38] (http://trimal.cgenomics.org/) are two widely used tools that heuristically search and eliminate the segments of contiguous, non-conserved, and non-informative positions in matrix and produce a "trimmed" version that would be more suitable for phylogenetic inferences.

2. Eliminate multiple, large portion of poorly aligned regions in the *AGO_msa* by using "trimAl," with the help of the following command:

```
$./trimal -in AGO_msa -out AGO_msa_trim -automated1
```

With "-automated1" option, "trimAl" heuristically selects the best automatic method between "gappyout" (removes gap-rich columns) and "strict" (removes columns having residue similarity score lower than the threshold) methods to trim a given alignment. Here, "trimAl" considers the number of sequences in the alignment and average identity score among them [38]. The improved MSA after trimming is stored in *AGO_msa_trim*.

3.3 Phylogenetic Analysis

A phylogeny represents the arrangement of sequences (or homologs) that are linked through branches according to their evolutionary relationship [39, 40]. The topology of a tree indicates the evolutionary relationship between the analyzed sequences. Several methods have been proposed for phylogenetic reconstruction to molecularly capture the process in the DNA sequences that significantly may lead to evolution of protein function. Three different types of methods, distance-based methods, character-based (discrete data) probabilistic methods, and the methods that are driven by Bayesian inferences (BI), have been maximally used for phylogeny reconstruction. Distance-based methods, such as neighbor joining (NJ), are simple and suitable for small datasets. NJ calculates pairwise distances between sequence (homologs) pairs and infers the phylogenetic relatedness among them. MEGA [41] (http://www.megasoftware.net/home) is one of the largely used tools for distance-based methods. On the other hand, character-based methods, such as maximum likelihood (ML), use the evolutionary information between sequences. It calculates the likelihood

of a tree to follow a given evolutionary model. The tree with high likelihood score is considered as best to explain the evolutionary relationship between the sequences. Finally, the robustness of best topology may be assessed with statistical measures such as bootstrapping [42] or likelihood ratio test (LRT) [43]. RAxML (randomized accelerated maximum likelihood) [26] is widely used for the ML phylogenetic inferences. On the other hand, BI uses a coupled Markov chain Monte Carlo method [44] (MCMC, as implemented in PhyloBayes [27]; http://www.atgc-montpellier.fr/phylobayes/) for searching tree space. It approximates the probability for posterior distribution of a tree topology on a given prior distribution and the parameter space. It then calculates the ratio of probabilities of posterior versus prior distributions and decides whether to accept or reject the topology distribution. Finally, a consensus tree is generated on the basis of all the sampled trees with posterior probability on each branch as its support value (*see* **Note 11**).

In this section, stepwise experiments are described for conducting a phylogenetic analysis of AGO protein family using ML (as implemented in the RAxML program). Further, the BI and NJ methods, as implemented in PhyloBayes and MEGA, respectively, are also described to construct the AGO phylogeny.

3.3.1 Selection of the Best Evolutionary Model

For the given set of sequences, the ML estimate of the phylogeny corresponds to the preferred explanation for the topology of tree. The ML score (L_D) for a given set of sequences (D) under a phylogenetic hypothesis (H) can be estimated by calculating the probability of D for a given H,

$$L_D = P(D / H)$$

Further, H also depends on the model selected for estimating sequence evolution used for the phylogenetic reconstruction. To infer phylogenetic relationship, each method requires a prior assumption of some "evolutionary models" to explain the distribution pattern of amino acids at different positions in the dataset (*see* **Note 12**). The models are represented by weight matrices [45], such as PAM [46] and BLOSUM [47]. Now, advance matrices have been developed, such as JTT [48] and WAG [49] model to estimate the evolutionary history of a gene family across wide range of genomes, including across kingdoms. All the models correspond to the types of dataset on which they have built, for example, the RTREV substitution matrix is derived from virally encoded amino acid data, whereas CPREV is derived from chloroplast-encoded amino acid data. Models consider several parameters, such as the frequencies of amino acids and their rates of substitution during estimation of evolutionary history of a protein family. Based on

evidences, another parameter of rate variation among sites (invariant sites (I)) has also been proposed. For example, sites critical for biochemical functions of a proteins tend to be highly conserved. Most commonly, for examining alternate hypothesis, the gamma (γ) model of rate variation among sites on the protein length is used, which approximates the distribution of rates in a protein MSA [50]. Depending on the value of its "shape parameter" (α), the γ-distribution can be either L shaped (extreme rate heterogeneity in the MSA) or bell shaped (minor rate heterogeneity in the MSA). Further, the evolutionary rate may also vary across the branches on a tree (called heterotachy), particularly if the homologs belong to a large family of proteins from very diverse lineages. Most of the modern methods consider these parameters during the estimation of evolutionary relatedness between the homologs.

"ProtTest" [51] is one of the recommended tools to optimize the substitution model for the evolution of a protein family. "ProtTest" uses four strategies, fixed BIONJ, BIONJ, ML, or user defined, to calculate the likelihood score under 120 candidate substitution models of evolution. After likelihood computation, it uses three statistical frameworks, Akaike Information Criterion (AIC) [52], Akaike Information Criterion corrected (AICc), and the Bayesian Information Criterion (BIC) [53] for determining the models that best fit the given data. ProtTest is a ".jar" file and opens as graphical interface with JDK run-time environment or by using the following command at the terminal:

```
$java -jar ProtTest.jar
```

A window-like console appears on the screen. The input for "ProtTest" is the MSA file in phylip format. The phylip format may be obtained by using the program, "readAl," which is supplied with "trimAl," with the help of the following command:

```
$./readal -in AGO_msa_trim.fasta -out AGO_msa_trim.phy -phylip
```

"readAl" converts the files in various formats, such as clustal, mega, and phylip, which could be needed for a variety of tools employed during the phylogenetic and evolutionary analysis. "ProtTest" has been used to optimize the Jones-Taylor-Thornton (JTT) model with an estimated γ-distribution parameter (G) and the proportion of invariant sites (I) "(JTT + G + I)" as the best fitting model for the analysis of AGO family of proteins according to Akaike Information Criterion (AIC) framework [8].

3.3.2 Phylogenetic Analysis Using the Maximum Likelihood Approach

RAxML tool is most popularly used for ML phylogenetic analysis. RAxML is computationally efficient for large data set analysis [26]. RAxML may use different types of character sets, including nucleotides, amino acids, multistate characters, and binary data. RAxML is invoked with the help of the following command:

```
$ raxmlHPC -h
```

AGO_msa_trim file, in phylip format, is used as input for the RAxML program with the following command:

```
$ raxmlHPC -s AGO_msa_trim.phy -m PROTGAMMAJTT -f a -x
12345 -p 12345 -# 1000 -o yeast_AGO -n AGO_ML_tree
```

Option "-s" specifies the input file, *AGO_msa_trim*, which is used in phylip format. Option "-m" corresponds to the model of amino acid substitutions, and PROT corresponds to types of sequences, which, here, are proteins. GAMMAJTT specifies the amino acid substitution model (JTT) with optimization of substitution rates, using γ-model of rate heterogeneity (this shall estimate the alpha parameter). Available amino acid substitution models in RAxML are DAYHOFF, DCMUT, JTT, MTREV, WAG, RTREV, CPREV, VT, BLOSUM62, MTMAM, LG, MTART, MTZOA, PMB, HIVB, HIVW, JTTDCMUT, FLU, STMTREV, DUMMY, DUMMY2, AUTO, LG4M, LG4X, PROT_FILE, GTR_UNLINKED, and GTR. The option "-f a" corresponds to rapid bootstrap analysis and searches for the best-scoring ML tree in one program run. The option "-x 12345" specifies an integer number (random seed) and turns on rapid bootstrapping. The option "-p 12345" specifies a random number seed for the parsimony inferences, whereas the option "-# 1000" specifies the number of alternate runs on distinct starting trees. In combination with the "-b" option, this invokes a multiple bootstrap analysis.

