AN ABSTRACT OF THE DISSERTATION OF

Elisabeth J. Chapman for the degree of Doctor of Philosophy in Molecular and Cellular Biology presented on April 24, 2007.

Title: Endogenous and Antiviral RNA Silencing Pathways in *Arabidopsis*.

Abstract approved:

_____

James C. Carrington

RNA silencing pathways are required for a wide variety of processes in most eucaryotes. In plants, small-interfering RNA (siRNA) arising from transposons and other repetitive sequences is associated with heterochromatin formation and maintenance. MicroRNAs and trans-acting siRNAs encoded at discrete loci function as negative regulators of gene expression by triggering cleavage or translational repression of mRNA transcripts with base complementarity to the small RNA. siRNA processed from viral RNA directs antiviral silencing that represses virus accumulation in plants and other organisms. Together, these pathways serve numerous functions in plants including genome maintenance, developmental timing and patterning and antiviral defense.

Virus-encoded RNA silencing suppressor proteins are viral pathogenicity factors and inhibit the antiviral silencing response through interaction with small RNA intermediates. In this work, small RNA duplex binding was demonstrated for unrelated suppressors from multiple viruses using molecular biology and biochemistry techniques. Sequestration of virus-derived siRNA and microRNA/microRNA* duplexes, inhibition of microRNA methylation, and perturbation of *Arabidopsis* development was demonstrated for several suppressors using transgenic approaches and molecular techniques. Suppressor inhibition of antiviral RNA silencing and endogenous microRNA pathways indicates that small RNA binding is a common strategy used by unrelated viruses, and suggests that interference with miRNA-directed processes may be a general feature contributing to pathogenicity of many viruses.

Finally, small RNA preparation and high-throughput sequencing procedures were developed for profiling small RNA populations in *Arabidopsis*. Genome-wide profiles of small RNA from wild-type *Arabidopsis thaliana* and silencing pathway mutants revealed dynamic

changes in expression of some microRNA families as well as genome-wide distribution patterns of small RNA in plants. These results establish high-throughput sequencing as a small RNA profiling tool and provide a comprehensive description of the major small RNA pathways in *Arabidopsis*. Together, the results presented here provide a basic understanding of the breadth of small RNA pathways in plants and show how interference with these pathways by viruses contributes to virus disease.

ENDOGENOUS AND ANTIVIRAL RNA SILENCING PATHWAYS IN *ARABIDOPSIS*

by
Elisabeth J. Chapman

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented April 24, 2007
Commencement June 2007

Doctor of Philosophy dissertation of Elisabeth J. Chapman presented on April 24, 2007.

APPROVED:

_____

Major Professor, representing Molecular and Cellular Biology

_____

Director of the Molecular and Cellular Biology Program

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

Elisabeth J. Chapman, Author

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to the following people for support and guidance throughout the preparation of this dissertation: James Carrington, for the opportunity to work in the Carrington lab; Kristin D. Kasschau for guidance and support throughout the course of this work; present and former members of the Carrington lab for support and many helpful discussions; and Karen Tutmark, Amy Shatswell and Sarah Dvorak for technical assistance. Thank you to my family, particularly my husband Jeff, without whom none of this would be possible or be any fun. Thanks also to the CGRB Core Lab members, CGRB office staff and to fellow graduate students in the Molecular and Cellular Biology Program for always being there when I needed help and friendship.

CONTRIBUTION OF AUTHORS

Alexey I. Prokhnevsky and Kodetham Gopinath produced cDNA clones and recombinant proteins used for production of transgenic plants and for p21 electrophoretic mobility shift assays in Chapter 2. RISC activity assays, RNA binding assays and TEV experiments in Chapter 3 were done by Lorant Lakatos. Yu-Ping Liu launched BYV virus infection, shown in Chapter 3. Bin Yu and Zhiyong Yang did filter hybridization analyses of the RNA and assisted with writing in Chapter 3. Noah Fahlgren assisted with small RNA library construction in Chapter 4. Noah Fahlgren, Kristin D. Kasschau, Christopher M. Sullivan, Jason Cumbie and Miya Howell assisted with experimental design and biocomputing, executed experiments and assisted with writing in Chapter 4. Additional data presented in the published manuscripts are not reproduced here.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

CHAPTER 1
GENERAL INTRODUCTION

Elisabeth J. Chapman & James C. Carrington

In preparation for submission to *Nature Reviews Genetics*
The Macmillan Building, 4 Crinan Street, London, N1 9XW, UK

Summary

Genome expansion in higher plants and animals is partially attributed to increased capacity for the spatio-temporal control of gene expression required for development of complex organisms. Throughout evolution of higher plants and animals, specificity modules for regulation of gene expression have proliferated for regulatory specialization at the transcriptional, post-transcriptional, and post-translational levels. Specificity modules for transcriptional regulation and proteasome-mediated degradation are proteins, and represent as much as 10% of the genome. In contrast, specificity for RNA silencing-based regulation is conferred by unique classes of small RNA guides (microRNA, siRNA, piRNA) that function with specialized protein-based effectors (members of the ARGONAUTE and PIWI families). Therefore, evolution of the RNA silencing mechanism as a multicellular regulatory device has involved proliferation and specialization of small RNA biogenesis and effector proteins, and *de novo* formation, proliferation and refinement of small RNA guides. How have the core biogenesis and effector modules proliferated and specialized to generate diverse regulatory pathways? How have new specificities emerged, and how has RNA silencing been integrated into regulatory networks? We review the evolution of the basic RNA silencing modules, focusing on three areas where regulatory proliferation and functional specialization has occurred: diversification and specialization of small RNA processing factors, proliferation and specialization of effectors and de novo evolution of novel small RNA regulators. We focus on plant systems, where the diversity and genome-wide deployment of silencing mechanisms may be greatest, but also include a broader perspective from other systems.

