Signaling and Communication in Plants

Ramanjulu Sunkar *Editor*

# MicroRNAs in Plant Development and Stress Responses

Springer

# Signaling and Communication in Plants

## Series Editors

František Baluška
Department of Plant Cell Biology, IZMB, University of Bonn, Kirschallee 1,
D-53115 Bonn, Germany

Jorge M. Vivanco
Center for Rhizosphere Biology, Colorado State University, 217 Shepardson Building,
Fort Collins, CO 80523-1173, USA

For further volumes:
http://www.springer.com/series/8094

Ramanjulu Sunkar
Editor

# MicroRNAs in Plant Development and Stress Responses

Springer

*Editor*
Ramanjulu Sunkar
Oklahoma State University
Dept. Biochemistry & Molecular Biol
Noble Research Center 246
Stillwater, Oklahoma
USA

ISSN 1867-9048 e-ISSN 1867-9056
ISBN 978-3-642-27383-4 e-ISBN 978-3-642-27384-1
DOI 10.1007/978-3-642-27384-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012932737



Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

Proteins are the building blocks of all living cells. Cell fate that includes the type of cell, its function, and the timing of its death are largely determined by which proteins are produced in the cell, and at what quantities and when they are produced. MicroRNAs (miRNAs) regulate the number of protein molecules produced by the messenger RNA molecule in specific cell type at particular developmental stage, thus, emerged as critical regulators of gene expression at posttranscriptional level. The lessons from plant miRNA biology are quite clear. These are the major regulators of gene expression by virtue of their preponderance to target transcription factors. The silencing or fine-tuning of miRNA target genes at appropriate places and times allows the plant to grow and complete its' life cycle normally. On the same lines, miRNA-controlled regulation of gene expression is necessary for plants' adaptation to biotic and abiotic stresses including the lack of nutrients. This book highlights the roles of various miRNAs that control and regulate these diverse plant processes, which are discussed in a detailed manner by expert contributors. Expert authors also emphasize the current challenges and outstanding questions for future research in this field. Thus, this is a comprehensive book on plant miRNA biology covering wide range of topics in the field.

This book begins with a chapter by Zhixin Xie and colleagues, who introduce the plant small RNA world. In this chapter, authors describe the diverse small RNAs and small RNA pathways in plants, including their biogenesis and mode of function. In chapter "Role of microRNA miR319 in plant development", Palatnik and colleagues discuss the role of miR319 and TCP factors in leaf morphogenesis. Plant developmental progression from one phase to the other seems to be controlled by two miRNAs, miR156 and miR172. In chapters "The roles of miR156 and miR172 in phase change regulation" and "Roles of miR156 and miR172 in reproductive development", Rebecca Schwab discusses the role of miR156 and miR172 in phase change transitions from juvenile-to-adult-to-reproductive stages. Plant small RNA pathways include conserved transacting siRNA pathway. In chapter "Trans-acting small interfering RNAs: biogenesis, mode of action and role in plant development", Maizel and Colleagues describe not only the biogenesis of transacting siRNAs but also their functions in controlling leaf polarity and lateral root growth including plant development.

Several developmental and physiological events controlled by gene regulatory networks govern the process of seed development, and miRNAs are also part of such regulatory networks. Nonegaki and colleagues describe the functions of miRNAs in this important process (Chapter "Role of miRNAs in seed development"). High-throughput sequencing of small RNAs to an unprecedented depth from diverse plant species led to the discovery of several novel miRNAs. Assessing their function is one of the major challenges now. Millar and colleagues address various strategies to dissect the functions of these miRNAs in plants (Chapter "Genetic and molecular approaches to assess microRNA function"). Rice is the most important crop in the world and is a model system for monocots, especially for cereals. Helliwell and colleagues summarize the progress that has been made with respect to miRNAs' discovery, target genes that miRNAs are regulating in rice, as well as functions of some of the rice miRNAs (Chapter "Functions of miRNAs in rice").

Legumes have established symbiotic relationship with the rhizobia in the specialized structures called nodules that are associated with the roots and the process is called "nodulation." Nodulation is a highly complex process, which is governed by the spatial and temporal expression of genes and gene products. Senthil Subramanian describes the importance of miRNA-controlled gene regulation during nodulation (Chapter "microRNA regulation of symbiotic nodule development in legumes"). Plant growth and development as well as reproduction depend on availability of adequate macronutrients (N, P, K, S) and micronutrients (Cu, Fe, Zn, Mn, etc.). Plants often are challenged by the inadequate supply of these nutrients, particularly the macronutrients. Recent studies have established a key role for miRNAs in nutrient homeostasis. Julia Kehr summarizes the latest findings on this important topic (Chapter "Roles of miRNAs in nutrient signaling and homeostasis").

Being sessile organisms, plants are often challenged with abiotic (drought, salinity, cold, heavy metals, and others) and biotic (bacteria, viruses, fungi, insects, and several others) stress factors that negatively impact crop productivity. Therefore, developing crop plants with increased abiotic or biotic stress resistance using molecular breeding or biotechnological approaches are of paramount importance. However, a major challenge has been to identify key genes/proteins or other molecules that play critical roles in stress tolerance. Recent exciting research implicated an important role for miRNAs in plant stress responses. Chapters "Role of microRNAs in plant adaptation to environmental stresses" and "Endogenous small RNAs and antibacterial resistance in plants" discuss the importance of miRNA-dependent gene regulation during abiotic stresses (Sunkar and colleagues) and bacterial pathogens (Katiyar-Agarwal and colleagues), respectively. Finally, Pooggin and colleagues describe the involvement of small RNAs in plant viral resistance (Chapter "Role of virus-derived small RNAs in plant antiviral defense: insights from DNA viruses").

I sincerely thank the contributors who made this assignment possible and rewarding and Dr. Frantisek Baluska for giving me this opportunity.

Stillwater, Oklahoma Ramanjulu Sunkar

# Contents

# Small RNAs in Plants

**Zhixin Xie, Gengxiang Jia, and Arnab Ghosh**

**Abstract** Small RNAs associated with RNA silencing have emerged as an essential regulatory component in eukaryotes. Although their widespread existence was revealed only a decade ago, remarkable progress has been made toward our understanding of their biogenesis and cellular function. In plants, the small RNA-mediated regulatory mechanisms are involved in many important biological processes including developmental timing, pattern formation, epigenetic silencing of transposable elements, response to environmental stress, and defense against invading pathogens. Emerging evidence also indicates the involvement of small RNAs in epigenetic reprogramming associated with germ cell and embryo development during sexual reproduction. In this chapter, we provide an overview on the conserved molecular machinery that has evolved to give rise to microRNAs (miRNAs) and several distinct classes of small interfering RNAs (siRNAs) in plants, including heterochromatin-associated siRNAs (hc-siRNAs), *trans*-acting siRNAs (ta-siRNAs), and natural *cis*-antisense transcripts-associated siRNAs (nat-siRNAs). These are followed by a description on the cellular function and regulatory targets for each class of these endogenous small RNAs. While the focus of the book is on miRNAs, it is our hope that this chapter will serve as a brief introduction to the plant small RNA world.

## 1 Introduction

Small RNAs of 21- to 24-nucleotide (nt) in size are important sequence-specific regulators in eukaryotes. Over the past decade, studies from diverse model systems have uncovered the genomic origin, biogenesis pathway, and cellular function for

Z. Xie (✉) • G. Jia • A. Ghosh
Department of Biological Sciences, Texas Tech University, P.O. Box 43131, Lubbock, TX 79409, USA
e-mail: zhixin.xie@ttu.edu

R. Sunkar (ed.), *MicroRNAs in Plant Development and Stress Responses*, Signaling and Communication in Plants 15,
DOI 10.1007/978-3-642-27384-1_1, 

many of these fascinating small RNA molecules. Plants such as *Arabidopsis thaliana* serve as important and unique systems for deciphering the genetic and functional diversification of small RNA-directed pathways (Baulcombe 2004; Chen 2010). This chapter will begin with a brief history of small RNA discovery, which is followed by a short description on the conserved molecular machinery and biogenesis pathways for distinct classes of small RNAs. We then present the cellular function and regulatory targets for each class of these endogenous small RNAs. The purpose of this chapter is to provide an overview on the biogenesis and cellular function of endogenous small RNAs in plants, which we hope will serve as a brief introduction to the plant small RNA world.