First, RAxML generates best starting trees under parsimony criterion by randomly adding the sequences one by one. By using the *lazy subtree rearrangement* (LSR) method, RAxML clips and reinserts the "subtrees" at all possible locations of the branches of a tree. With the option "-# 1000," RAxML generates 1000 best starting trees, by sampling the sequences randomly. Hence, all 1000 searches will begin from different starting trees, which will increase the confidence of best-scoring ML tree. The output of RAxML analysis would be stored in files suffixed, *AGO_ML_tree*. A total of five output files are generated after the completion of RAxML execution: the bootstrapped trees (*RAxML_bootstrap.AGO_ML_tree*), the best-scoring ML tree (*RAxML_bestTree.AGO_ML_tree*), the best-scoring tree with BS support values as node labels (*RAxML_bipartitions.AGO_ML_tree*), the branch labels (*RAxML_bipartitionsBranchLabels.AGO_ML_tree*), and one summary file (*RAxML_info.AGO_ML_tree*) that contains information like likelihood values for each bootstrap and frequency of residues. The best ML tree would be in "newick" format that can be visualized with "FigTree" (http://tree.bio.ed.ac.uk/software/figtree/) (Fig. 2) or "Dendroscope" (http://dendroscope.org/).

To get insight into the direction in which evolution of AGO took place during the expansion of plant lineages, an AGO is

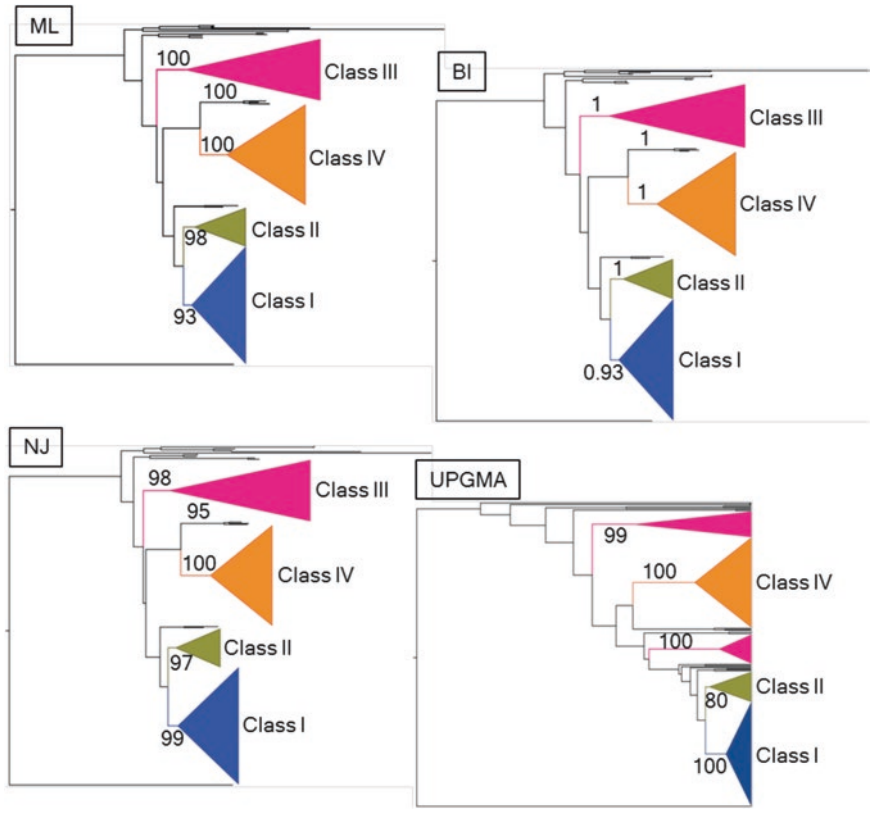


Fig 2 Phylogeny of plant ARGONAUTE (AGO) family using maximum likelihood (ML), Bayesian inferences (BI), neighbor joining (NJ), and unweighted pair group method with arithmetic mean (UPGMA) approaches. The values on the branches show the bootstrap (on the ML, NJ, and UPGMA trees) and posterior probability (on the BI tree). "Yeast AGO" was used to root the trees. All the three phylogeny methods (ML, BI, and NJ) produced trees showing the similar topologies, which suggest four classes of AGOs. The UPGMA clustering shows differences in topology, for instance, clade III has split into two separate clusters. Raw sequences used for constructing trees were the same as have been detailed in the "AGO_dataset" from Singh et al. (2015) [8]

selected as an out-group to root the phylogenetic tree (*see* **Note 13**) [54]. An out-group is specified by adding the option, -o *name_of_sequence* in the RAxML argument, where *name_of_sequence* is in the MSA. "yeast_AGO" was selected as out-group to root the plant AGO phylogeny. However, in the presence of more than one out-group species, two or more out-group sequences may also be used (command may be executed by adding the sequence name after comma in the command line option), for instance, "-o name_of_sequence1, name_of_sequence2."

3.3.3 Phylogenetic Analysis Using Bayesian Inference (BI)

PhyloBayes 4.1 [27] uses Bayesian Monte Carlo Markov Chain (MCMC) sampler for phylogenetic reconstruction. It is an open-source, command line program that can be easily installed at any Linux-like OS. The details about the options and parameters of

PhyloBayes can be obtained by invoking the program at the terminal (which contains the path in their directory):

```
$./pb
```

AGO_msa_trim file, in phylip format, is used as input for the PhyloBayes program with the help of the following command:

```
$./pb -d  AGO_msa_trim.phy -jtt -ncat 1 -dgam 4 -dc -
nchain 2 1000 0.3 50 -s AGO_BI_tree
```

Here, option "-d" specifies the dataset; "-jtt," the substitution model; "-ncat," the number of categories for the dataset; "-dgam," the discrete gamma values in which dataset may be divided on the basis of differences in evolutionary properties; and "-nchain," the number of chains in which dataset is executed in parallel to test the robustness of the analysis. Further, the robustness is evaluated with the help of convergence score between the chains; the lowest value corresponds to the strong convergence. Convergence score of ≤ 0.3 corresponds to the significant cut-off of the convergence value between two chains. The above command argument would run two chains in parallel, for a minimum of 5000 cycles. Then, every 1000 cycles, the two chains are automatically compared (with a burn-in equal to one fifth of the total length of the chain), and the run stops once all the discrepancies are lower or equal to 0.3 and all the effective sizes are larger than 50. The Bayesian consensus tree of the AGOs is reconstructed using four discrete gamma categories and JTT substitution models to fix exchange rates. Two independent chains were run to observe lower largest discrepancy (that indicates a good qualitative of the posterior consensus) across all bipartitions in the consensus topology (indicating that convergence between the two chains has been achieved) for searching tree space. Then, the probability of posterior distribution is approximated on the basis of most parsimonies of prior distribution and the parameter space. The BI consensus tree is stored in *AGO_BI_tree* with the posterior probability values for each node. This tree in the file can be visualized with "FigTree" (Fig. 2).

3.3.4 Phylogenetic Analysis Using Neighbor Joining

MEGA 5.2 [41] is a package of programs that allows investigating various evolutionary properties using a variety of models. It is often used to construct NJ-based phylogeny. MEGA can be downloaded from http://www.megasoftware.net/home. The input for MEGA is an alignment file, in "mega" format, such as *AGO_msa_trim.meg*. MEGA provides a user-friendly interface to work, where various options such as "construct/test neighbor-joining tree" may be selected from a drop-down menu in the "phylogeny" tab of the MEGA console. The tool prompts the user for choosing the file as well as for selecting the parameters for the phylogenetic analysis. JTT model with γ-distribution of rates (G) among invariant sites (I) ("JTT + I + G") has been tested to yield optimum results

during analysis of the evolutionary relatedness among the sequences. The option of partial deletion for gaps/missing data is used to account for highly variable sites. The robustness of tree topology is estimated with 1000 bootstrap replicates. The final phylogenetic tree (along with appropriate statistics) is displayed, which along with the complete session can be saved; the tree can be further visualized with the help of "FigTree" (Fig. 2).