Proliferation and Specialization of Small RNA Biogenesis Factors

Initial descriptions of RNA silencing in a variety of organisms were reported just over a decade ago [1-5]. Observations that RNA silencing can be experimentally induced led to intensive investigation of experimental applications of RNA silencing (RNA interference [RNAi]) as well of the endogenous roles of RNA silencing pathways. The canonical RNA silencing pathway involves processing of a double-stranded RNA by a member of the Dicer family, producing small RNA that incorporates into an Argonaute protein-containing complex. The small RNA serves as a sequence-specificity determinant that guides the Argonaute complex to targets containing sequence complementary to the small RNA (Fig. 1.1). Fates of small RNA-Argonaute targets include messenger RNA cleavage, translational repression, and recruitment of additional small RNA biogenesis or DNA methylation factors. Initial studies of

RNA silencing in plants, *Caenorhabditis elegans* and *Drosophila melanogaster* indicated that accumulation of small RNA, 21-25 nucleotides in length, was a hallmark of silencing associated with enzymatic degradation of messenger RNA [6-11]. Small RNA intermediates are common to all RNA silencing pathways and are now recognized as the core specificity determinant in the broad range of small RNA regulatory pathways found in eucaryotes. Several different classes of small RNA have been described.

Small RNA classes and genomic origin

Small interfering RNA (siRNA; Table 1.1) arises from a variety of triggers including experimentally introduced double-stranded RNA, viruses, transgenes, endogenous repeat loci, transposon and retroelement loci, and in plants, from specialized trans-acting siRNA (tasiRNA)-encoding loci termed *TAS* genes. Plants also produce natural antisense transcript-derived siRNA (nat-siRNA) from loci with overlapping, bidirectional transcripts from two genes [12, 13]. SiRNA [21 nucleotides (nt)] processed from double-stranded RNA or transgenes, *TAS* locus transcripts, and some transposon/retroelement loci by RNAse III-like enzymes in the Dicer family (described below) is incorporated into RISC, the RNA-induced silencing effector complex initially described in *Drosophila* cells [7, 10, 14]. Thus programmed, RISC is guided to RNA containing complementarity to the guide siRNA and catalyzes cleavage or translational repression of the targeted RNA (reviewed in [15]). Heterochromatin-associated siRNAs (24 nt) arise from loci with repeat sequences in *Arabidopsis* and are involved in DNA methylation and establishment or maintenance of transcriptionally silent chromatin at some loci (reviewed in [16]).

MicroRNAs (miRNAs; Table 1.1) are a class of endogenous ~21-22 nucleotide RNAs that control a wide variety of growth and developmental processes in eucaryotes through post-transcriptional inhibition of messenger RNA translation (reviewed in [15, 17]). MiRNAs are encoded at *MIRNA* loci, and miRNA biogenesis requires transcription by RNA polymerase II and processing of nascent transcripts by RNAse III-like enzymes and double-stranded RNA binding partners to release a miRNA duplex intermediate (reviewed in [17]). Once assembled into RISC, the miRNA guides base-pairing interactions and cleavage or translational repression of mRNAs containing regions of sequence complementary to the miRNA (reviewed in [15]). The first *MIRNA* gene identified, encoding the lin-4 miRNA, was found through a forward-genetics screen for mutants in developmental timing in *Caenorhabditis elegans* [18, 19]. Additional screens as well as sequencing of small RNA populations has since revealed the

existence of hundreds of distinct miRNAs in numerous organisms, but biological function and target genes are ascribed to only a fraction of these [20-22]. *MIRNA* genes and target gene families are not conserved between plants and animals, but miRNA-mediated regulation of cell fate decisions and developmental timing and patterning is found in both kingdoms.

Small RNA precursor processing yields a duplex intermediate containing perfectly (siRNA) or imperfectly (miRNA) paired strands with two unpaired base overhangs at each 3' end [10, 23]. Within a small RNA duplex, each 5' terminus contains a phosphate residue, and each 3' terminus contains hydroxyl residues on the 2' and 3' carbons [10]. The 5' phosphates are required for siRNA function [24]. In *Arabidopsis,* the 2' and 3' hydroxyl groups are required for methylation of the 3' terminus at the 2' carbon by the HEN1 methyltransferase, and miRNA populations *in vivo* are found in the methylated state [25-27]. The small RNA duplex intermediate is separated during incorporation of one strand into the RISC in a process dependent upon ATP, the phosphorylated 5' end of the selected RNA, and additional protein factors [24]. Strength of base-pair interactions differ at each end of the small RNA duplex, and this thermodynamic asymmetry ensures that one strand, the "guide" strand or strand antisense to the target, is selected for incorporation into the RISC while the non-selected strand (siRNA* or miRNA*) is rapidly degraded [23, 28]. In *Drosophila*, siRNA selection involves the double-stranded RNA binding protein R2D2 [29, 30]. Proteins with homology to R2D2 in *Arabidopsis* include HYL1, which contributes to miRNA accumulation, DRB4, involved in accumulation of trans-acting siRNA, and other members of the DRB family [31-34].