## 2 Discovery and Classification

### *2.1 RNA Silencing and Associated Small RNAs*

Although phenomena of RNA silencing [known as cosuppression, or posttranscriptional gene silencing (PTGS)] in plants were documented in the early 1990s (Napoli et al. 1990; van der Krol et al. 1990), it took nearly another decade for the small RNAs and their link to gene silencing to be discovered. In 1999, a seminal paper from the Baulcombe laboratory reported the detection of an approximately 25-nt small RNA species that is associated with transgene- and virus-induced PTGS in plants (Hamilton and Baulcombe 1999). A few months later, a biochemical analysis in *Drosophila melanogaster* in vitro system demonstrated that double-stranded RNAs, the trigger of RNA interference (RNAi) (Fire et al. 1998), are processed into 21- to 23-nt small RNAs during RNAi (Zamore et al. 2000). These and other early findings from diverse eukaryotic model systems collectively pointed to a unifying mechanism of RNA silencing that appears to be evolutionarily conserved (Cogoni and Macino 2000; Fagard et al. 2000). The idea that RNA silencing might operate as an endogenous regulatory mechanism in the normal life of eukaryotes had prompted the search for endogenous small RNAs [reviewed in (Zamore and Haley 2005)]. The discovery of distinct classes of endogenous small RNAs and their important regulatory functions has since revealed a hidden small RNA world and opened a new era of small RNA biology (Chapman and Carrington 2007).

### *2.2 microRNAs and Other Endogenous Small RNAs in Plants*

The effort in searching for endogenous small RNAs through molecular cloning revealed an abundance of these molecules in *Arabidopsis* (Llave et al. 2002a; Park et al. 2002; Reinhart et al. 2002), shortly after the first reports on animal models

(Lagos-Quintana et al. 2001; Lau et al. 2001; Lee and Ambros 2001). A subset of these endogenous small RNAs, predominately 21-nt long, were named microRNAs (miRNAs) for their distinct biogenesis features. They arise from noncoding precursor transcripts that are capable of forming the characteristic foldback stem-loop structure—an intramolecular dsRNA structure with extensive base pairing in the "stem" portion. The primary transcripts of an miRNA (pri-miRNAs) originate from a defined genetic locus known as a *MIRNA* gene, typically located in an intergenic region (IGR). Some of the *MIRNA* loci may give rise to mature miRNAs with identical or nearly identical sequences, forming a multimember miRNA family (Jones-Rhoades et al. 2006). The genome of the reference plant *A. thaliana* contains nearly 200 *MIRNA* loci. While in many cases each of these loci gives rise to a unique mature miRNA (single-member family), there are over 20 *Arabidopsis* miRNA families with 2–14 members, each arising from a distinct locus (Rajagopalan et al. 2006; Fahlgren et al. 2007; Backman et al. 2008; Meyers et al. 2008; Kozomara and Griffiths-Jones 2010). Initial cloning and sequencing effort in multiple species have revealed other important features of plant miRNAs (Sunkar and Zhu 2004; Arazi et al. 2005; Axtell and Bartel 2005; Lu et al. 2005; Sunkar et al. 2005). For instance, while many known plant miRNAs appear to be lineage- or species-specific, some miRNAs are deeply conserved across all land plant species so far examined, reflecting an ancient origin of miRNA-based regulatory mechanism [reviewed in (Cuperus et al. 2011)]. It is worth noting that conservation of miRNAs across species is generally limited to the mature miRNA sequences, with the remaining portion of miRNA precursors showing little sequence homology. This is also true among the paralogous *MIRNA* loci in a given plant species (Jones-Rhoades et al. 2006). These observations suggest that the sequence of a mature miRNA and the stem-loop structure of its precursor may constitute two most important parameters for the functionality of a *MIRNA* locus.

Interestingly, the *Caenorhabditis elegans lin-4* locus, which was genetically identified as an essential developmental regulator that negatively regulates the level of LIN-14 protein, turned out to be the very first miRNA gene that had been functionally characterized. In their work published in 1993, Victor Ambros and colleagues showed that *lin-4* is a noncoding RNA locus which produces two size species of transcripts in vivo, a 61-nt lin-4L and a 22-nt lin-4S, respectively (Lee et al. 1993). Based on the sequence complementarity found between the lin-4 RNAs and a segment in the 3′UTR of LIN-14 mRNA, an antisense RNA–RNA interaction-mediated mechanism was proposed for the negative regulation of LIN-14 by lin-4 RNA (Lee et al. 1993). These early observations provided important clues for uncovering the regulatory mode of miRNAs in both plants and animals.

It was obvious even from the early cloning effort that many endogenous small RNAs in plants do not belong to miRNA. With the application of next-generation DNA sequencing (NGS) technologies in small RNA discovery, it becomes even more clear that miRNAs account for only a small fraction of the endogenous small RNA complexity (defined as the sum of distinct small RNA sequences), although some miRNAs may be present in extremely high abundance in certain tissue types or at specific developmental stages (Lu et al. 2005; Rajagopalan et al. 2006;

Howell et al. 2007; Kasschau et al. 2007). Those endogenous small RNAs other than miRNAs are collectively known as small interfering RNAs (siRNAs) in plants, for they are generally derived from perfectly base-paired dsRNA precursors. At least three distinct biogenesis pathways are known to give rise to endogenous siRNAs, as discussed below.

## 3 Biogenesis Pathways

### *3.1 Conserved Machinery*

The core components of the RNA silencing machinery involve several evolutionarily conserved protein families. These include the Dicer (DCR; nomenclature used in animals and the fission yeast) or Dicer-like (DCL; nomenclature used in plants and fungi), Argonaute (AGO), and RNA-dependent RNA Polymerase (RDR) protein families, among others. Identification of these proteins as key components of RNA silencing was a result of collective effort made in diverse model systems. A Drosophila DCR protein, a multidomain RNase III-like ribonuclease, for example, was first identified in an in vitro RNAi system as a key enzyme for processing dsRNA into small RNAs (Bernstein et al. 2001). This work also serves as an interesting example for successful identification of an RNAi pathway component using RNAi. A role for both the RDR and AGO family proteins in RNA silencing was identified genetically. Screens for RNA silencing-defective mutants independently identified genes encoding proteins homologous to a tomato RDR (Schiebel et al. 1998) in the filamentous fungus *Neurospora crassa* (QDE-1) (Cogoni and Macino 1999), the nematode worm *Caenorhabditis elegans* (EGO-1)(Smardon et al. 2000), and the reference plant *A. thaliana* (SDE1/SGS2, now known as RDR6) (Dalmay et al. 2000; Mourrain et al. 2000). Similar screens also led to the independent identification of genes encoding AGO family proteins in *N. crassa* (QDE-2) (Catalanotto et al. 2000), *C. elegans* (RDE-1) (Tabara et al. 1999), and *A. thaliana* (AGO1) (Fagard et al. 2000). These works provided strong genetic evidence for conservation of RNA silencing mechanisms across the kingdoms and formed the foundation for investigating the role of these conserved proteins in biogenesis and function of endogenous small RNAs.

### *3.2 Distinct Pathways*

The rich genetic resources for *Arabidopsis*, including molecular markers for powerful forward genetic screen, plus the large collections of sequence-indexed T-DNA insertion mutants (Sessions et al. 2002; Alonso et al. 2003; Rosso et al. 2003; Woody et al. 2007) for reverse genetics approach have substantially facilitated the

identification of genes involved in biogenesis and function of small RNAs in plants. For example, the *Arabidopsis* DCL1, which encodes a homolog of the Drosophila DCRs, was independently recovered from multiple genetic screens for mutants that are defective in embryo, ovule, or flower development, respectively, prior to the discovery of small RNAs (Schauer et al. 2002). The discovery of miRNAs and the availability of multiple *dcl1* mutant alleles allowed a direct examination on the role of DCL1 in miRNA biogenesis, leading to the identification of DCL1 as a key component of the miRNA pathway (Park et al. 2002; Reinhart et al. 2002). The observation that a miRNA-deficient *dcl1* mutant was capable of PTGS induced by a dsRNA-producing transgene and accumulating PTGS-associated siRNAs was indicative for more than one functional DCL proteins in *Arabidopsis*, which suggests more than one small RNA pathway operating in plants (Finnegan et al. 2003). Indeed, at least four distinct endogenous small RNA pathways have been identified in plants, each involving a subset of conserved RNA silencing components.