Different phylogenetic methods may reconstruct the phylogenetic trees with differences in tree topologies for a given dataset. However, all the above methods produced overall similar tree topologies for *AGO_msa_trim* (Fig. 2). This suggests that the predicted evolutionary events, such as duplication, speciation, and expansion, in the AGO phylogeny are highly robust. Further, to rule out any effect of clustering properties among the sequences in the dataset, unweighted pair group method with arithmetic mean (UPGMA), as implemented in MEGA, has been used to cluster the dataset. The UPGMA produced a different topology than the trees by ML, BI, and NJ methods (Fig. 2).

3.4 Evolutionary Analysis

Knowledge of phylogeny helps in revealing the evolutionary history of AGO family in the plant kingdom. Members of a gene family may evolve through varied numbers of duplication, speciation, and transfer (such as horizontal gene transfer) events among the lineages, resulting in variations in the number of paralogs among plant genomes. Further, paralogs may have variations in evolutionary constraints across sites. Change in evolutionary rates may lead to substitution of residues at the functionally important sites with the residues of different physico-chemical properties. These differences at the functionally important sites may lead to the structural and functional diversification within the AGO gene family across different plant lineages.

3.4.1 Expansion of AGO Family During Evolution

Insights into duplication and how it has influenced the expansion of gene family are gained by reconciling the AGO gene family tree (GFT) with the species tree. The ML tree of the AGO, generated by RAxML, is used as GFT. The species tree shows actual evolutionary relationships between species. rRNA sequences could be a good marker of species phylogeny. Species tree (comprising of all the representative species used in GFT construction) may be generated with the help of NCBI Taxonomy Browser (http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi). Alternatively, a previously generated species tree [55, 56] from the literature (and included in various database [57]) may also be used for reconciliation. Notung [58] program may be used to identify events of duplications and losses of the AGO genes during evolution. Notung is a java-based graphical interface that may be utilized for reconciliation of gene tree with the species tree to

reconstruct the history of gene duplications and losses. The program may be downloaded from http://www.cs.cmu.edu/~durand/Notung/download.html as a zip file. After unzipping, Notung may be invoked using the following command:

```
$java -jar Notung.jar
```

Notung would appear as a window-like console on the screen. It requires a gene tree and a species tree as input in "newick" format. The content of the species information should be identical between the two trees. This can be estimated by evaluating the correspondence between the leaves of the species and gene trees. Here, each leaf label in the gene tree must include a substring of that species from which the gene was sampled. Notung compares gene tree with species tree and displays a reconciled tree in the tree panel with the inferred duplications (D) and losses (L) indicated on the tree. The D/L score of the reconciled tree is also displayed in the lower-left corner of the console.

After reconciliation, Notung infers orthologous and paralogous relationships and determines the lower and the upper bounds on the duplication events, where bounds are represented in terms of internal nodes in the species tree (which is relative to speciation events). The upper bound on the time of duplication indicates the most recent species in which the duplication may yet not be present. The lower bound indicates the oldest species in which the duplication must have been present. All such information, along with statistics on losses, may be viewed in a pop-up window by selecting "event summary" from "about this tree" menu. Duplications and bounds are identified by internal node names. The numbers of losses are listed with each associated node and taxon (*see* **Note 14**).

3.4.2 Evolutionary Diversification Among the AGO Sequences

After duplication, the AGO members of various phylogenetic clades may evolve independently. The evolutionary patterns within a phylogenetic clade of the AGO family could be inferred by calculating the position-specific evolutionary rate across sites on proteins. The rates of variations/substitutions of different amino acids on the AGO sequences are compared for each phylogenetic clades using a maximum likelihood method. MEGA 5.2 could be used to model the position-by-position likelihood of the substitution of residues. By invoking the "Rates" tab in the MEGA console, a drop-down menu shall appear from where the option of "estimate position-by-position rates (ML)" could be selected for estimation of site-specific relative evolutionary rates. Next, MSA (in mega format as input) and appropriate parameters/models (JTT + I + G) for the evolutionary analysis are selected. The position-by-position relative evolutionary rate for the each column of the MSA is computed. The option "partial deletion for gaps/missing" data is used to account for highly variable sites (all positions with <90% site

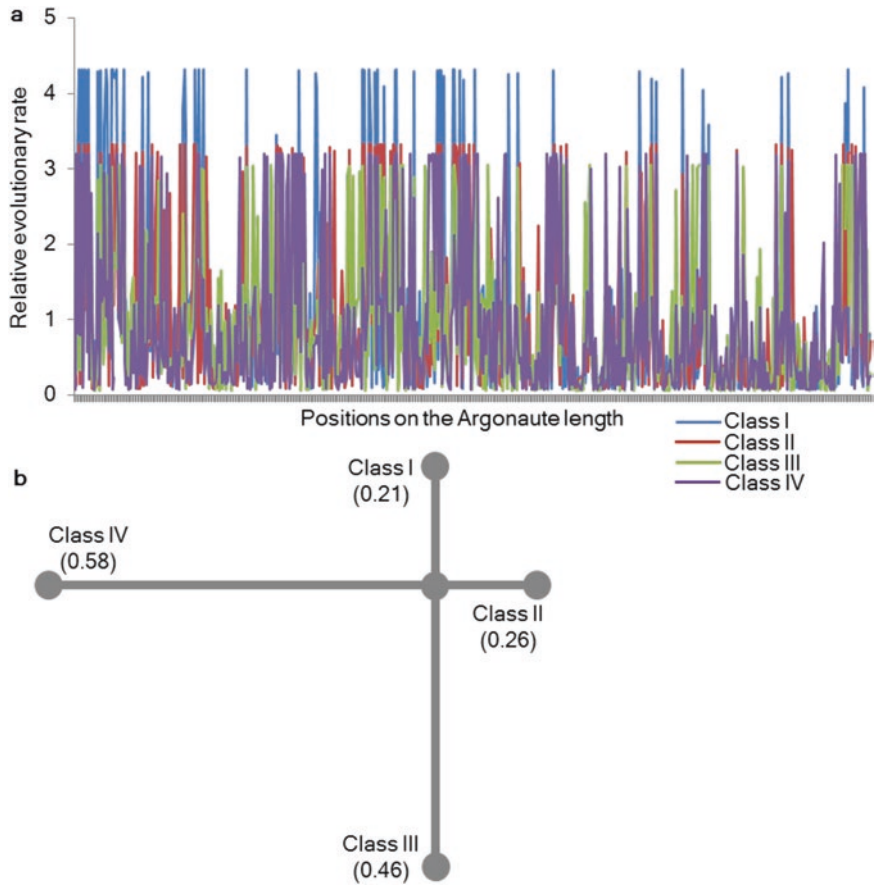


Fig. 3 Evolutionary diversification among four classes of plant AGOs. (**a**) Site-specific relative evolutionary rates of AGOs across classes I–IV. Sites showing rates <1 are evolving slower than average, and those with rates >1 are evolving faster. Sites in class I AGOs show more variation in the evolutionary rates than the AGOs in other classes. (**b**) Evolutionary distance (type I shift in substitution rate) between four AGO classes the form of a tree-like topology is presented. Class IV has the longest branch length, indicating that substitution rate has been shifted at maximum number of sites from ancestor after duplication. Sequences as described in Singh et al. (2015) [8] were used for the analysis

coverage are eliminated) (*see* **Note 15**). Mean (relative) evolutionary rates are scaled such that the average evolutionary rate across all sites is 1: sites showing a rate of <1 are evolving slower than average, and those with a rate of >1 are evolving faster than average. The tabulated values (in excel or CSV formats) of the relative evolutionary rates for each of the sites show that the position-specific relative rates are variable between various clades of AGOs (Fig. 3a).