Recently an additional class of small RNA has been identified in *Drosophila* and mammals. Piwi-interacting RNA (piRNA; Table 1.1) is found in the germline in association with members of the Piwi subfamily of Argonaute proteins [35-39]. piRNAs are longer than other small RNA classes, 24-29nt, and are modified at the 3'end, unlike other classes of small RNA in animals [40]. piRNA biogenesis is unaffected by mutations in RNAi pathway genes [40], suggesting that piRNA biogenesis occurs through a mechanism distinct from that of siRNA and miRNA ([38]; discussed further below).

The Dicer family of small RNA processing enzymes

The majority of small RNAs are processed from precursor RNA by members of the Dicer family. Dicers are large, ~200 kDa proteins that generally contain DExD and –C ATPase/helicase domains, a domain of unknown function (DUF283), a PAZ domain that is unique to Dicer and Argonaute proteins as discussed below, two catalytic RNase III domains and

a double-stranded RNA binding (dsRB) domain (Fig. 1.2; [41-43]. Genomes of *Schizosaccharomyces pombe, Chlamydomonas reinhardtii, C. elegans* and vertebrate animals encode one Dicer protein, while *Drosophila, Tetrahymena* and *Neurospora crassa* each encode two Dicers. The single *Chlamydomonas* Dicer DCR1 lacks the PAZ and dsRB domains, and Dicers of the ciliate *Tetrahymena* lack these as well as the DUF283 domain [44]. In other organisms, all seven elements characteristic of Dicer are represented, although only plants encode Dicer proteins that contain all seven elements within a single protein. Plants encode four or more DICER-LIKE (DCL) proteins, and of these, DCL1, DCL3, and DCL4, but not DCL2, contain duplicated dsRB domains [41]. Thus, plant DCLs have proliferated to a greater extent than Dicers in other eucaryotes.

Specialization among Dicer proteins

Mutation of the single *S. pombe* Dicer gene, *dcr1,* prevents transcriptional silencing and heterochromatin establishment at the centromeres and the mating-type region *mat* [45, 46]. It is proposed that RNA products of *dcr1* are involved in the assembly of heterochromatin in this organism. Roles for RNA silencing components in genome structure and epigenetic modification have also been described in diverse organisms such as *Tetrahymena, Drosophila,* and *Arabidopsis* [47-50].

C. elegans and most other animals use a single Dicer to produce endogenous and experimentally-induced siRNA as well as miRNA. Activities of a single Dicer nevertheless generate small RNA with diverse applications. In *C. elegans,* specification of Dicer function requires interactions between DICER-1 (DCR-1) and other proteins to produce siRNA from exogenous double-stranded RNA, miRNA processed from hairpin precursors, tncRNA arising from non-hairpin-forming loci, or endogenous siRNAs that regulate protein coding genes [51-54]. Notable among the DCR-1 interacting proteins are ALG-1 and ALG-2, Argonaute proteins that are required for miRNA processing (discussed further below). RDE-4, a double-stranded RNA binding protein that interacts with DCR-1, was required for accumulation of RNAs from a locus on the X chromosome [51]. Dicer-interacting double-stranded RNA binding proteins are also implicated in production of specific small RNA classes in *Arabidopsis* and *Drosophila* (see discussion below). This indicates that in organisms limited to a single Dicer protein, multiple small RNA classes may be generated through contributions of specialized double-stranded RNA binding proteins or other processing cofactors.

Another interesting group of DCR-1 interactors consists of proteins previously identified in mutants with an enhanced RNAi response (*eri* mutants; [55, 56]. The observation that ERI proteins interact with DCR-1 suggests a mechanism for the enhanced response of *eri* mutants in which loss of one ERI protein releases DCR-1 for entry into another RNA silencing pathway, thus boosting activity in that pathway [51]. This raises the question of whether RNA silencing components may be limiting, and competition among these factors prevents aberrant dicing activities by titrating Dicers out to suitable templates [51, 54]. A similar competition model has been proposed to explain small RNA accumulation patterns in *Arabidopsis,* in which specific DCL proteins faithfully generate specific small RNA classes from viruses and certain trigger loci, even though several DCL functions are redundant [57-60].

In fungi, flies, and plants, Dicer genes have duplicated or proliferated and specialized to various degrees. The fungus *Magnaporthe oryzae* contains two Dicers, MDL-1 and MDL-2, and mutation of *mdl-2* is sufficient to inhibit hairpin-induced silencing of a reporter transgene [61]. Interestingly, *Neurospora crassa* also contains two members of the Dicer gene family, and inhibition of transgene-induced or hairpin-induced silencing (termed "quelling" in this organism) requires mutation of both *dcl-1* and *dcl-2* [62]. These data suggest a degree of specialization in *Magnaporthe* that is not observed in *Neurospora*.