### 3.2.1 miRNAs

Biogenesis of miRNA begins with transcription at a *MIRNA* locus by RNA Polymerase II (Pol II), producing 5′-capped, 3′-polyadenylated pri-miRNAs which may contain introns (Aukerman and Sakai 2003; Kurihara and Watanabe 2004; Xie et al. 2005a). Known factors involved in the initial processing of a pri-miRNA in *Arabidopsis* include subunits of the nuclear cap-binding complex (CBC) CBP20 and CBP80 (Gregory et al. 2008; Laubinger et al. 2008); SERRATE (SE) (Grigg et al. 2005; Lobbes et al. 2006; Yang et al. 2006a), a C2H2-type zinc finger domain-containing protein; HYPONASTIC LEAVES 1 (HYL1) (Han et al. 2004; Vazquez et al. 2004a), a member of dsRNA-binding protein family; and DCL1 (Park et al. 2002; Reinhart et al. 2002), one of the four DCL proteins in *Arabidopsis* (Fig. 1a) [for recent reviews, see (Voinnet 2009) and (Xie et al. 2010)]. The current model proposes that CBC facilitates miRNA biogenesis, likely through direct interaction with the 5′ cap of a nascent pri-miRNA, consistent with reduced accumulation of mature miRNAs in *cpb20* and *cbp80* mutants (Gregory et al. 2008; Kim et al. 2008; Laubinger et al. 2008). DAWDLE (DDL), a forkhead-associated domain-containing protein (Morris et al. 2006), also plays a role in stabilizing the foldback structure of a pri-miRNA and in the recruitment of DCL1 (Yu et al. 2008). SE also facilitates miRNA biogenesis, likely through its direct interaction in the nucleus with HYL1 and DCL (Fang and Spector 2007; Fujioka et al. 2007; Song et al. 2007). Interestingly, SE, as well as CPB20 and CPB80, appear to also play a more general role in intron splicing for both pre-mRNA and pri-miRNA (Laubinger et al. 2008). One possibility is that SE may serve as a common bridging factor not only between CBC and spliceosome components but also between CBC and DCL1 or HYL1, as has been proposed (Laubinger et al. 2008), although a direct interaction between CBC and SE has yet to be demonstrated. A key feature of a pri-miRNA is that they are self-complementary and capable of forming the characteristic foldback

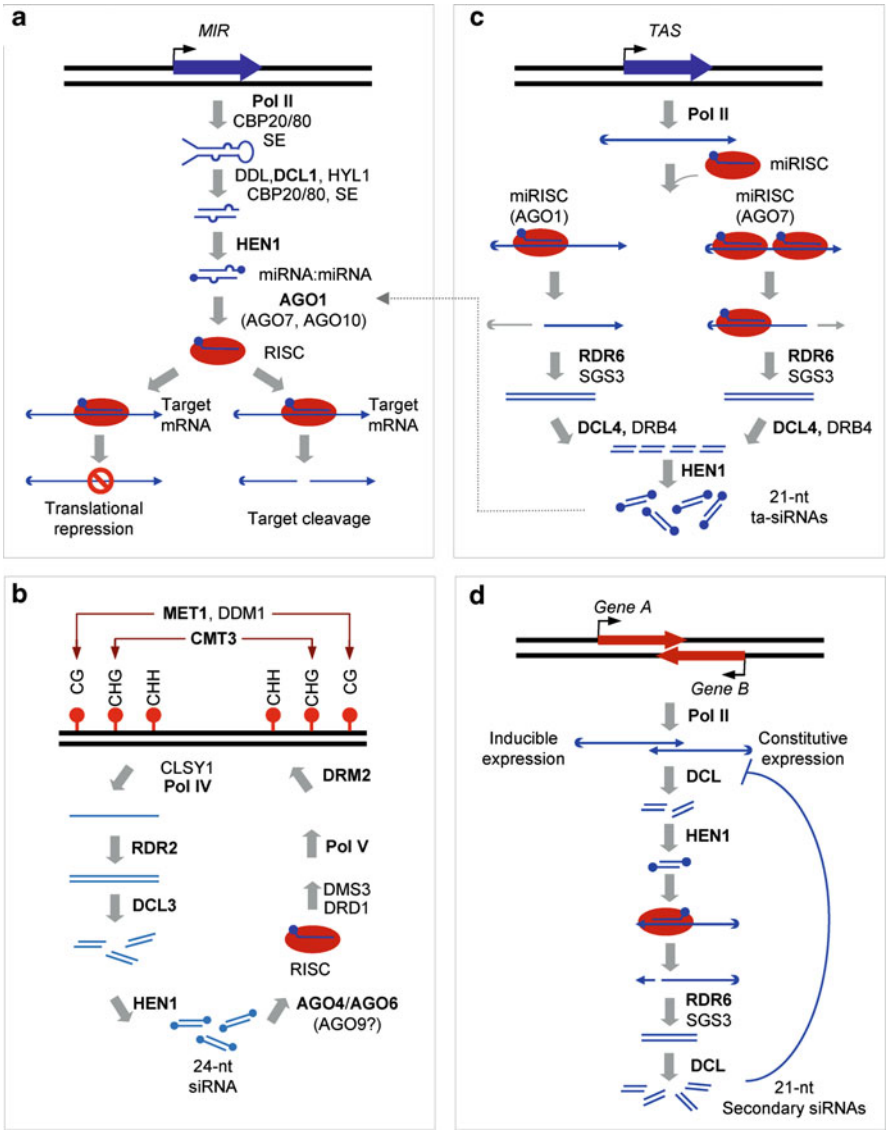


**Fig. 1** Pathways for small RNA biogenesis and function in plants. (**a**) microRNA (miRNA) biogenesis and miRNA-directed posttranscriptional regulation of gene expression through target cleavage or translational repression. (**b**) Heterochromatin-associated small interfering RNA (hc-siRNA) and RNA-directed DNA methylation (RdDM). (**c**) Trans-acting siRNA biogenesis and ta-siRNA-directed posttranscriptional regulation of gene expression. (**d**) A pathway for biogenesis and function of natural *cis*-antisense transcripts-associated siRNAs (nat-siRNAs). Components with a catalytic activity in each pathway are indicated by a *bold typeset*

hairpin-like structure recognized by DCL1. HYL1, as well as SE, are thought to stabilize the foldback structure of a pri-miRNA and facilitate accurate processing by DCL1 (Kurihara et al. 2006; Dong et al. 2008; Tagami et al. 2009). Processing of a pri-miRNA by the DCL1 complex ultimately releases a small RNA duplex consisting of a miRNA and its passenger strand (termed miRNA*), with a 2-nt 3′ overhang in each strand (Fig. 1a). This DCL1-catalyzed excision of a miRNA: miRNA* duplex involves at least two cleavage events which typically follow a "base-to-loop" processing mode, with the first cut occurring at a loop-distal site, approximately 15 bases away from the end of stem (Cuperus et al. 2009; Mateos et al. 2009; Song et al. 2009; Werner et al. 2009), releasing a partially processed stem-loop precursor known as pre-miRNA. However, for pri-miRNAs with a longer foldback structure (e.g., miR159 and miR319), DCL1 appears to make more than two cuts in a noncanonical "loop-to-base" processing mode (Addo-Quaye et al. 2009; Bologna et al. 2009). HUA ENHANCER1 (HEN1), a small RNA methyltransferase, recognizes small RNA duplexes resulting from DCL processing and deposits a methyl group at the 2′-OH position of the 3′ terminal ribose in each small RNA (Fig. 1a) (Park et al. 2002; Yu et al. 2005; Yang et al. 2006b). This 2′-O-methyl group is believed to stabilize miRNAs in vivo by protecting them from exonucleolytic attack or from alternative end modification such as 3′-uridylation which has been observed in *hen1* mutants (Li et al. 2005).

### 3.2.2 Heterochromatin-Associated siRNAs

Sequencing of *Arabidopsis* small RNA libraries revealed endogenous small RNA populations arising through mechanisms that differ from that of miRNAs. One class of the endogenous small RNAs are predominantly 24-nt and associated with sequences derived from transposable elements (TE), highly repetitive rDNA regions, and other uncharacterized intergenic sequences (Tang et al. 2003; Xie et al. 2004). These longer species of TE-derived small RNAs were previously observed in blot-based assays from tobacco and *Arabidopsis*, and appeared to correlate with DNA methylation at the corresponding genomic loci (Hamilton et al. 2002). Genetic analysis in *Arabidopsis* using T-DNA insertion mutants identified DCL3 and RDR2 as key components for the biogenesis of 24-nt small RNAs, suggesting a genetically distinct pathway that involves an RDR2-dependent dsRNA precursor (Xie et al. 2004). A nomenclature of siRNA is therefore justified. AGO4, an *Arabidopsis* AGO family member which was genetically identified as a suppressor for epigenetic silencing of the *Superman* (*SUP*) locus (Zilberman et al. 2003), turned out to be another key component of the RDR2- and DCL3-dependent small RNA pathway. RDR2, DCL3, and AGO4 were shown to be required for accumulation of 24-nt siRNAs derived from several endogenous loci typically associated with TEs and other repetitive sequences including the 5S rDNA array (Zilberman et al. 2003, 2004; Xie et al. 2004). Loss-of-function mutations in these RNA silencing components led to concomitant loss of 24-nt siRNA accumulation, epigenetic marks characteristic of heterochromatin [e.g., cytosine methylation in

DNA, histone H3 dimethylation at lysine 9 (H3K9)], and release of transcriptional silencing at specific genomic loci, suggesting a role for 24-nt siRNAs in directing chromatin silencing in a sequence-specific manner (Zilberman et al. 2003; Chan et al. 2004; Xie et al. 2004; Zilberman et al. 2004). The 24-nt siRNAs were therefore considered as heterochromatin-associated siRNAs (hc-siRNAs) that function in RNA-directed DNA methylation (RdDM) (Fig. 1b). These observations supported the earlier speculations that the 24-nt siRNAs may act as sequence determinant in maintenance of TE silencing and genome integrity in plants (Hamilton et al. 2002).