These evolutionary diversification among the AGO sequences suggested the possibility that different classes of AGOs undergo site-specific rate shifts. The change in evolutionary rates may impact the structure and hence the biochemical functions of the AGOs. Hence, it will be interesting to examine the sites at which

the rates have been shifting among the various AGO clades. A likelihood ratio test, as implemented in DIVERGE 2.0 [59], could be used to identify the sites that show changes in the amino acids substitution rates across the clades (Type I divergence). DIVERGE is also a "Windows"-compatible user interface. The input file for such analysis is the MSA in fasta format. First, it generates an NJ tree using one of the various implemented parameters, such as the Kimura distance correction measure. Alternately, one may use the customized ML tree as input. Then, the DIVERGE performs pairwise comparisons of each of the monophyletic clades. All of such comparisons are used to calculate the coefficient of type I (θ_I) divergence and the posterior probability (PP) of shift in substitution rates for sites using the likelihood ratio test. The θ_I between different clades provides the statistical evidence for supporting the hypothesis of rate shift between AGO clades. Rejection of the null hypothesis ($\theta_I > 0$) indicates that after duplication, selection constraints may have altered many sites differently across the clades (thus, shifts in substitution rates). Tree-like topology of the AGO clades would also be generated, which represents the functional distance among the clusters (Fig. 3b). Such results point to evolutionary histories of various clades of the AGOs, which could have resulted in diversity in structural and functional properties.

3.4.3 Diversity in Selection Constraints Among AGOs

The rate of natural selection on the protein-coding genes varies across various families [13]. It has been noted that the substitution rate at the synonymous sites (d_S) is most frequent than at the nonsynonymous sites (d_N) [20, 60, 61]. Also, substitutions at the nonsynonymous site are less likely to be tolerated than at the synonymous sites within the genes. Hence, the ratio of d_N to the d_S ($\omega = d_N/d_S$) is a useful measure to estimate the strength of natural selection. Selections in favor of d_N (i.e., $\omega > 1$), called as positive (adaptive or diversifying) selection, may lead to "novelty" in the sequence and hence diversification in the protein family [61]. On the other hand, $\omega = 1$ and $\omega < 1$ indicates the neutral evolution and negative (purifying) selection, respectively.

To study the diversity in selection constraints (strength of purifying selection) across various phylogenetic clades of AGO family, we have used the CODEML program of the PAML package [61]. CODEML calculates the d_N, d_S and ω for the given set of genes by estimating the observed changes in the alignment of codons of the genes and from their phylogeny. The codon alignments may be generated using the program, "PAL2NAL" [62]. "pal2nal.pl" is a command line perl program that can be downloaded from http://www.bork.embl.de/pal2nal/. Using a protein sequence alignment along with the corresponding DNA sequences, "pal2nal.pl" constructs a multiple codon alignment:

```
$pal2nal.pl AGO_msa.aln AGO_cds.fasta -nogap -output
paml >AGO_codon_msa.nuc
```

In the above command, *AGO_msa.aln* and *AGO_cds.fasta* are the alignment (in clustal format) and the CDS sequence files, respectively, of AGOs from each clade. The option of "-nogap" would eliminate all the columns that have any gaps in the alignment file, *AGO_msa.aln*. The output codon alignment would be stored in *AGO_codon_msa.nuc* that could be directly used as an input in CODEML. For testing various models, CODEML also requires a phylogenetic tree. The ML trees of AGO have been used as input for the CODEML.

It may be assumed that adaptive evolution ($(d_N/d_S) > 1$) may have occurred at specific sites in certain branches of the phylogeny of the AGOs. Alternative hypotheses have been proposed for the variation in selection constraints over the null hypothesis of equal substitution rate across sites along all branches (*see* **Note 16**). CODEML [61] tests different categories of codon models. The main classes of models are "branch" models (that assume that ω may vary over different branches in the phylogeny), "site" models (assuming ω may vary at different sites in the gene), and "branch-site" models (where ω varies at particular sites, in particular branches). The following are the highly recommended alternative models in CODEML [61, 63–65]:

1. The *one-ratio model* (M0) calculates a single average ω over all branches in the phylogeny [66, 67]. However, it is not thought to be a realistic model to detect adaptive evolution.
2. The *nearly neutral model* (M1a) classifies codon sites into two groups: one for purifying selection ($\omega < 1$) and the other for neutral evolution ($\omega = 1$) [63, 65]. There is no assumption of positive selection ($\omega > 1$).
3. The *positive selection model* (M2a) is similar to M1a, with an extra group of codons sites subjected to positive selection ($\omega > 1$) [63, 65].
4. In the *beta model* (M7), the ω follows the beta distribution for describing the variation among sites [64]. The ω distribution takes a variety of shapes in the range between 0 and 1; hence, positive selections are not allowed.
5. The *beta and ω model* (M8) is same as the M7, except it allows some codons sites to be subjected to positive selection ($\omega > 1$) [64].

CODEML attempts to fit the above models, either on the alignment (site models), or the tree (branch models), or both (branch-site models). Then, it estimates the parameters (d_N, d_S, and ω) and a likelihood score (L) for each model. L is presented in the natural log ($\mathrm{Ln}L$). The CODEML program may be run with the following command:

```
$codeml codeml.ctl
```

"codeml.ctl" (detailed below) is the CODEML control file that specifies the alignment file, tree file, and all the options of models as well as other parameters:

```
seqfile = AGO.nuc * sequence data filename
treefile = AGO.nwk * tree structure file name
outfile = AGO_M1-out * main result file name

noisy = 9  * 0,1,2,3,9: how much rubbish on the screen
verbose = 1  * 0: concise; 1: detailed, 2: too much
runmode = 0  * 0: user tree;  1: semi-automatic;  2:
automatic
                  * 3: StepwiseAddition;
                    (4,5):PerturbationNNI; -2: pairwise

seqtype = 1  * 1:codons; 2:AAs; 3:codons-->AAs
CodonFreq = 2  * 0:1/61 each, 1:F1X4, 2:F3X4, 3:codon table

*      ndata = 10
clock = 0  * 0:no clock, 1:clock; 2:local clock;
3:CombinedAnalysis
aaDist = 0  * 0:equal, +:geometric; -:linear,
1-6:G1974,Miyata,c,p,v,a
aaRatefile = dat/jones.dat  * only used for aaseqs with
model=empirical(_F)
                  * dayhoff.dat, jones.dat, wag.dat, mt-
                    mam.dat, or your own

model = 0
                  * models for codons:
                      * 0:one, 1:b, 2:2 or more dN/dS
                        ratios for branches
                  * models for AAs or codon-translated
                    AAs:
                      * 0:poisson, 1:proportional,
                        2:Empirical, 3:Empirical+F
                      * 6:FromCodon, 7:AAClasses,
                        8:REVaa_0, 9:REVaa(nr=189)

NSsites = 0  * 0:one w;1:neutral;2:selection;
3:discrete;4:freqs;
                  * 5:gamma;6:2gamma;7:beta;8:beta&w;9:
                    beta&gamma;
                  * 10:beta&gamma+1; 11:beta&normal>1;
                    12:0&2normal>1;
                  * 13:3normal>0

icode = 0  * 0:universal code; 1:mammalian mt; 2-10:see
below
Mgene = 0
                  * codon: 0:rates, 1:separate; 2:diff
                    pi, 3:diff kapa, 4:all diff
                  * AA: 0:rates, 1:separate
```

```
fix_kappa = 0  * 1: kappa fixed, 0: kappa to be estimated
kappa = 2  * initial or fixed kappa
fix_omega = 0  * 1: omega or omega_1 fixed, 0: estimate
omega = .4 * initial or fixed omega, for codons or codon-
based AAs

fix_alpha = 1  * 0: estimate gamma shape parameter; 1: fix
it at alpha
alpha = 0. * initial or fixed alpha, 0:infinity (constant
rate)
Malpha = 0  * different alphas for genes
ncatG = 8  * # of categories in dG of NSsites models

getSE = 0  * 0: don't want them, 1: want S.E.s of estimates
RateAncestor = 1  * (0,1,2): rates (alpha>0) or ancestral
states (1 or 2)

Small_Diff = .5e-6
cleandata = 1  * remove sites with ambiguity data (1:yes,
0:no)?
*  fix_blength = -1  * 0: ignore, -1: random, 1: initial,
2: fixed
method = 0  * Optimization method 0: simultaneous; 1: one
branch a time

* Genetic codes: 0:universal, 1:mammalian mt., 2:yeast
mt., 3:mold mt.,
```