Specialization of miRNA maturation pathways is evident in flies and mammals, where miRNA biogenesis is a two-step process. In *Drosophila,* the RNAse III-like enzyme Drosha and its RNA binding partner Pasha are required for pri-miRNA recognition in the nucleus [63], and Dicer-1 (Dcr-1) and Loquacious act in the cytoplasm to further process pre-miRNA into mature miRNA [64-66]. A second Dicer exists in flies, Dcr-2, which acts in concert with the RNA binding protein R2D2 to process siRNA precursors and effect siRNA-mediated silencing [29, 30]. Therefore, it appears that the miRNA and siRNA maturation pathways have specialized in *Drosophila* and that interactions between double-stranded RNA binding proteins and Dicers guide Dicer specificity toward an siRNA or miRNA precursor substrate. Mammals, however, produce siRNA and miRNA through the activity of a single Dicer. The mammalian miRNA biogenesis pathway seems similar to that in flies: the Pasha homolog DGCR8 is required for pri-miRNA recognition by Drosha and processing into pre-miRNAs in the nucleus [67] and pre-miRNAs are subsequently processed by in the cytoplasm by Dicer and TRBP (a Loquacious homolog) [68, 69]. It remains to be determined whether or not additional double-stranded RNA binding proteins are involved in siRNA precursor recognition by Dicer in mammals.

Specialization of RNA silencing factors for miRNA biogenesis is also apparent in plants (Fig. 1.3). MiRNA accumulation in *Arabidopsis* requires DCL1 and the double-stranded RNA binding protein HYL1 [70-74]. DCL1 and HYL1 co-localize in the nucleus where HYL1 interacts directly with miRNA precursors [32, 75, 76]. The *Arabidopsis* genome contains three additional *DICER-LIKE* genes, but mutations in these genes do not affect miRNA accumulation [59, 73]. Drosha homologs have not been detected in the *Arabidopsis* genome, and purified DCL1 catalyzes accumulation of 21 nt small RNA *in vitro* [77]. Thus, DCL1 is the primary miRNA processing enzyme in *Arabidopsis*. Consequently, weak *dcl1* mutant alleles trigger severe developmental and reproductive defects, and strong alleles are embryonic lethal [42]. Mutations in *HYL1* are mild in comparison, and many miRNAs accumulate to detectable levels in *hyl1* mutants, suggesting that HYL1 is dispensible for miRNA production or may be redundant. Plants encode several other double-stranded RNA binding proteins in the HYL1/DRB family, and one or more of these proteins may substitute for HYL1 function [57, 78].

DCL1 appears to be highly specialized for miRNA production in *Arabidopsis*. DCL1 is not required for transgene silencing [79] or accumulation of heterochromatin-associated endogenous siRNA [73] and does not contribute significantly to accumulation of virus-derived siRNA from several RNA viruses [57, 73]. However, DCL1 is required for accumulation of virus-derived siRNA arising from a predicted imperfect hairpin RNA transcript in the pararetrovirus *Cauliflower mosaic virus* [80]. High-throughput sequencing analysis of DCL1-dependent small RNA populations in *dcl2/dcl3/dcl4* triple mutants demonstrated that DCL1 retains the ability to process 21nt siRNA from imperfect hairpin-like RNA [60]. These findings suggest that DCL1 has specialized for production of 21nt from imperfect hairpins, the majority of which are likely transcripts from *MIRNA* genes.

In *Arabidopsis*, endogenous 24 nt siRNAs originate from transposons and other retroelements and are associated with RNA-directed DNA methylation and chromatin remodeling (Fig. 1.3; [47, 48, 59, 73]. On a genome-wide scale, production of retroelement-associated 24 nt siRNA is dependent upon DCL3 [59]. Although DCL3 is the primary processing enzyme that generates siRNA from retroelement loci, DCL2 and DCL4 can substitute for these activities and produce off-size siRNA [59, 81]. Interestingly, DCL2 and DCL4 activities may be required together to repress retroelement loci; *dcl2/dcl3* and *dcl3/dcl4* double mutants, but not a *dcl3* single mutant, accumulate stochastic phenotypic abnormalities over subsequent generations, as is predicted to accompany reactivation of transposable elements from

heterochromatic loci [81]. DCL3 does not contribute to miRNA or tasiRNA accumulation and is dispensable for production of virus-derived siRNA, and thus appears to have specialized for production of heterochromatin-associated siRNA [73].

Elegant experiments by Pontes and colleagues [82] show that DCL3 colocalizes in nucleolar bodies with RNA-DEPENDENT RNA POLYMERASE2 (RDR2) and acts downstream of RDR2 and Pol IVa, a plant-specific RNA polymerase. In these experiments, Pol IVa colocalized with repeat loci in the nucleus and was required, along with DCL3 and RDR2, for siRNA production from these loci. These authors propose a model for chromatin modification in which primary transcripts stalled by methylation are recognized as aberrant and attract RNA synthesis by Pol IVa. Pol IVa products then move to a nucleolar processing center where they are converted by RDR2 into a template for DCL3, which processes 24 nt siRNA from these precursors. DCL3 is required for de novo methylation and maintenance of heterochromatin at some loci [73, 83] and is proposed to generate effector molecules for DNA methylation mediated by ARGONAUTE4 (discussed below).

Arabidopsis encodes two additional DICER-LIKE proteins, DICER-LIKE2 (DCL2), and DICER-LIKE4 (DCL4). A full catalog of DCL2 functions is likely not complete, however, DCL2 is required for processing of nat-siRNA, which suggests that it may be related to stress responses [12]. DCL4 is required for processing of 21nt tasiRNA that regulates the vegetative-to-adult phase transition in *Arabidopsis* [81, 84, 85] and for production of virus-derived siRNA that programs antiviral RISC and propagates an antiviral silencing signal [58, 86]. DCL4 activity at TAS loci is associated with conversion of single-stranded RNA to a double-stranded RNA precursor by RDR6. Interestingly, RDR6 is also implicated in limiting virus spread in infected plants [87]. DCL4 may, therefore, be specialized for processing of RDR6 products.