Additional components required for hc-siRNA biogenesis and function have been identified genetically. Besides the three nuclear RNA polymerases (Pol I, Pol II, and Pol III) that are common to eukaryotes, plants possess two additional RNA polymerases (termed Pol IV and Pol V; formerly Pol IVa and Pol IVb) which are most closely related to Pol II and specifically act in the chromatin silencing pathway (Fig. 1b). Loss-of-function mutations in NRPD1 (formerly NRPD1a), the largest subunit of Pol IV, or NRPD2 (formerly NRPD2a), the second largest subunit common to Pol IV and PolV, largely eliminated the hc-siRNA accumulation and impaired chromatin silencing which also requires RDR2, DCL3, and AGO4 (Herr et al. 2005; Kanno et al. 2005b; Onodera et al. 2005; Pontier et al. 2005). Immunofluorescence-based localization studies showed that NRPD1 signals were detected throughout the nucleoplasm as punctate foci, but were absent from the nucleolus (Pontes et al. 2006). Pairwise detection of fluorescence signals for RDR2, DCL3, AGO4, and NRPE1 (formerly NRPD1b, the largest subunit of Pol V) indicated that these proteins, along with hc-siRNAs, colocalize in nucleolus-associated bodies, although punctate signals for each of these proteins were also seen in the nucleoplasm (Li et al. 2006; Pontes et al. 2006). These AGO4-containing nucleolar bodies, which were initially thought to overlap with Cajal bodies where maturation of multiple ribonucleoprotein complexes takes place, were designated as nuclear RNA processing center (Li et al. 2006; Pontes et al. 2006). A later study showed that AGO4 protein can be found in two distinct types of nuclear bodies: the Cajal body and the AGO4-NRPE1 body (also known as the AB body) which also contains NRPD2 and the de novo DNA methyltransferase DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) (Li et al. 2008). The fact that loss-of-function mutations in one component may induce mislocalization of downstream components in the same pathway allowed Pikaard and colleagues to establish that Pol IV acts upstream of RDR2 in hc-siRNA biogenesis (Pontes et al. 2006). CLASSY1 (CLSY1), an SNF2-domain-containing protein that exhibited a nuclear localization pattern highly similar to that of RDR2, is also involved in hc-siRNA biogenesis, likely acting at a step upstream of RDR2 (Smith et al. 2007). Consistent with a model in which Pol IV gives rise to single stranded transcripts for RDR2- and DCL3-dependent processing into hc-siRNAs (Fig. 1b), genome-wide small RNA profiling revealed that Pol IV activity is required for biogenesis of a vast majority of the endogenous 24-nt siRNAs in *Arabidopsis* (Zhang et al. 2007; Mosher et al. 2008). Examples for Pol IV-independent 24-nt siRNAs include those derived from genomic loci harboring large inverted repeats, from which presumed Pol II

transcription may give rise to transcripts capable of forming long dsRNA precursors accessible for multiple DCL processing (Zhang et al. 2007), as has been shown in a transgenic inverted repeat (IR) locus (Kanno et al. 2005b). Unlike Pol IV, Pol V is required for hc-siRNA accumulation at only a subset of Pol IV-dependent loci (Mosher et al. 2008) and is thought to function at a downstream step to reinforce hc-siRNA biogenesis in a genomic context-dependent manner (Pontier et al. 2005). Two nonmutually exclusive potential mechanisms exist for a role of Pol V in directing hc-siRNA–AGO4 complexes to a target locus during RdDM. First, the carboxyl-terminal domain (CTD) of NRPE1 contains reiterated WG/GW motifs that were shown to confer a direct protein–protein interaction with AGO4 (Li et al. 2006). Second, Pol V transcription generates intergenic noncoding (IGN) transcripts at certain RdDM target loci (Wierzbicki et al. 2008). Experimental evidence supports the idea that AGO4 may be guided to target loci through base pairing between its associated siRNA and the nascent Pol V transcripts (Wierzbicki et al. 2009). Two other proteins, defective in RNA-directed DNA methylation1 (DRD1) which is a putative SNF2 domain-containing chromatin remodeling factor (Kanno et al. 2004, 2005a) and defective in meristem silencing3 (DMS3) which encodes a protein similar to the hinge-domain region of structural maintenance of chromosome (SMC) proteins (Kanno et al. 2008), are required for initiation of Pol V transcription (Wierzbicki et al. 2008). Interestingly, Pol II has also been implicated in RdDM at certain endogenous loci, presumably by generating intergenic noncoding transcripts that may serve as scaffolds for recruitment of RdDM factors such as AGO4 (Zheng et al. 2009). Current models for RdDM involve Pol IV transcription to generate a single-stranded transcript which is acted upon by RDR2 and DCL3 to produce hc-siRNAs. The formation of hc-siRNA–AGO4 complex and its subsequent interaction with either NRPE1 or Pol V-generated transcripts allows assembly of the RdDM effector complex at the target loci, which ultimately recruits DRM2 to the scene (Fig. 1b) (Law and Jacobsen 2010).

### 3.2.3 *Trans*-acting siRNAs

Analysis of small RNAs from *Arabidopsis* identified yet another class of endogenous small RNAs, now known as *trans*-acting siRNAs (ta-siRNAs) which are generated through a distinct, miRNA-dependent mechanism. Initial analysis of these predominantly 21-nt small RNAs revealed that their accumulation requires all known factors involved in miRNA biogenesis, but their putative precursor transcripts do not seem to have the potential of adopting a hairpin-like structure, suggesting a biogenesis pathway that differs from that of miRNA (Peragine et al. 2004; Vazquez et al. 2004b). In addition, accumulation of these small RNAs also requires RDR6 and suppressor of gene silencing 3 (SGS3) (Peragine et al. 2004; Vazquez et al. 2004b), two factors previously identified as components of PTGS (Dalmay et al. 2000; Mourrain et al. 2000). Important clues for uncovering the puzzling ta-siRNA biogenesis came from two additional observations. First, the ta-siRNA precursor transcripts, which arise from Pol II transcription (Vazquez et al. 2004b)

at defined noncoding RNA loci, are bona fide miRNA targets and subject to miRNA-directed cleavage (Fig. 1c) (Allen et al. 2005). The miRNA-directed cleavage appears to serve two functions: it stimulates RDR6-dependent conversion of the cleaved transcript into a dsRNA and sets the register for subsequent DCL processing of the resulting dsRNA into a phased array of 21-nt siRNAs (Fig. 1c). Second, loss-of-function mutations in DCL4 lead to mild developmental phenotypes reminiscent of *sgs3* and *rdr6* mutants, suggesting the involvement of DCL4 in ta-siRNA biogenesis. Indeed, DCL4 turned out to be required for ta-siRNA accumulation, presumably acting on the RDR6-generated dsRNA substrates to yield a phased array of 21-nt siRNAs (Fig. 1c) (Gasciolli et al. 2005; Xie et al. 2005b; Yoshikawa et al. 2005). DRB4, another member of the dsRNA-binding protein family of which HYL1 is the founding member, also plays a role in ta-siRNA biogenesis (Fig. 1c) (Adenot et al. 2006). DRB4 is thought to facilitate DCL4 processing of a dsRNA substrate by directly binding to the dsRNA and specifically interacting with DCL4 (Nakazawa et al. 2007; Fukudome et al. 2011). As expected, genetic lesions in *Arabidopsis* DRB4 also result in mild developmental perturbations as seen in the *Arabidopsis sgs3*, *rdr6*, and *dcl4* mutants (Adenot et al. 2006; Nakazawa et al. 2007). This miRNA targeting-initiated, SGS3-, RDR6-, DRB4-, and DCL4-dependent ta-siRNA biogenesis mechanism therefore defines an endogenous pathway for efficient secondary small RNA production that involves an RDR-mediated amplification step. These siRNAs appear to be selectively recruited by AGO1 to direct cleavage of their target mRNAs in a mechanism that is otherwise indistinguishable from the miRNA-directed target cleavage. The term ta-siRNA was so coined because a ta-siRNA and its target RNA arise from distinct genetic loci (a ta-siRNA therefore acts in *trans*) and share little sequence similarity outside the siRNA–target interacting site.

A total of eight *Trans-acting small interfering RNA* (*TAS*) loci belonging to four families have been identified in the *Arabidopsis* genome. The miRNAs that trigger ta-siRNA biogenesis from these loci include miR173 (for *TAS1a, b, c;* and *TAS2*), miR390 (for *TAS3a, b, c*), and miR828 (for *TAS4*) (Peragine et al. 2004; Vazquez et al. 2004b; Allen et al. 2005; Yoshikawa et al. 2005; Rajagopalan et al. 2006). However, ta-siRNAs-like siRNA biogenesis and regulatory mechanisms are not limited to these loci. Genome-wide analysis of small RNAs in *Arabidopsis* revealed over a dozen of *TAS*-like loci that give rise to RDR6-dependent, phased siRNAs (Howell et al. 2007). They differ from the canonical *TAS* loci in that they are protein-coding genes and that a majority of them encode pentatricopeptide repeat (PPR) proteins.