Here, "seqfile," "treefile," and "AGO_M1-out" are the codon alignment, the tree, and the output files, respectively. "model=0" is the "null model" that specifies that CODEML will not allow ω to vary among lineages in the phylogeny, whereas "NSsites" specifies whether the model uses ω to vary among the sites. To evaluate whether the data fits to an alternate evolutionary model, CODEML is rerun after specifying the use of another model. For example, to specify the use of branch model, the control file is edited with "model=2"; the tree file could be modified by putting the "hash (#)" tag on to the branch where positive selection has to be evaluated (the output file may be saved with a different name). After having estimated parameters (d_N, d_S, and ω) and a likelihood score (L) for both, the null model and the alternate (branch) model, likelihood ratio test (LRT) is used to calculate the LRT statistic to evaluate if data significantly fits better to the alternate model. LRT statistic of the alternate and null model is compared to the critical value from χ^2 distribution with an appropriate degree of freedom (differences in the number of parameters from alternate and null model) and the significance level (0.05). If the LRT statistic is greater than the critical value, the null hypothesis will be rejected: there are branches (marked with #) that have undergone adaptive evolution.

3.4.4 Correlated Evolution Within AGO Sequences

The evolution of residues in a protein sequence is "context dependent": the substitutions at a given site are affected by local structures, residues at the other sites, and related functions. Such context-dependent substitutions result in coevolution of residues at various sites that have implications for overall protein structure and function. Multiple approaches have been suggested to uncover the coevolving residues in a protein family. Here, we shall detail two such methods to examine the coevolving residues in the AGO sequences.

CAPS (coevolution analysis using protein sequences) [68] (http://bioinf.gen.tcd.ie/~faresm/software/software.html) is aimed at measuring the coevolution between amino acid sites that belong to the same protein (intramolecular coevolution). This software is written in C++ and may be used across various operating systems including Linux, Windows, and Mac OS. The following command is used to call the CAPS:

```
$./caps -F data_folder/ --intra -r 1000
```

where the option "-F" corresponds to the directory name where the data file is saved, the option "--intra" toggles to the intramolecular coevolution analysis, and the option "-r" allows the number of random cycles that are carried out for the input data; here 1000 simulated alignments of the same size and phylogenetic distributions as the input alignment were specified.

The input for CAPS is the AGO-MSA of the dataset in "fasta" format. CAPS, first, calculates a phylogenetic relationship among the sequences using an improved NJ method (BIONJ [69]), followed by reconstruction of ancestral sequences at all inner nodes using the ML approach. Then, mean transition parameter, θ, is calculated for each column using Blosum-corrected values for the transition between two amino acids [68]:

$$\bar{\theta} = \sum_{S=1}^{n} \left(\theta_{ij}\right) S$$

where θ_{ij} is the transition score for an amino acid substitution from state i to state j and n is the number of transitions in each column. Further, the variability of each amino acid transition from the mean transition parameter is estimated as

$$D_{ij} = \left[\theta_{ij} - \bar{\theta}\right] 2$$

Correlation coefficients between columns are then calculated:

$$\rho AB = \frac{\sum_{S=1}^{n} \left[\left(D_A\right)_S - \bar{D}_A\right]\left[\left(D_B\right)_S - \bar{D}_B\right]}{\sqrt{\sum_{S=1}^{n} \left[\left(D_A\right)_S - \bar{D}_A\right]^2} \sqrt{\sum_{S=1}^{n} \left[\left(D_B\right)_S - \bar{D}_B\right]^2}}$$

Only coevolving sites, showing correlation coefficient of ≥0.5, are considered. The significance of the correlation coefficients may be estimated by comparing the correlation coefficient to the distribution of 100 randomly sampled correlation coefficients, against which *P*-values are calculated.

Mutual information (MI)-based measure may also be used to identify correlated residue positions during the evolution of AGOs. Here, MI scores are calculated between pairs of columns in the MSA using MISTIC web server [70] with default parameters (http://mistic.leloir.org.ar/index.php). It assumes that mutations at two positions are uncorrelated. The frequency for each amino acid pair at the given pair of columns is calculated using sequence weighting and low-count corrections [71]. A weighted sum of the log ratios between the calculated and expected amino acid pair frequencies, at the corresponding columns, is calculated as the MI score. Using APC (average product correction) method [71], the background of the MI signal for each pair of residues is corrected. The MI score is translated into MI *Z*-scores by comparing the MI values with the distributed prediction scores obtained from a large set of randomized and permuted MSA. Finally, MI values with threshold of >6.5 are considered significant. It also calculates the cMI score (the sum of corresponding MI *Z*-score values over all residues within the MSA) and PMI score (a proximity average calculated as the average of cMI of all the residues within 5 Å of the evaluated amino acid) for each position in the MSA.

4 Notes

1. As several analyses are computationally intensive (e.g., likelihood estimates for phylogeny), we recommend using a minimum of 64 bit machine with 8 cores [likelihood analysis of a peptide (620 residues) alignment of 300 taxa, on a 2.8 GHz computer with 8 cores, could take around 72 h].
2. Most of the tools are dedicated for and follow a specific method, for instance, RAxML uses maximum likelihood (ML) method for phylogenetic inferences [26], while PhyloBayes is a Bayesian Monte Carlo Markov Chain (MCMC) sampler [27].
3. "PThread" version of RAxML is 10–11 times faster in reconstructing the ML tree from a DNA alignment of 20,000 bases from 125 taxa on an 8-core machine as compared to a single-core computer [26]. Although, both online and command line (stand-alone) versions are available for several tools (Table 1), the use of stand-alone/local versions may be preferred to maintain high speed and low run-times of the submitted "jobs."

4. The header (name) or annotation of each sequence (gene) should be compatible between the CDS and the peptide files for a particular species.
5. The FASTA format text files are widely used to store the molecular sequences. "fasta" (".fasta" or ".fa") format is one of the standard formats of the sequence files that is accepted by most of the sequence analysis tools.
6. Two approaches, based on homology search and profile search, are commonly used to find the homologs. Profile search methods work better when the query sequences and the database sequences are from distant species (including from different kingdoms).
7. The "BLAST" package has different types of search modules, such as "blastn," "blastp," "blastx," "tblastn," and "tblastx" [29], to assist various types of query and subject database sequences (nucleotides or peptides). The "blastn" program is used to search the nucleotide database against the nucleotide query sequences. Similarly, "blastp" program searches the protein database using a protein query sequence. While "blastx" program searches protein database using a translated nucleotide query, "tblastn" program searches translated nucleotide database using a protein query, and "tblastx" program searches translated nucleotide database using a translated nucleotide query.
8. If the query sequence and the database sequences are from different kingdoms, the "-e" value may be altered to detect the homologs, such as to $\leq 10^{-2}$. Further, it is helpful to parse the output to include only hits that have alignment length of $\geq 70\%$ of the query length and the identity of $\geq 30\%$ with the query.
9. CLUSTAL [72] (http://www.clustal.org/) is one of the widely used non-iterative progressive alignment programs. The iterative progressive alignment-based heuristic methods are MUSCLE [73] (http://www.drive5.com/muscle/), T-Coffee [74] (http://www.tcoffee.org), and MAFFT [75] (http://mafft.cbrc.jp/alignment/software/). Probabilistic approaches, such as maximum likelihood and Bayesian Markov chain process, have also been recommended for the alignment of sequences with high accuracy.
10. The positions corresponding to poorly aligned regions may not be homologous or may have been saturated or eroded. These uninformative sites may produce artifacts during construction of a robust phylogeny. The positions that correspond to the uninformative sites are excluded during trimming. The inferences of the phylogenetic relatedness among the AGO family members (Fig. 2) are based on the trimmed dataset.