Antiviral defense and proliferation of plant Dicers

Mutations in *A. thaliana DCL1* do not affect accumulation patterns of *Cabbage leaf curl virus*-, *Turnip crinkle virus* (TCV)- or *Cucumber mosaic virus* (CMV)-derived siRNA alone or in combination with mutations in other DCLs, suggesting that DCL1 does not contribute substantially to siRNA accumulation from these viruses [57, 88]. However, DCL1 produces processes a predicted imperfect hairpin RNA transcript in *Cauliflower mosaic virus*, as mentioned above [80]. Thus, DCL1 does interact with RNA from some viruses. Roles for each of the three remaining DCL proteins in antiviral silencing in *Arabidopsis* have been demonstrated, although DCL3-dependent 24nt siRNAs that arise in plants infected with *Tobacco*

rattle virus are not necessary for antiviral defense, and DCL2 seems largely to serve a supportive role for DCL4 activity [57, 58].

The notable proliferation of plant Dicers may in part be a consequence of the role of Dicer activities in antiviral defense [58]. Plant Dicers process viral RNA into siRNA that programs antiviral RISC activity [58, 89] and propagates an antiviral signaling signal to distal tissues [58, 86, 90]. Accumulation of virus-derived siRNA is associated with reduced virus accumulation [57, 80]. Consequently, plant viruses frequently encode RNA silencing suppressor proteins that inhibit Dicer activities or sequester specific sizes of small RNAs [57, 91-95]. A successful RNA silencing-based antiviral defense strategy therefore likely includes the activity of multiple Dicers and production of multiple size classes of virus-derived siRNA [58]. Thus, selection favoring individuals with a strong antiviral defense response may also favor duplication and diversification of functions of plant Dicers. Antiviral silencing also occurs in cultured cells of *Drosophila* and *C. elegans,* where Dicer proteins have not proliferated [96-98]. However, no naturally occurring viral infections of *C. elegans* have been reported, suggesting little pressure to evade viral RNA silencing suppression mechanisms. Also, both of these organisms have robust innate immunity pathways, which may lessen the need for antiviral RNA silencing [99].

Proliferation and Specialization of Small RNA Effector Complexes

Small RNA serves as the specificity determinant in RNA silencing pathways by incorporating into the RNA-induced silencing complex (RISC) [8]. Once programmed with a small RNA, RISC acts as a sequence-specific translational repressor or RNA degradative enzyme or recruits RNA-dependent RNA polymerases or DNA methylation machinery [15]. The decision among these fates is determined by a number of factors, including the degree of complementarity between the small RNA and its target, the nature of the programming RNA, and interaction with other proteins. In general, small RNA with a high degree of complementarity to the target RNA, such as trigger-derived siRNA or plant miRNA, confers RNA cleavage activity ("Slicing") to RISC, while base-pairing with a lower degree of complementarity, as is found in most animal miRNA-target pairs, results in translational repression [100, 101]. Mechanisms of translational repression are reviewed elsewhere [102], and involve mRNA cleavage and decapping, inhibition of translation initiation, and relocalization of mRNA-effector complexes into P-bodies, sites of mRNA destabilization [103]

Early biochemical fractionation experiments revealed a range of sizes for RISC purified from cells [104-108]. Analysis of purified RISC from various species revealed the presence of several proteins, but a member of the Argonaute family of proteins was ubiquitously present. Subsequent assays using recombinant Argonaute proteins or highly purified preparations demonstrated that the Argonaute protein and a small RNA guide alone were sufficient for RNA cleavage activity [109-112]. Members of the Argonaute family were named for *ARGONAUTE1* (*AGO1*), a gene required for leaf polarity and development in *Arabidopsis thaliana* [113]. Strong *ago1* alleles identified by Bohmert and colleagues [113] result in plants that develop abnormally, with an undifferentiated vegetative phenotype that reminded the authors of a "small squid". The initial Argonaute protein identified in *Drosophila,* Piwi (P-element-induced wimpy testis) is expressed in the germline [114] and is required for stem cell division [115]. An additional seminal Argonaute protein was also identified in *Arabidopsis,* the ZWILLE (PINHEAD) protein required for embryonic development and organization of the shoot meristem [116, 117].

Recognition that AGO1, Piwi, and ZWILLE are related proteins lead to identification of a domain conserved in this family, the PAZ domain (for Piwi/AGO/ZWILLE) [118]. PAZ domains are also found in Dicer proteins and are unique to these families (Fig. 1.2; [44, 118]. Biochemical experiments and co-crystal structures of Argonaute PAZ domains bound to RNA demonstrate that the PAZ domain is sufficient for small RNA binding, which is non-sequence specific, and forms a pocket for binding the 3' end of a guide siRNA [119]. Other domains characteristic of Argonaute proteins include an N-terminal and middle domains, and the PIWI domain (Fig. 1.2), which binds the 5' end of a guide siRNA and the target RNA. The PIWI domain folds into an RNAse H-like structure and contains an Asp-Asp-Asp/His/Glu/Lys motif associated with, but not sufficient for, catalytic cleavage of the target RNA [109, 110, 119]. Argonaute-mediated cleavage of an RNA target, guided by a complementary guide small RNA, has been termed "slicing" and thus, Argonaute proteins with catalytic function are termed "Slicers".