The discovery of ta-siRNA biogenesis pathway had elicited several intriguing questions. Although miRNA-directed target cleavage is common in plants, the RDR6-dependent generation of secondary siRNAs derived from the cleaved target appears to be the exception rather than the rule. What, then, are the molecular determinants that channel a cleaved RNA fragment into the RDR6-dependent pathway? Answers to this question would provide mechanistic insights into PTGS of transgenes which also involves RDR6 (Dalmay et al. 2000; Mourrain et al. 2000; Luo and Chen 2007), because the long speculated “aberrant” RNA, a

hypothetical trigger of PTGS, has yet to be defined. Moreover, miRNA-directed cleavage on a ta-siRNA precursor transcript yields two fragments, selection of which for RDR6-dependent processing varies among ta-siRNA-generating loci (Fig. 1c). What molecular features dictate the 5′ or 3′ cleaved fragment to be selected for ta-siRNA production? Works addressing these questions have uncovered intricate molecular mechanisms that govern RDR6-dependent ta-siRNA biogenesis following the initial miRNA-directed cleavage of a *TAS* transcript. The ta-siRNA biogenesis from *TAS1* and *TAS2* transcripts triggered by miR173-directed cleavage represents a mechanism by which the 3′ fragment (downstream from the cleavage site) is specifically selected as the RDR6 substrate for entry into the pathway. The interaction between miR173–AGO1 complex and a *TAS* transcript which results in a cleavage on the *TAS* transcript was shown to be specifically required, as mutations that render a noncleavage interaction or a cleavage directed by several other miRNAs failed to initiate ta-siRNA biogenesis from a *TAS1* transcript (Montgomery et al. 2008b; Felippes and Weigel 2009). Curiously, genome-wide surveys for 21-nt siRNA generation from miRNA-targeted transcripts in Arabidopsis and rice (*Oryza sativa*) revealed a surprisingly unique feature that is associated with a miRNA trigger for RDR6-dependent siRNA biogenesis: a 22-nt miRNA (Chen et al. 2010; Cuperus et al. 2010). While mature miRNAs in plants are typically 21-nt in length, canonical processing of certain miRNA foldback precursors that contain an asymmetric, single-nucleotide bulge in the miRNA arm yields 22-nt mature miRNAs, which is the case for miR173, miR828, and several other miRNAs in *Arabidopsis* (Chen et al. 2010; Cuperus et al. 2010). Rigorous experimental analysis on siRNA generation in response to manipulation of trigger miRNAs confirmed the unique functionality of 22-nt miRNAs in initiating RDR6-dependent secondary siRNA biogenesis (Chen et al. 2010; Cuperus et al. 2010). Exactly how a 22-nt miRNA–AGO1 complex specifically routes the cleaved 3′ target fragment into the RDR6-dependent pathway remains a mystery.

The ta-siRNA biogenesis from *TAS3* transcripts triggered by miR390-directed cleavage, on the other hand, represents a different mechanism involving the selected entry of a 5′ fragment (upstream from the cleavage site) into the RDR6-dependent pathway (Fig. 1c). Intriguingly, miR390 which is a 21-nt miRNA with a 5′ terminal adenosine (5′-A) is preferentially recruited to AGO7 (Montgomery et al. 2008a). A closer inspection of *TAS3* transcripts identified a second, noncleavable miR390-interacting site proximal to the 5′ end of the transcript (Axtell et al. 2006). Both the 5′-noncleavable and the 3′-cleavable miR390 target sites are required for ta-siRNA biogenesis (Fig. 1c). Curiously, while the miR390-directed cleavage at the 3′ proximal target site may be functionally mimicked by a different miRNA–AGO complex, the noncleavage interaction between the miR390–AGO7 complex and a *TAS3* transcript at the 5′ proximal target site appears to be specifically required for ta-siRNA biogenesis (Axtell et al. 2006; Montgomery et al. 2008a). It is possible that the noncleavage interaction between miR390–AGO7 complex and *TAS3* transcript may help stabilize the 5′ fragment following the 3′ proximal cleavage event, thereby facilitating the selected entry of the 5′ fragment into the RDR6-dependent pathway (Fig. 1c), although the specific requirement for miR390–AGO7 complex is currently not understood.

### 3.2.4 Natural *cis*-Antisense Transcripts-Associated siRNAs

Natural *cis*-antisense transcripts (*cis*-NATs) refer to transcripts arising from two neighboring protein-coding genes that occupy partially overlapping genomic loci, in either convergent or divergent configuration. When present in the same cell, pairs of *cis*-NATs have the potential to form dsRNAs that can be processed into siRNAs (*nat-siRNAs*) by a DCL activity (Fig. 1d). A few nat-siRNAs in *Arabidopsis* have been functionally analyzed so far (Borsani et al. 2005; Katiyar-Agarwal et al. 2006; Ron et al. 2010), although hundreds of potential nat-siRNA-generating loci exist in the genomes of *Arabidopsis* and rice (German et al. 2008; Jin et al. 2008; Zhou et al. 2009). Based on a genome-wide computational analysis, a vast majority of the identifiable potential nat-siRNA-generating loci (over 900) in *Arabidopsis* involve *cis*-NATs from two convergent neighboring genes (Jin et al. 2008).

Several common features for nat-siRNA biogenesis can be drawn from the few case studies in *Arabidopsis*. First, one of the two neighboring genes at a nat-siRNA locus may be constitutively transcribed whereas expression of the other may be inducible in response to environmental or developmental cues. Initiation of nat-siRNA biogenesis may therefore occur only upon certain environmental stimuli or in specific cell types (Fig. 1d). This may account for the scarcity of nat-siRNAs in certain biological samples that are prepared either from plants grown under normal growth conditions or from tissues in which a specific cell type is poorly represented. In the two earlier reports, for example, the specific nat-siRNAs were shown to be detectable only upon high-salinity stress (Borsani et al. 2005) or infection by a bacterial pathogen (Katiyar-Agarwal et al. 2006), respectively. Secondly, nat-siRNA biogenesis appears to involve an RDR-mediated amplification step followed by secondary siRNA production (Fig. 1d). Based on the two reports in which accumulation of locus-specific nat-siRNAs was examined in a panel of small RNA-deficient mutants, RDR6 and SGS3 are required for secondary siRNA generation at the nat-siRNA loci, likely involving steps that are mechanistically similar to ta-siRNA biogenesis (Borsani et al. 2005; Katiyar-Agarwal et al. 2006). The dicing activity presumably involved in the initial processing of the dsRNA may be locus-dependent because accumulation of the primary nat-siRNA species was shown to be dependent on DCL2 in one case (Borsani et al. 2005) and on DCL1 in another case (Katiyar-Agarwal et al. 2006). Curiously, nat-siRNA biogenesis also appears to involve Pol IV because a loss-of-function mutation in NRPD1 abolished nat-siRNA accumulation (Borsani et al. 2005; Katiyar-Agarwal et al. 2006). The nature of nat-siRNA biogenesis at the third reported locus is less clear because the low abundance of the nat-siRNAs, presumably due to the sperm-specific expression pattern, prevented a direct genetic dissection (Ron et al. 2010).

## 4 Regulatory Functions

Small RNAs generally function as negative, sequence-specific regulators for gene expression, at either transcription or posttranscription level. They are known to function through formation of effector complexes in which an AGO protein is the core component. A small RNA-loaded, AGO-containing multiprotein complex is termed RNA-induced silencing complex (RISC), which was first described in Drosophila (Hammond et al. 2000). Certain AGO family members, such as the mammalian AGO2 (Liu et al. 2004a), possess a catalytically active PIWI domain and are able to direct target cleavage, an endonucleolytic activity that has been dubbed as a "slicer". Depending on the slicer activity of a specific AGO to which a small RNA is incorporated, the outcome of an interaction between a small RNA and its target may or may not involve target cleavage. The genome of *Arabidopsis* contains genes for ten putative AGO proteins which appear to form three distinct phylogenetic clades (Vaucheret 2008). A slicer activity for AGO1, AGO4, AGO7, and AGO10 has been shown (Baumberger and Baulcombe 2005; Qi et al. 2005, 2006; Montgomery et al. 2008a; Ji et al. 2011). Although the biochemical nature of the plant "holo RISCs" remains to be elucidated, genetic and RNA immunoprecipitation (RIP)-based studies, coupled with NGS technologies, have been informative in uncovering the molecular basis underlying the regulatory function of distinct classes of small RNAs in plants.