Hence, trimming of poorly aligned regions in large dataset have been recommended.

11. ML and BI are extensive, and they are suitable particularly for large datasets. Also, these methods are more computationally demanding, thus needing more run-time.
12. An evolutionary model represents the tendency of a residue to substitute for another during the course of evolution. This helps in tracing the phylogenetic history of each of the homologs.
13. An out-group can be selected on the basis of prior knowledge. It is generally a homolog from a distant (out-group) species, such as from unicellular lineages [54, 76]. In the absence of such out-group species (for instance, while working with a highly diverse set of sequences), the last sequence in the MSA may be used as an out-group [77].
14. The expansion history of the AGO family is based on the duplication and loss events in the genes from ancestor to descendent. Since, the dataset comprised of only the species whose genomes have been sequenced, the measured number of duplication and loss events at the ancestor node may still differ from the actual.
15. The estimated relative evolutionary rate at positions with <90% site coverage for small dataset (such as alignment having <200 sequences that may not have experienced repeated changes sufficient number of times) may not be robust.
16. The inferences of the selection constraints largely depend on the source of data. It has also been suggested that sampled dataset should not be from widely distributed species. Since in the absence of intermediate species, the exact path of evolution of gene of interest might not be estimated, and hence it can either overestimate the sites having positive selection or high substitution rate at the non-synonymous sites. Conversely, if the sequences do not have significant diversity, such as from the similar genera, the sites having selection constraints may not be significantly identified. Further, with the optimum dataset, the estimated selection constraint at the terminal positions in the genes may be artifacts as the terminal positions accumulate more substitutions among the orthologs than in the middle or functional domains.

Acknowledgments

Financial assistance from MPG India partner program of Max Planck Society and Department of Science and Technology, India, the WHEAT Competitive Grants Initiative, CIMMYT and the CGIAR (A4031.09.10), and core funding from IISER-Kolkata is thankfully acknowledged.

References

1. Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev 16(21):2733–2742. doi:10.1101/gad.1026102
2. Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9(1):22–32. doi:10.1038/nrm2321
3. 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
4. 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(11):7829–7840. doi:10.1074/jbc.M112.441030
5. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305(5689):1434–1437. doi:10.1126/science.1102514
6. Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159. doi:10.1146/annurev-arplant-050312-120043
7. Baulcombe D (2004) RNA silencing in plants. Nature 431(7006):356–363. doi:10.1038/nature02874
8. Singh RK, Gase K, Baldwin IT, Pandey SP (2015) Molecular evolution and diversification of the Argonaute family of proteins in plants. BMC Plant Biol 15(1):1–23. doi:10.1186/s12870-014-0364-6
9. Hock J, Meister G (2008) The Argonaute protein family. Genome Biol 9(2):210. doi:10.1186/gb-2008-9-2-210
10. Singh RK, Pandey SP (2015) Evolution of structural and functional diversification among plant Argonautes. Plant Signal Behav 10(10):e1069455. doi:10.1080/15592324.2015.1069455
11. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. Cell 133(1):116–127. doi:10.1016/j.cell.2008.02.034
12. Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13(7):350–358. doi:10.1016/j.tplants.2008.04.007
13. Chothia C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. EMBO J 5(4):823–826
14. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW (2002) Evolutionary rate in the protein interaction network. Science 296(5568):750–752. doi:10.1126/science.1068696
15. Waxman D, Peck JR (1998) Pleiotropy and the preservation of perfection. Science 279(5354):1210–1213. doi:10.1126/science.279.5354.1210
16. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40(Database issue):D1178–D1186. doi:10.1093/nar/gkr944
17. Mewes HW, Frishman D, Guldener U, Mannhaupt G, Mayer K, Mokrejs M, Morgenstern B, Munsterkotter M, Rudd S, Weil B (2002) MIPS: a database for genomes and protein sequences. Nucleic Acids Res 30(1):31–34. doi:10.1093/nar/30.1.31
18. Pruitt KD, Tatusova T, Maglott DR (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 33(Database issue):D501–D504. doi:10.1093/nar/gki025
19. Clarke JT, Warnock RC, Donoghue PC (2011) Establishing a time-scale for plant evolution. New Phytol 192(1):266–301. doi:10.1111/j.1469-8137.2011.03794.x
20. Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. Nat Rev Genet 11(8):572–582. doi:10.1038/nrg2808
21. Eddy SR, Durbin R (1994) RNA sequence analysis using covariance models. Nucleic Acids Res 22(11):2079–2088. doi:10.1093/nar/22.11.2079
22. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95(25):14863–14868. doi:10.1073/pnas.95.25.14863
23. Gotoh O (1982) An improved algorithm for matching biological sequences. J Mol Biol 162(3):705–708. doi:10.1016/0022-2836(82)90398-9
24. Krogh A (1998) An introduction to hidden Markov models for biological sequences. New Compr Biochem 32:45–63. doi:10.1016/S0167-7306(08)60461-5
25. Bach MJ (1986) The design of the UNIX operating system, vol 5. Prentice-Hall Englewood Cliffs, NJ
26. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses

with thousands of taxa and mixed models. Bioinformatics 22(21):2688–2690. doi:10.1093/bioinformatics/btl446

27. Lartillot N, Rodrigue N, Stubbs D, Richer J (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol 62(4):611–615. doi:10.1093/sysbio/syt022
28. Fitch WM (2000) Homology: a personal view on some of the problems. Trends Genet 16(5):227–231. doi:10.1016/S0168-9525(00)02005-9
29. 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
30. Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14(9):755–763. doi:10.1093/bioinformatics/14.9.755
31. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39(Web Server issue):W29–W37. doi:10.1093/nar/gkr367
32. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ (2011) Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12(1):124. doi:10.1186/1471-2105-12-124
33. Boussau B, Daubin V (2010) Genomes as documents of evolutionary history. Trends Ecol Evol 25(4):224–232. doi:10.1016/j.tree.2009.09.007
34. Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33(2):511–518. doi:10.1093/nar/gki198
35. Atkinson GC, Baldauf SL (2011) Evolution of elongation factor G and the origins of mitochondrial and chloroplast forms. Mol Biol Evol 28(3):1281–1292. doi:10.1093/molbev/msq316
36. Christin PA, Spriggs E, Osborne CP, Stromberg CA, Salamin N, Edwards EJ (2014) Molecular dating, evolutionary rates, and the age of the grasses. Syst Biol 63(2):153–165. doi:10.1093/sysbio/syt072
37. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17(4):540–552. doi:10.1093/oxfordjournals.molbev.a026334
38. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25(15):1972–1973. doi:10.1093/bioinformatics/btp348
39. Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155(3760):279–284. doi:10.1126/science.155.3760.279
40. Page RD, Holmes EC (2009) Molecular evolution: a phylogenetic approach. John Wiley & Sons, New York, NY
41. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739. doi:10.1093/molbev/msr121
42. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39(4):783–791. doi:10.2307/2408678
43. Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst Biol 55(4):539–552. doi:10.1080/10635150600755453
44. Larget B, Simon DL (1999) Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol Biol Evol 16(6):750–759. doi:10.1093/oxfordjournals.molbev.a026160
45. Tomii K, Kanehisa M (1996) Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins. Protein Eng 9(1):27–36. doi:10.1093/protein/9.1.27
46. Dayhoff MO, Schwartz RM (1978) A model of evolutionary change in proteins. In: Dayhoff MO (ed) Atlas of protein sequence and structure, vol 5. National Biomedial Research Foundation, Washington DC, pp 345–358
47. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A 89(22):10915–10919. doi:10.1073/pnas.89.22.10915
48. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8(3):275–282. doi:10.1093/bioinformatics/8.3.275
49. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18(5):691–699. doi:10.1093/oxfordjournals.molbev.a003851
50. Yang Z (1996) Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. J Mol Evol 42(5):587–596. doi:10.1007/BF02352289
51. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein

evolution. Bioinformatics 21(9):2104–2105. doi:10.1093/bioinformatics/bti263

52. Akaike H (1973) Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika 60(2):255–265. doi:10.1093/biomet/60.2.255
53. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464. doi:10.1214/aos/1176344136
54. Maddison WP, Donoghue MJ, Maddison DR (1984) Outgroup analysis and parsimony. Syst Biol 33(1):83–103. doi:10.1093/sysbio/33.1.83
55. Hedges SB, Kumar S (2009) The timetree of life. OUP Oxford,
56. Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu YL, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donoghue MJ, Soltis PS (2011) Angiosperm phylogeny: 17 genes, 640 taxa. Am J Bot 98(4):704–730. doi:10.3732/ajb.1000404
57. Piel WH, Donoghue M, Sanderson M, Netherlands L TreeBASE: a database of phylogenetic information. In: Proceedings of the 2nd International Workshop of Species 2000, 2000.
58. Chen K, Durand D, Farach-Colton M (2000) NOTUNG: a program for dating gene duplications and optimizing gene family trees. J Comput Biol 7(3-4):429–447. doi:10.1089/106652700750050871
59. Gu X, Vander Velden K (2002) DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. Bioinformatics 18(3):500–501. doi:10.1093/bioinformatics/18.3.500
60. Gaucher EA, Gu X, Miyamoto MM, Benner SA (2002) Predicting functional divergence in protein evolution by site-specific rate shifts. Trends Biochem Sci 27(6):315–321. doi:10.1016/S0968-0004(02)02094-7
61. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24(8):1586–1591. doi:10.1093/molbev/msm088
62. Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res 34(suppl 2):W609–W612. doi:10.1093/nar/gkl315
63. Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148(3):929–936
64. Yang Z, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155(1):431–449
65. Yang Z, Wong WS, Nielsen R (2005) Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol 22(4):1107–1118. doi:10.1093/molbev/msi097
66. Yang Z, Nielsen R (1998) Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J Mol Evol 46(4):409–418. doi:10.1007/PL00006320
67. Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 11(5):725–736. doi:10.1093/oxfordjournals.molbev.a040153
68. Fares MA, McNally D (2006) CAPS: coevolution analysis using protein sequences. Bioinformatics 22(22):2821–2822. doi:10.1093/bioinformatics/btl493
69. Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14(7):685–695. doi:10.1093/oxfordjournals.molbev.a025808
70. Simonetti FL, Teppa E, Chernomoretz A, Nielsen M, Marino Buslje C (2013) MISTIC: mutual information server to infer coevolution. Nucleic Acids Res 41(Web Server issue):W8–14. doi:10.1093/nar/gkt427
71. Buslje CM, Santos J, Delfino JM, Nielsen M (2009) Correction for phylogeny, small number of observations and data redundancy improves the identification of coevolving amino acid pairs using mutual information. Bioinformatics 25(9):1125–1131. doi:10.1093/bioinformatics/btp135
72. Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
73. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797. doi:10.1093/nar/gkh340
74. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302(1):205–217. doi:10.1006/jmbi.2000.4042
75. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30(14):3059–3066. doi:10.1093/nar/gkf436

76. Pearson T, Hornstra HM, Sahl JW, Schaack S, Schupp JM, Beckstrom-Sternberg SM, O'Neill MW, Priestley RA, Champion MD, Beckstrom-Sternberg JS (2013) When outgroups fail; phylogenomics of rooting the emerging pathogen, Coxiella burnetii. Syst Biol 62(5):752–762. doi:10.1093/sysbio/syt038

77. Jill Harrison C, Langdale JA (2006) A step by step guide to phylogeny reconstruction. Plant J 45(4):561–572. doi:10.1111/j.1365-313X.2005.02611.x

INDEX

A

Adapter..... 97, 98, 105–107, 110, 114–116, 118–119, 122, 123, 125, 127
AGO1 .....2–6, 8–12, 24–31, 40, 45–46, 48, 51, 52, 55, 56, 62, 85, 94, 101, 109, 138, 139, 141, 142, 145, 159, 167, 174, 191, 233, 268
AGO2 ..... 2, 3, 6, 9, 11, 75, 85, 159, 167, 174, 241
AGO3 .....2, 6, 8, 11, 159, 167
AGO4 .....2, 3, 5, 6, 8–12, 129, 130, 132–134, 146, 174, 219, 222, 223, 243, 268
AGO5 .....2, 8, 9
AGO6 ..... 2, 6, 8, 129–134, 219, 222, 223
AGO7 ..... 2, 4, 6, 8, 9
AGO8 2
AGO9 .....2, 6, 8, 10, 11, 129, 268
AGO10 ..... 2–4, 6, 8, 9, 11, 12, 145, 268
AGO18 ..... 2, 3, 6, 9, 11
AGO–binding platform 257–265
AGO hook 5, 10, 257, 259–261
Agotron 12
Agrobacterium tumefaciens.....24, 28, 74, 76, 83, 176, 183, 192
Agroinfiltration 74–76, 83–85, 108
Antibody.....24–26, 28, 29, 35, 51, 60, 68, 78, 81, 87, 88, 90, 95, 99, 101, 109, 130–131, 133, 134, 138, 139, 141, 142, 148, 149, 152–154, 201–203, 205, 206
Antiviral 2, 6, 9, 41, 75, 93, 174
Arabidopsis thaliana (*Arabidopsis*) 1, 40, 55, 94, 101, 108, 109, 129, 130, 132, 137, 140–142, 145–156, 174, 203, 212, 219, 227, 233, 271, 273
Argonaute (AGO)..... 1, 7–8, 23, 39, 55, 73, 93, 129, 137, 145, 159, 174, 191, 219, 227, 241, 257, 267

B

Bacteria.....1, 11, 74, 79–81, 83, 84, 89, 90, 183, 184, 186, 192, 228, 230, 237, 255
Bayesian inference (BI)..... 274, 275, 278–280, 290
BLAST..... 170, 175, 177, 178, 185, 242, 258, 260, 264, 270, 272, 273, 289
BRAT-nova 220–222, 224
BY-2 ..... 40, 41, 43, 51, 56–57, 60–62, 65, 68

C

Catalytic mutant 56, 62, 64, 65, 69
Chiasma.....146, 153
Chlamydomonas reinhardtii 1, 5, 159
Chromosome..... 10, 93, 145–147, 150–153, 155, 156, 220, 221
Cleavage 1–4, 6, 9, 10, 23–25, 40, 41, 50–52, 55, 64, 65, 81, 94, 102, 113, 115, 118, 120, 121, 123–125, 137, 159, 174, 229, 238
Cross-linking.....11, 146, 205, 214–215
Crystallography 2, 227, 228, 235
Cytosine 50, 220, 222, 268

D

Deadenylation 5, 23, 24, 56, 65, 137
Decapping 5, 34, 35
Degradome.....2, 10, 11, 113, 115, 118, 120, 121, 123–125
Dicer–like (DCL).....3, 6, 9, 39, 40, 174
Differentially methylated region (DMR) 220, 223, 224
Digoxigenin (DIG) 58, 59, 66, 67, 147, 152, 199–209, 212
DNA extraction..... 162–163, 193
DNA repair 145, 150, 151, 153
DnaSP 259, 263, 264
Domain rearranged methyltransferase (DRM) 3, 5, 6

E

Escherichia coli 76, 79–81, 83, 88, 193, 194, 228
Evolution.....2, 6, 241, 257–265, 267–290

F

FASTA 177, 178, 220, 221, 248, 250, 253, 254, 272, 283, 287, 289
Fluorescence in situ hybridization (FISH).....145–148, 152, 153, 156

G

Geneious 259–261
Genetic transformation 173, 176, 177, 184
Genotoxic agents 149

Alberto Carbonell (ed.), *Plant Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1640, DOI 10.1007/978-1-4939-7165-7, © Springer Science+Business Media LLC 2017