Of the genes involved in RNA silencing, members of the Argonaute family have proliferated most in a number of organisms. *C. elegans* contains 27 Argonautes, *A. thaliana* contains 10, *Drosophila* contains four members, and *Homo sapiens,* eight members. *S. pombe* is unusual in having only one Argonaute protein. Phylogenetic analysis of the Argonaute family identifies three subfamilies: the Ago-like members, similar to *Arabidopsis* ARGONAUTE1, the Piwi-like members, similar to *Drosophila* Piwi, and "Group 3" Argonautes, specific to *C.*

elegans and containing RDE-1, the first Argonaute protein to be associated with RNA silencing [119-121]. Ago-like and Group 3 members are associated with gene silencing and RNAi, while the Piwi-like members are expressed in the germline of *C. elegans, Drosophila*, and mammals and are required for sperm cell development [35, 37-40, 122]. Interestingly, no Piwi-like members are represented in plant genomes and no Ago-like members are present in amoeba, *Tetrahymena,* and other unicellular eucaryotes. Thus, it has been proposed that common progenitors of these organisms carried at least one Ago-like member and one Piwi-like member, and that throughout evolution, one or the other subfamily was lost or proliferated in some clades [44]. Identified or predicted Slicers are distributed throughout the subfamilies. Sequence and structural analyses suggest that Argonautes from Archeae and primitive eucaryotes may be Slicers as well [111, 123, 124]. Functions have been ascribed to only a few of the Argonautes, but already our limited understanding of this family indicates that significant diversification and specialization of function among the Argonautes has supported the evolution of novel small RNA classes and regulatory cascades in a number of organisms.

Effector specificity through association with small RNA classes

Arabidopsis thaliana encodes 10 members of the Ago-like subfamily of Argonaute proteins, and no Piwi-like members. Each of the 10 *Arabidopsis* AGO proteins contains the Arg-Arg-Arg/His motif associated with Slicer function, and Slicer activity has been demonstrated for AGO1 [112] and AGO4 [125]. AGO1 is required for miRNA-mediated regulation and transgene silencing, and is implicated in antiviral defense responses [112, 126, 127]. AGO1 interacts *in vivo* with miRNA and tasiRNA, as well as with silencing-generated siRNA from transgenes, but excludes DCL3-dependent heterochromatin-associated siRNA [112, 125]. Isolated AGO1 complexes containing endogenous miR165 cleaved a *PHV* target RNA *in vitro*, and this slicer activity was dependent on conserved residues within the catalytic PIWI domain [112]. These results, coupled with the phenotype of *ago1* mutants, suggest that AGO1 plays a major role in RNA silencing pathways in *Arabidopsis* and has specialized to function in miRNA-mediated regulation.

Virus-derived siRNA was not detected in complexes with AGO1 extracted from plants infected with CMV, TCV, or *crucifer Tobacco mosaic virus* [112]. While this might suggest that AGO1 is not integral in antiviral RNA silencing, results of other studies show that *ago1* mutants are hypersusceptible to CMV [126] and that AGO1 is targeted and inhibited by 2b, the RNA silencing suppressor protein encoded by CMV [128]. Indeed, CMV-derived siRNA, as

well as siRNA derived from an unrelated virus *Turnip yellow mosaic virus*, were detected in complexes with AGO1 in experiments done by Zhang and colleagues [128]. Thus, the role of AGO1 in antiviral silencing requires further characterization.

Is there a molecular basis for the selective interaction of AGO1 with 21nt RNAs, specifically miRNA? While not yet clear, it does not appear to be based solely on size or subcellular distribution [112]. An intriguing model is one proposed by Baumberger & Baulcombe [112], in which small RNA biogenesis and effector programming are linked such that specific Dicer products are funneled to the corresponding Argonaute through Dicer-Argonaute interactions. This model is analogous to the R2D2-mediated programming of Ago2 with siRNA products of DCR-2 in *Drosophila* [29, 129] and predicts that DCL1 products will be the predominant RNA associated with AGO1. Indeed, miRNA represents over 90% of AGO1-associated small RNA [125] and tasiRNA and CMV-derived siRNA, also detected in complexes with AGO1, while predominantly DCL4 products, are processed by DCL1 in the absence of DCL4 [57]. The model predicts also that DCL3-dependent siRNA will interact with a specific Argonaute. Indeed, 24nt siRNA represents the majority of RNA complexed with AGO4 [125], which is required for maintenance of DNA methylation at inverted repeats [48, 130]. In light of these putative DCL1-AGO1 and DCL3-AGO4 interactions, do similar interactions exist for the other DCL proteins and Argonautes? Good candidates for such an interaction are DCL4 and ARGONAUTE7 (AGO7). Accumulation of tasiRNA from the *TAS3* locus is dependent upon DCL4, but also requires AGO7 [34]. Although AGO7-tasiRNA interactions have not been reported and accumulation of tasiRNA from *TAS1* and *TAS2* does not require AGO7, DCL4-AGO7 interaction remains an open possibility. The role of AGO-DCL interactions in effector programming and associated downstream uses of specific DCL products is an interesting area for further study.