### *4.1 miRNA- and ta-siRNA-Directed Gene Regulation*

RIP analysis of AGO1-associated small RNAs in *Arabidopsis* by recovered predominantly miRNAs and ta-siRNAs which are typically 21-nt in size (Baumberger and Baulcombe 2005; Qi et al. 2005). This preferred association is attributable to the preferential recruitment by AGO1 for small RNAs with a 5′ terminal uridine (5′-U) (Mi et al. 2008; Montgomery et al. 2008a; Takeda et al. 2008), which is a characteristic feature for most miRNAs. RIP analysis has also revealed preferred small RNA association for several other AGO family proteins. For example, AGO2 was shown to preferentially recruit small RNAs with a 5′-A, which include several miRNA and miRNA* species, as well as some ta-siRNAs (Mi et al. 2008; Montgomery et al. 2008a; Takeda et al. 2008). However, the 5′-terminal nucleotide identity is obviously not the sole determinant in the formation of distinct small RNA–AGO complexes. The recently reported specific interactions for *Arabidopsis* miR390 (a miRNA with a 5′-A) with AGO7 and miR165/166 (miRNAs with a 5′-U) with AGO10, respectively, clearly indicate the involvement of additional structural features that can affect RISC assembly (Montgomery et al. 2008a; Zhu et al. 2011).

In general, plant miRNAs exert their regulatory functions through extensive, near-perfect base-paring with their target RNAs (Rhoades et al. 2002), typically

leading to cleavage of targets at the middle of the base-paired region (Fig. 1a) (Llave et al. 2002b; Kasschau et al. 2003). These observations are consistent with the overall preferential association of miRNAs with AGO1, for which a slicer activity has been demonstrated (Baumberger and Baulcombe 2005; Qi et al. 2005). Several reports have suggested that miRNA-mediated regulation in plants also involves translational repression (Fig. 1a) (Aukerman and Sakai 2003; Chen 2004; Schwab et al. 2005; Gandikota et al. 2007; Brodersen et al. 2008; Lanet et al. 2009; Beauclair et al. 2010), a mode of action that is prevalent for animal miRNAs. These observations raised the interesting possibility that at least some of the plant miRNAs may have adopted two alternative modes of action since a target cleavage mode has been previously demonstrated for the miRNAs implicated in translational repression. It will be interesting to find out if the two modes of action for a miRNA is mediated by distinct AGO proteins. A third type of miRNA:target interaction involves a noncleavage interaction between a miRNA and an endogenous noncoding transcript that appears to serve as a miRNA decoy, a phenomenon termed target mimicry which has been implicated in miR399 homeostasis in *Arabidopsis* (Franco-Zorrilla et al. 2007). It is currently unclear if similar mechanisms are involved in attaining proper homeostasis for other plant miRNAs. The known targets of plant miRNAs include mRNAs for diverse groups of developmentally important proteins, such as transcription factors, components of hormone signaling pathways, and enzymes involved in nutrient assimilation [reviewed in (Jones-Rhoades et al. 2006)]. Intriguingly, the mRNAs for DCL1 and AGO1, key components of the miRNA pathway, are also subject to miRNA-directed posttranscriptional regulation in *Arabidopsis* (Xie et al. 2003; Vaucheret et al. 2004), suggesting negative feedback regulatory loops in miRNA biogenesis and function. The importance of this feedback regulatory mechanism is further supported by evidence for its conserved operation in basal plant lineages including the moss *Physcomitrella patens* (Axtell et al. 2007).

The essential regulatory role for miRNAs in plant biology is well supported by multiple lines of evidence, including the deeply conserved nature of several miRNA:target regulatory pairs (Cuperus et al. 2011). Among the most convincing evidence may be the embryonic lethality associated with null mutations in the *Arabidopsis* DCL1 and AGO1, as well as cases in which perturbation in expression of a single miRNA leads to severe developmental defects [reviewed in (Jones-Rhoades et al. 2006), and (Chen 2010)]. The later chapters of this book will provide a good sample of interesting cases that illustrate the regulatory role of specific miRNAs and ta-siRNAs.

## *4.2 nat-siRNA-Directed Gene Regulation*

Current knowledge on the regulatory role of nat-siRNAs has been limited to the few functionally characterized loci. In the case of a high-salinity-responsive nat-siRNA locus, a constitutively expressed transcript of $\Delta^1$-PYRROLINE-5-CARBOXYLATE

DEHYDROGENASE (P5CDH, At5g62530) and a salt-inducible transcript of SIMILAR TO RCD ONE 5 (SRO5, At5g62520) form the *cis*-NATs. Upon salt stress, the induced expression of *SRO5* triggers the initiation of *SRO5-P5CDH* nat-siRNA biogenesis. A primary *SRO5-P5CDH* nat-siRNA is thought to target the *P5CDH* transcripts for cleavage, which in turn triggers RDR6- and DCL1-dependent secondary siRNA formation from the cleaved *P5CDH* transcript (Borsani et al. 2005). The downregulation of P5CDH, which is likely reinforced by the RDR6-mediated secondary siRNA generation, was shown to promote proline accumulation and salt tolerance, a process also involving reactive oxygen species (ROS)-mediated signaling (Borsani et al. 2005). In the case of a bacterial pathogen-responsive nat-siRNA locus, the *cis*-NATs consist of a transcript for RAB2-LIKE small GTP-binding protein (ATGB2, At4g35860), which is highly inducible upon infection by *Pseudomonas syringae* pathovar tomato (*Pst*) carrying *avrRpt2* and a transcript for PENTATRICOPEPTIDE REPEAT-LIKE (PPRL, At4g35850). The induced expression of *ATGB2* upon pathogen challenge is thought to trigger the DCL1-dependent biogenesis of ATGB2 nat-siRNA, which was shown to downregulate the expression of PPRL, a putative negative regulator of a signaling pathway in disease resistance (Katiyar-Agarwal et al. 2006). These data suggest that nat-siRNA-mediated regulation could conceivably function as an adaptive protection mechanism in coping with biotic or abiotic stress in plants.

A report on a sperm-specific nat-siRNA-generating locus involves transcripts for KOKOPELLI (KPL, At5g63720) and ARIADNE14 (ARI14, At5g63730), respectively. Repression of ARI14 (a putative ubiquitin E3 ligase) expression in sperm by the *KPL-ARI14* nat-siRNA, which is thought to form specifically in sperm, was shown to be required for proper male gametophyte formation, as evidenced by the defective double fertilization observed in *kpl* mutants that showed elevated levels of *ARI14* mRNA (Ron et al. 2010). Although formation of the putative *KPL-ARI14* nat-siRNA has not been shown in wild-type *Arabidopsis*, transgenic coexpression of *KPL* and *ARI14* did result in the accumulation of 21-nt *KPL-ARI14* nat-siRNA species as well as cleaved ARI14 transcripts (Ron et al. 2010). These data suggest the involvement of nat-siRNA-mediated regulation in fundamental developmental processes in plants.

## *4.3 Hc-siRNA-Directed Epigenetic Regulation and Genome Defense*

Genetic identification of the *Arabidopsis* AGO4 as a suppressor of epigenetic silencing, as well as a component of the Pol IV-, RDR2-, and DCL3-dependent hc-siRNA pathway suggests that the 24-nt hc-siRNAs function in RdDM through formation of AGO4-containing effector complexes. Consistent with this idea, RIP analysis revealed preferential association of 24-nt siRNAs with AGO4 (Qi et al. 2006). As expected, siRNAs derived from TEs and other genomic repetitive

sequences were highly enriched in the AGO4-associated small RNA population (Qi et al. 2006; Mi et al. 2008). Interestingly, AGO4 appears to preferentially recruit 24-nt siRNAs with a 5′-A (Mi et al. 2008), which may account for an earlier observation that 5′-A small RNAs predominate the cellular 24-nt small RNA population (Tang et al. 2003). As mentioned earlier, involvement of AGO4 in RdDM was further supported by immunofluorescence-based localization studies in which AGO4 was shown to colocalize with NRPE1, NRPD2, and DRM2 in distinct nuclear bodies (Li et al. 2008). AGO6, another AGO family protein in *Arabidopsis* was first identified as a suppressor of transcriptional gene silencing (TGS) in a *ros1* mutant background (Zheng et al. 2007). ROS1 (for repressor of transcriptional gene silencing) is a 5-methylcytosine DNA glycosylase/lyase which catalyzes DNA demethylation through base excision. Molecular analysis revealed a role for AGO6 that is partially redundant with AGO4 in 24-nt siRNA accumulation and RdDM at multiple endogenous loci (Zheng et al. 2007). Consistent with a role for AGO6 in RdDM, immunofluorescence-based assay with a C-terminal MYC-tagged AGO6, as well as YFP-AGO6 fusion-based analysis indicated nuclear localization for AGO6 (Zheng et al. 2007). The functional similarity between AGO4 and AGO6 is not surprising, as a phylogenetic analysis for the ten putative *Arabidopsis* AGO proteins has clearly placed AGO4, AGO6, AGO9, and AGO8 [which is most likely a pseudogene (Takeda et al. 2008)] to the same clade (also known as the AGO4 clade) (Vaucheret 2008). Although an AGO9 function has not been reported from forward genetic screens in *Arabidopsis*, analysis of T-DNA insertion mutants has revealed a regulatory role of AGO9 in cell fate determination during female gametogenesis (Olmedo-Monfil et al. 2010). Interestingly, a genetic screen for mutants with apomictic development in maize identified a putative ortholog of *Arabidopsis* AGO9 (Singh et al. 2011), suggesting a likely conserved AGO9 function in germ cell fate determination. Like AGO4, RIP analysis for AGO6- and AGO9-associated small RNAs in *Arabidopsis* also revealed a remarkable enrichment for 24-nt siRNAs derived from TEs and other genomic repetitive sequences, mostly with a 5′-A (Havecker et al. 2010; Olmedo-Monfil et al. 2010).