Green fluorescent protein (GFP) 24–26, 28–30, 35, 40, 48, 74, 75, 78, 81–83, 85, 87, 90, 196
GW182 ..... 5, 242, 255, 257

H

Heat shock protein 90 (HSP90) 10, 40, 41
Hidden Markov model (HMM) 271, 272
High-throughput sequencing 11, 93–111, 114, 178, 199, 211
HMMER 242, 271, 272
Hybridization 59, 67, 97, 102, 109, 111, 122, 145, 147–148, 152–153, 201–203, 205–209, 211, 213, 215, 216

I

Immunoprecipitation (IP) 10, 11, 93, 94, 98, 100, 101, 106, 107
Immunostaining 130, 133
In silico 114, 248, 254–255
Intron 12, 178, 264
In vitro 24–25, 39–52, 55–69, 227
In vivo 9, 11, 24–25, 40, 41, 73–90, 114, 228
IUPred 258, 260
iWsearch 248, 250, 251

L

Library 94, 98, 100, 103–110, 114, 115, 118, 122–127, 159, 164–165, 170, 199, 222, 253
Luciferase 74, 81–83, 85, 87, 90

M

MAFFT 272, 273, 289
Maximum likelihood (ML) 259, 261, 262, 265, 274–278, 280, 281, 283, 284, 287–289
MEGA 259, 274, 275, 279–281
Megagametogenesis 9, 174
Meiosis 7, 9–10, 145, 150, 151, 153
Metaphase 151, 153
Methylation 3, 5–6, 10, 23, 129–134, 174, 219–224
MicroRNA (miRNA) 2–4, 10–12, 23–25, 39, 55–58, 63–64, 74, 75, 80–85, 87, 90, 113, 114, 137, 138, 159, 173, 174, 191, 200, 208, 209
Microscopy 45, 60, 61, 131, 133, 134, 146–149, 151, 193, 195, 237
MID 2, 40, 174, 227–238
MIPS 270
Multiple sequence alignment (MSA) 272–274, 276, 278, 281, 283, 287, 288, 290
Multiplex 106–108, 114, 123, 125
Muscle 258, 261, 289

N

NCBI 177, 178, 193, 220, 221, 223, 268, 270, 280
Neighbor joining (NJ) 274, 275, 278–280, 283, 287
Nick translation 147, 152
Nicotiana benthamiana 9, 24–25, 28–29, 40, 73–75, 78–81, 83, 84, 89, 90, 108, 118
Nicotiana tabacum 40, 41
Nonradioactive 199–209, 211–216
Northern blot 58–59, 65–68, 96, 101–102, 109, 199, 211, 212, 214, 216
NRPD1 ..... 130, 133–134
NRPE1 ..... 130, 133–134, 242–245, 257
Nuclear magnetic resonance (NMR) 227–229, 231–235, 237
Nuclease 32–34, 61, 95, 97, 99, 100, 103, 107–109, 111, 116–123, 126
Nuclei 45, 73, 130–134, 151, 171, 201, 212, 213
Nucleus 5, 129, 130, 151

O

Oryza sativa 212, 268

P

Parallel analysis of RNA ends (PARE) 114, 125
PAZ 2, 174
Percoll 41, 42, 44, 51, 57, 61
Perl script 178, 220
Phased secondary small interfering RNA (phasiRNA) 4, 114
PhyloBayes 275, 278, 288
Phylogenetic analysis 259–262, 274–280
PhyML 261
Physcomitrella patens 1
Phytozome 170, 268, 270
PIWI 2, 174, 191, 237
Pollen 9, 145, 146, 150, 151, 196
Polyadenylation 62–63
Polysome 4, 137–143
Position–specific scoring matrix (PSSM) 245, 247, 248, 253–255
Posttranscriptional gene silencing (PTGS) 1, 2, 6, 267
Protoplast 5, 39, 48, 51, 52, 56, 60–62, 68, 192, 194–195
ProtTest 276
Python 242, 253

R

RADAR 258, 260, 261
Randomized accelerated maximum likelihood (RAXML) 259, 261, 275, 276, 278, 280, 288
Rapid amplification of cDNA ends (RACE)
cRACE 24–25, 27–28, 33–35
3' RACE 24, 27, 32, 194, 196
5' RACE 194, 196
RLM-3' RACE 24–25, 27, 32–33
Renilla 74, 81–83, 85, 87, 90

Repeat array 260–264
Reporter 5, 62, 73, 74
RIP–Seq 11, 94, 108, 109
RNA decay 23
RNA-dependent DNA methylation (RdDM) 3, 5–6, 9, 129, 132, 219
RNA-dependent RNA polymerase (RDR) 5, 6, 173, 174
RNA extraction 27, 95, 99–101, 108, 117–118, 175, 179, 212, 213
RNA–induced silencing complex (RISC) . 39, 48, 51, 52, 56–58, 63–66, 68, 73, 85, 137, 174, 243, 254, 257, 267
RNA polymerase V (Pol V) 3, 5, 6, 10, 129
RNA polymerase IV (Pol IV) 5, 129
RNA–Seq 96–97, 110, 178
RNA silencing 2, 41, 56, 73, 93, 137, 241, 257, 264

S

Salt stress 6, 11
Salvia miltiorrhiza 173–186
Sensor 74, 75, 81–85, 87, 90
Setaria italica 8
Shoot apical meristem (SAM) 6, 7, 9, 174
Small interfering RNA (siRNA) 3–6, 8, 9, 24, 39–41, 46–48, 50–52, 113, 114, 173, 174, 191, 200, 209, 216
Small interfering RNA independent of DCLs (sidRNA) 3, 6, 12
Small RNA (sRNA) 1–3, 5, 6, 9, 11, 12, 23, 39–41, 55, 56, 63, 73, 93, 100–102, 109, 113–127, 129, 137, 145, 146, 173, 174, 191, 199–209, 211–216, 219, 227, 228, 232, 257, 267
Solanum lycopersicum 191
Stripping 86, 90, 104, 201–202, 206–207, 209
Structure 2, 9, 12, 62, 74, 81, 93, 173, 227–238, 242, 264, 280, 282, 287
Sucrose gradient 138
Suppressor of TY insertion 5 (SPT5) 5, 10, 242, 245, 257
Suppressor of TY insertion 6 (SPT6) 248–251, 253, 255

T

Tag-insertion 159–161, 164–165
Tandem repeat 258–260
Target cleavage 4, 23, 24, 50–52, 55, 64, 65
Target RNA 2–4, 10–12, 23–25, 40, 41, 48–50, 66, 93, 94, 98, 100, 101, 106, 107, 137
Ternary complex 11, 12, 93
Trans-acting small interfering RNA (tasiRNA) 4, 8, 200, 209
Transcriptional gene silencing (TGS) 1, 6, 139, 141, 219, 257
Transcription factor 12
Transfer 29, 31, 43–46, 49, 50, 59, 60, 62, 63, 65, 67, 68, 77, 81, 85–87, 89, 95, 98–102, 104–107, 109, 110, 117, 119, 121, 131, 133, 134, 140–142, 150, 154, 156, 162, 164–166, 169, 171, 179, 180, 183–185, 194–196, 199, 201, 204–208, 212–216, 230, 231, 236, 237, 270, 280
Translational repression 1–5, 9, 12, 41, 55, 64, 65, 137, 138, 159, 174
T-REKS 259–261
TrimAl 274, 276

U

Uridylation 23–35

V

Vacuole 44, 45, 60, 61
Viroid 9
Virus 9, 41, 73, 108, 191, 241, 243

W

Web portal (Whub) 242, 245, 248, 250, 253
Western blot 59–60, 64–66, 68, 75, 77–78, 81, 85, 95–96, 99, 101, 142
WG/GW 241–255
Whole–genome bisulfite sequencing (WGBS) 219, 220
Wsearch 248–250, 253, 255

Z

Zea mays 8