The Argonaute family in *C. elegans* has proliferated to 27 predicted members, all of which have been analyzed using forward or reverse genetics or biochemical approaches [120]. Several proteins among these appear to have specialized functions. ALG-1 and ALG-2 are highly similar proteins and are dispensable for RNAi but are required for growth, development, and germline maintenance [131]. ALG-1 and ALG-2 affect accumulation of miRNA and are proposed to aid in miRNA processing or function [131]. CSR-1 is required for chromosome segregation and germline RNAi, and PRG-1 is postulated to function in germline maintenance [120, 132]. Other Argonaute proteins in *C. elegans* serve specialized functions within the RNAi response, which in the model proposed by Yigit and colleagues [120] involves production of

primary siRNA from an RNAi trigger, followed by initial guide siRNA-directed cleavage of the target mRNA by RDE-1. This cleavage event recruits an RNA-dependent RNA polymerase (RDR) to the target, and RDR synthesizes double-stranded RNA using the target mRNA as template. The newly synthesized RNA-target mRNA hybrid is then processed into secondary siRNA which mediates degradation of additional copies of the target mRNA.

RDE-1, one of the first Argonautes identified in this organism, is required for the RNAi response and antiviral RNAi, but not for development [96, 121, 133, 134]. RDE-1 interacts with low-abundance primary siRNA derived from trigger double-stranded RNA but does not significantly interact with abundant secondary siRNA arising from RDR-dependent amplification of silencing at the target locus [120, 135, 136]. Secondary siRNAs arise from regions within and upstream of the target sequence, and interact with SAGO-1 and SAGO-2, Argonaute proteins that appear to be limiting factors in the RNAi response and lack the catalytic residues associated with slicer function [120, 137]. This Argonaute protein hierarchy is proposed to protect against off-targeting effects that may accompany transitivity, the production of secondary siRNA from regions outside the target sequence [120]. According to the model of Yigit and colleagues [120], interaction between secondary siRNA and RDE-1 is predicted to recruit RDR activity and generate additional siRNA. Such an amplification system could quickly spread into non-targeted, essential genes if allowed to propagate unchecked. Limiting primary siRNA-Argonaute interactions to non-Slicer Argonautes would prevent secondary siRNA biogenesis from regions distal to the initially targeted element. Thus, specialization among the *C. elegans* Argonaute proteins allows the animal to mount a potent, amplified RNAi response against target loci while protecting neighboring genes.

Perhaps the most intriguing example of Argonaute protein specificity lies in the proposed role of Piwi-like Argonaute proteins in piRNA biogenesis and silencing of transposon/retroelement loci in the germline. In *Drosophila,* protection of the germline from expression of transposons and other mobile genetic elements is essential to prevent transposon/retroelement reactivation and eventual hybrid dysgenesis [138]. piRNA (formerly recognized as repeat-associated RNA, rasiRNA) arises in the germline and silences retrotransposons and repetitive elements such *gypsy* transposons and *Stellate* (*Ste*) [38, 40]. Piwi family members Aubergine (Aub), Piwi and Ago3 are expressed in the germline [38, 132], and Aub is required for silencing of *Ste* and repetitive loci, such as transposons containing LTRs [40]. This indicates expression patterns and activities distinct from those of Ago1, which is broadly expressed and interacts with miRNA to mediate miRNA-mediated silencing, and of

Ago2, which interacts with exogenous siRNA to mediate RNA silencing but is not required for silencing at the *Ste* locus [139, 140]. All three of these Piwi-type proteins associate with piRNA, and although size classes of RNA associated with each protein overlap, each protein displays a unique size preference. Piwi associates with piRNA with a median size of ~25-26nt, while Aub-associated piRNA is ~24-25nt and Ago3-associated piRNA is ~24nt. In addition, piRNAs associated with Piwi and Aub frequently contain a 5'terminal uridine, while Ago3 does not show this preference. Finally, Piwi and Aub bind piRNA antisense to transposons, while Ago3 complexes contain sense piRNA [38, 141].

What is the function of these piRNA-interaction preferences among the Piwi proteins? Brennecke and colleagues [38] and Gunawardane and colleagues [141] propose that piRNA interactions with Piwi proteins are a consequence of the mechanism of piRNA biogenesis. In these studies, the authors observed that populations of piRNA do not contain complementary RNA with 5' ends offset by 2 nt, as is predicted for populations of Dicer products. Instead, piRNA guide strands and complementary strands offset by 10 nucleotides were predominant. They also observed that guide strands and complementary strands were frequently divided between Ago3- and Aub-interacting pools, and between Ago3- and Piwi-interacting pools. The model proposed for piRNA biogenesis involves cleavage of sense transposon transcript by a guide piRNA-Piwi or piRNA-Aub complex, generating a new mRNA 5' end near the middle of the paired region, at a position opposite nucleotides 10 and 11 of the piRNA. The cleavage product would then be incorporated into Ago3 and undergo further 3'end processing to release a new piRNA. This model suggests that hierarchical, specialized Argonaute proteins working in concert have acquired a novel function in biogenesis of small RNA.

Specificity through interactions with protein partners

In *Arabidopsis thaliana,* the majority of endogenous siRNA arises from repetitive DNA elements in heterochromatic regions [59, 142]. These heterochromatin-associated siRNAs are implicated in the establishment and maintenance of heterochromatin through a pathway that requires methyltransferases and histone-modifying enzymes as well as the RNA silencing factors AGO4, RDR2, DCL3, and a newly identified DNA-dependent RNA polymerase specific to plants, PolIV [143, 144]. PolIV consists of two subunits, NRPD1, which contains domains divergent from those in the paralogous subunits in plants, and NRPD2 [143, 144]. Two genes encode forms of NRPD1, *NRPD1a* and *NRPD1b,* and Pol IVa and Pol IVb forms are both active, each form containing a different NRPD1 subunit [143-145].