Little is known regarding the possible mechanistic differences among hc-siRNA functions mediated through AGO4, AGO6, or AGO9. Current data suggest that hc-siRNA association with a specific AGO4 clade protein may be affected by the expression domain of an AGO protein. *AGO6* promoter: β-glucuronidase (GUS) fusion-based analyses indicated that AGO6 expression is primarily restricted in shoot and root meristems (Zheng et al. 2007; Havecker et al. 2010), which contrasts with the widely expressed pattern of AGO4 (Fig. 2a) (Havecker et al. 2010; Mallory and Vaucheret 2010). Interestingly, among the *defective in meristem silencing* (*dms*) mutants that were identified from a genetic screen involving a meristem-specific silencing system, four of the *dms* mutations have been recently mapped to *AGO6* by whole genome sequencing (Eun et al. 2011), further supporting a meristem-specific function for AGO6 and its associated siRNAs. AGO9, on the other hand, exhibits a highly localized expression pattern in developing ovules, as has been shown in both *Arabidopsis* and maize (Fig. 2a) (Havecker et al. 2010; Olmedo-Monfil et al. 2010; Singh et al. 2011). The developing ovules of

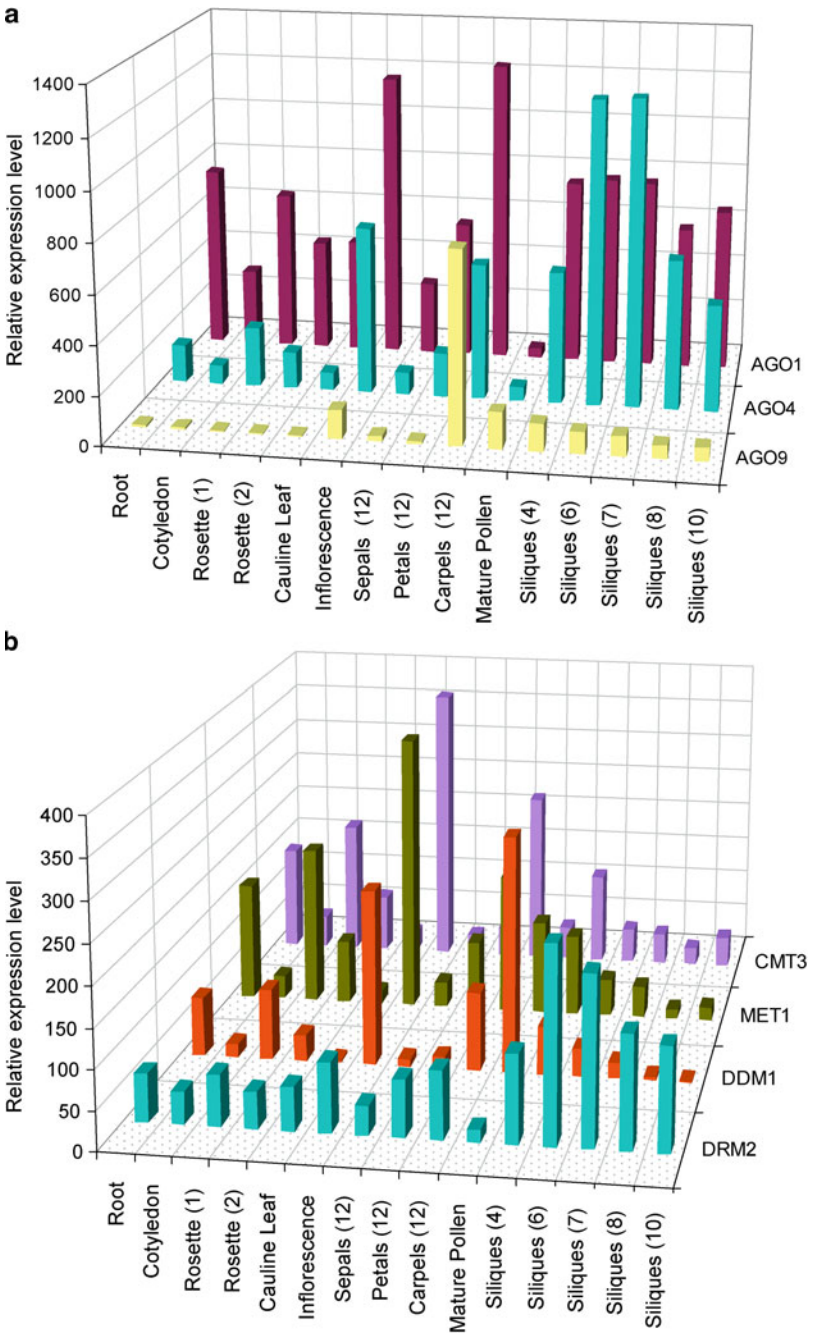


**Fig. 2** Spatial and temporal expression patterns for selected *Arabidopsis* genes based on transcriptome data. (**a**) A developmental expression pattern for *AGO1*, *AGO4*, and *AGO9*. While AGO1 functions in miRNA- and ta-siRNA-directed pathways for posttranscriptional

*Arabidopsis ago9* mutant were found to produce abnormal gametic cells in addition to the functional megaspores, suggesting a role for AGO9 and its associated siRNAs in restricting the specification of gametophyte precursor cells (Olmedo-Monfil et al. 2010). Conceivably, the spatial and temporal expression patterns of siRNA loci may also contribute to their differential incorporation into distinct AGO complexes, as indicated by a recent study involving *AGO* promoter swap coupled with RIP analysis (Havecker et al. 2010).

A role for hc-siRNAs in chromatin silencing and maintenance of genome integrity has been well established genetically, although the biochemical details remain to be elucidated. Loss-of-function mutations that affect hc-siRNA biogenesis or function are often associated with loss of repressive heterochromatic marks and releasing of TGS. Known targets of hc-siRNA-mediated silencing in *Arabidopsis* include loci representing all three types of TEs, the 5S rDNA array, as well as several uncharacterized intergenic regions. A few protein-coding genes including *FLOWERING WAGENINGEN* (*FWA*, At4g25530) which encodes a homeodomain-containing transcription factor (Kinoshita et al. 2003; Lippman et al. 2004; Chan et al. 2006), *FLOWERING LOCUS C* (*FLC*, At5g10140) which encodes a MADS-box transcription factor (Liu et al. 2004b; Swiezewski et al. 2007; Crevillen and Dean 2011), and a gene encoding the ribosomal protein RPL18 (At5g27850) (Huettel et al. 2006) have also been identified as targets of hc-siRNA-mediated regulation in *Arabidopsis*. As TEs and genomic repetitive sequences are also known to be epigenetically silenced through DNA methylation maintained by DNA METHYLTRANSFERASE1 (MET1, for CG methylation), and DECREASE IN DNA METHYLATION1 (DDM1) which is a SWI/SNF chromatin-remodeling

**Fig. 2** (continued) regulation of gene expression, AGO4 functions in the nuclear hc-siRNA pathway during RdRM. AGO9 is thought to function in a similar pathway with that of AGO4, but with a highly localized expression in the developing ovules. Data for *AGO6* was not available because it was not included in the ATH1 arrays. (**b**) A developmental expression pattern for *Arabidopsis* genes encoding key proteins involved in DNA methylation. DNA METHYLTRANSFERASE1 (MET1) maintains DNA methylation in CG context. CHROMOMETHYLASE3 (CMT3) maintains CHG methylation. DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) is a de novo DNA methyltransferase and maintains CHH methylation. DRM2 provides the catalytic activity for DNA methylation during RdDM. DECREASE IN DNA METHYLATION1 (DDM1), a SWI/SNF chromatin-remodeling ATPase, is involved in cytosine methylation in both CG and non-CG contexts. All data were extracted from the AtGenExpress dataset (Schmid et al. 2005) through the Botany Array Resource (BAR) server (Toufighi et al. 2005). *Root* and *Cotyledon* were from 7-day-old soil-grown seedlings; *Rosette (1)* and *Rosette (2)* refer to vegetative rosette from 14-day-old (short day period of 10-h light) and 21-day-old (continuous light) soil-grown plants, respectively; *Cauline leaf* was from 21-day-old (continuous light) soil-grown plants; *Inflorescence* including the shoot apex (after bolting) was also from 21-day-old plants; The *Sepals (12)*, *Petals (12)*, and *Carpels (12)* were from stage 12 flowers; *Mature Pollen* was from 6-week-old plants; *Siliques (4)* and *(6)* refer to siliques with seeds at stages corresponding to early to late heart embryos (4), and mid to late torpedo embryos (6), respectively; *Siliques (7)*, *(8)*, and *(10)* actually refer to seeds at the corresponding stages without siliques, which are late torpedo to early walking-stick embryos (7), walking stick to early curled cotyledons embryos (8), and green cotyledons embryos (10), respectively

factor (Fig. 1b) (Slotkin and Martienssen 2007; Law and Jacobsen 2010), the contribution of RdDM to TGS at endogenous loci has been a subject of broad interest (Lippman et al. 2003, 2004; Blevins et al. 2009; Mirouze et al. 2009; Teixeira et al. 2009). Release of TE silencing as evidenced by transcriptional activation is often seen in *met1* or *ddm1* genetic backgrounds, as well as in hc-siRNA-defective mutants. However, although actual TE transposition has been observed in *ddm1* background (Hirochika et al. 2000; Tsukahara et al. 2009), it has not been reported in hc-siRNA-deficient mutants. Thus, RdDM appears to serve as a backup or complementary system for reinforcement of DNA methylation maintained by MET1 and DDM1 (Fig. 2b). This view is further supported by the fact that hc-siRNAs are able to direct DRM2-catalyzed de novo DNA methylation in all sequence context and that efficient TE reactivation may require a combined loss of both systems (Mirouze et al. 2009; Teixeira et al. 2009). Of note, the developmental expression pattern of DRM2 is highly similar to that of AGO4, but substantially differs from those of MET1, DDM1, and CHROMOMETHYLASE3 (CMT3) which maintains DNA methylation in CHG context (Fig. 2b). Interestingly, several lines of emerging evidence suggest that regulated operation of hc-siRNAs and RdDM may play an important role in germ cell and embryo development during plant reproduction. During male gametogenesis, coincidental downregulation of DDM1 and RdDM machinery correlated with TE reactivation and mobilization in pollen vegetative nucleus (VN) (Slotkin et al. 2009). TE-derived siRNAs (mostly 21-nt) which accumulated in pollen, presumably resulting from a PTGS surveillance mechanism in the vegetative nucleus in response to TE reactivation, are speculated to migrate into the sperms and direct TE silencing (Slotkin et al. 2009). In principle, a similar mechanism could also operate in female gametogenesis. In this scenario, TE relaxation and TE-derived siRNA production may occur in the central cells which are known to express elevated levels of DEMETER (DME) (Choi et al. 2002), a DNA glycosylase which is similar to ROS1 and catalyzes DNA demethylation (Morales-Ruiz et al. 2006). These TE-derived siRNAs could then direct TE silencing in the egg cell. After fertilization, maternally expressed, TE-derived hc-siRNAs that accumulate in the developing seeds (Mosher et al. 2009), presumably resulting from TE derepression in the endosperm due to massive DNA demethylation (Gehring et al. 2009; Hsieh et al. 2009), are speculated to direct epigenetic silencing of TEs in the embryo. These ideas are consistent with the observations that small RNAs are mobile in plants (Brosnan et al. 2007; Dunoyer et al. 2010; Molnar et al. 2010; Melnyk et al. 2011) and that TE sequences are hypermethylated in the sperm and developing embryos (Gehring et al. 2009; Hsieh et al. 2009; Slotkin et al. 2009). Maternally derived hc-siRNAs may also function to suppress the expression of certain paternally inherited alleles in early embryogenesis (Autran et al. 2011). These observations collectively suggest an emerging picture in which a burst of siRNAs in supporting cells may serve as a mobile signal for proper reprogramming of the genomes in the germ cells as well as the developing embryos.

## 4.4 Concluding Remarks

The past decade has witnessed the discovery of a hidden small RNA world and the rapid unveiling of its many facets. In this chapter, we have summarized some of the important aspects regarding the biogenesis and function of small RNAs in plants. The remarkable progresses in the emerging field of small RNA biology were made possible, to a great extent, by the numerous sequenced genomes for diverse model organisms and have been accelerated by the new enabling technologies including NGS. It is interesting to note that plants, in which RNA silencing was first discovered, have continued to be a rich source for discovery of many surprising and unique features associated with small RNA-mediated regulatory mechanisms. With new genomic and epigenomic tools being added to the conventional genetic and biochemical toolboxes, perhaps along with the advancement of cell separation technology in plants, we have every reason to foresee that the small RNA field will continue to be one of the fastest moving frontiers in plant biology in the years to come.

**Acknowledgments** We thank Matthew Olson for critically reading the manuscript. We apologize to colleagues whose relevant work could not be covered due to space limitations. Research in our laboratory is currently supported by grants from the US National Institutes of Health and the Southwest Consortium on Plant Genetics and Water Resources (to Z.X.).

## References

Addo-Quaye C, Snyder JA, Park YB, Li YF, Sunkar R, Axtell MJ (2009) Sliced microRNA targets and precise loop-first processing of MIR319 hairpins revealed by analysis of the *Physcomitrella patens* degradome. RNA 15:2112–2121

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 16:927–932

Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121:207–221

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 301:653–657

Arazi T, Talmor-Neiman M, Stav R, Riese M, Huijser P, Baulcombe DC (2005) Cloning and characterization of micro-RNAs from moss. Plant J 43:837–848

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 15:2730–2741

Autran D, Baroux C, Raissig MT, Lenormand T, Wittig M, Grob S, Steimer A, Barann M, Klostermeier UC, Leblanc O, Vielle-Calzada JP, Rosenstiel P, Grimanelli D, Grossniklaus U (2011) Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis. Cell 145:707–719

Axtell MJ, Bartel DP (2005) Antiquity of microRNAs and their targets in land plants. Plant Cell 17:1658–1673

Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006) A two-hit trigger for siRNA biogenesis in plants. Cell 127:565–577
Axtell MJ, Snyder JA, Bartel DP (2007) Common functions for diverse small RNAs of land plants. Plant Cell 19:1750–1769
Backman TW, Sullivan CM, Cumbie JS, Miller ZA, Chapman EJ, Fahlgren N, Givan SA, Carrington JC, Kasschau KD (2008) Update of ASRP: the *Arabidopsis* Small RNA project database. Nucleic Acids Res 36:D982–985
Baulcombe D (2004) RNA silencing in plants. Nature 431:356–363
Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci USA 102:11928–11933
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 62:454–462
Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366
Blevins T, Pontes O, Pikaard CS, Meins F Jr (2009) Heterochromatic siRNAs and DDM1 independently silence aberrant 5S rDNA transcripts in Arabidopsis. PLoS One 4:e5932
Bologna NG, Mateos JL, Bresso EG, Palatnik JF (2009) A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. EMBO J 28:3646–3656
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 123:1279–1291
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 320:1185–1190
Brosnan CA, Mitter N, Christie M, Smith NA, Waterhouse PM, Carroll BJ (2007) Nuclear gene silencing directs reception of long-distance mRNA silencing in *Arabidopsis*. Proc Natl Acad Sci USA 104:14741–14746
Catalanotto C, Azzalin G, Macino G, Cogoni C (2000) Gene silencing in worms and Fungi. Nature 404:245
Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE (2004) RNA silencing genes control de novo DNA methylation. Science 303:1336
Chan SW, Zhang X, Bernatavichute YV, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol 4:e363
Chapman EJ, Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. Nat Rev Genet 8:884–896
Chen X (2004) A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. Science 303:2022–2025
Chen X (2010) Small RNAs - secrets and surprises of the genome. Plant J 61:941–958
Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH (2010) 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc Natl Acad Sci USA 107:15269–15274
Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell 110:33–42
Cogoni C, Macino G (1999) Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature 399:166–169
Cogoni C, Macino G (2000) Post-transcriptional gene silencing across kingdoms. Curr Opin Genet Dev 10:638–643
Crevillen P, Dean C (2011) Regulation of the floral repressor gene FLC: the complexity of transcription in a chromatin context. Curr Opin Plant Biol 14:38–44
Cuperus JT, Montgomery TA, Fahlgren N, Burke RT, Townsend T, Sullivan CM, Carrington JC (2009) Identification of MIR390a precursor processing-defective mutants in *Arabidopsis* by direct genome sequencing. Proc Natl Acad Sci USA 107:466–471