Experimental evidence supports a recently proposed model for the nuclear siRNA pathway in *Arabidopsis* in which AGO4 serves a role in chromatin modification through its interaction with Pol IVb. Pol IVa localizes to endogenous repeat loci in an RNA-dependent manner and is required for DCL3- and RDR2-dependent siRNA production from these loci [82]. The siRNA products of DCL3 incorporate into AGO4, which colocalizes with Cajal bodies in the nucleolus [146]. AGO4 interacts with the C-terminal domain of Pol IVb [146] and is required for localization of NRPD1b to the nucleolus [82]. NRPD2, however, does not localize to the nucleolus, suggesting that the NRPD1b-AGO4 interaction is independent of NRPD2. Pontes and colleagues propose that NRPD1b is a subunit of an AGO4-containing RISC that acts to mediate chromatin modification at endogenous repeat loci. Although the mechanism of AGO4-mediated chromatin modification is still unclear, AGO4 slicer activity is dispensible for maintenance of heterochromatin at some loci [125] suggesting roles for AGO4 that are distinct from those of other plant Argonautes.

De Novo Generation of Novel Small RNA Regulators

Various classes of small RNA can be divided into two categories based on genomic origin; miRNA and tasiRNA are encoded at *MIRNA* and *TAS* genes, and siRNA, nat-siRNA and piRNA are processed de novo from trigger loci that become the target of siRNA regulation. Biogenesis of small RNA from an RNA target allows for specificity in the small RNA pool that is equal in diversity to the target sequences themselves. This mechanism in theory has the potential to generate an unlimited reservoir of specificities that are fluid and can instantaneously be adapted in response to stimuli. The regulatory potential of such a reservoir can be inferred from the staggering number of unique small RNA species identified by high-throughput sequencing of small RNA populations that is not yet saturating [59, 60, 125, 142, 147-149]. Additionally, genesis of a small RNA pool from a trigger locus can provide regulators that target loci distinct from, but containing sequence homology to, trigger loci [38]. How are specific loci recognized by the RNA silencing machinery, and how have trigger loci evolved into potent regulators of gene regulation?

Double-stranded RNA production through bi-directional transcription

Initial RNA silencing experiments in *C. elegans* induced RNAi by injecting double-stranded RNA synthesized *in vitro* by transcription from opposing promoters [1]. Logically then, transcription from opposing promoters *in vivo* will also generate double-stranded RNA

inducers of RNA silencing. One of the few reported examples of this is the *Arabidopsis thaliana P5CDH* gene and the overlapping *SRO5,* a gene of unknown function [12]. Abiotic stress induces transcription of *SRO5,* and a 24nt nat-siRNA is processed from the resulting double-stranded RNA consisting of *SRO5* and *P5CDH* transcripts. Accumulation of the *SRO5-P5CDH* nat-siRNA is dependent upon DCL2, SGS3, RNA-DEPENDENT RNA POLYMERASE6 (RDR6) and NRPD1a, and correlates with cleavage of the *P5CDH* transcript.

An intriguing feature of this system is that cleavage of the *P5CDH* transcript was detected not only at the site predicted based on base-pair interactions with the *SRO5-P5CDH* nat-siRNA, but also at additional DCL1-dependent sites downstream and in 21nt register from the initial cleavage site. The mechanism of phased cleavage of 21nt RNA initiated from a small RNA-guided cleavage site is also observed in the production of tasiRNA in *Arabidopsis,* a process that also requires SGS3 and RDR6. In the nat-siRNA example, this mechanism results not only in degradation of the *P5CDH* target, but also in the production of small RNAs with sequences, therefore specificities, distinct from the initial *SRO5-P5CDH* nat-siRNA [12]. The potential for such secondary small RNA species to target additional transcripts has not been explored.

RNA-dependent RNA polymerases in generation of novel small RNA regulators

Production of small RNA by the nat-siRNA mechanism described above requires co-expression of genes whose transcripts contain significant sequence complementarity. How do loci that lack these features generate targets for RNA silencing? In these cases, an RNA-dependent RNA polymerase may be recruited to a transcript to generate a double-stranded RNA. In addition to its roles in tasiRNA and nat-siRNA biogenesis, RDR6 is also required for silencing of single copy transgenes in plants [150, 151] and is involved in resistance to CMV [152]. Accumulation of heterochromatin-associated 24nt siRNA and siRNA from *FWA* tandem repeats in *Arabidopsis* requires RDR2 [59, 153]. A maize protein similar to RDR2, MOP1, is required for epigenetic modifications (paramutation mediated by tandem repeats) [154]. In these examples, activity of an RNA-dependent RNA polymerase likely converts RNA transcripts into structures sufficient to trigger DCL processing and an RNA silencing response. Mechanisms for recruitment of RDRs to these loci are not yet clear, although evidence emerging from characterization of Arabidopsis *TAS* loci suggests that, similar to amplification of RNAi in *C. elegans* and nat-siRNA biogenesis in *Arabidopsis*, initial small RNA targeting plays a role in routing transcripts into an RDR-dependent small RNA biogenesis pathway [142,

155, 156]. *Arabidopsis* tasiRNA is encoded at multiple non-protein-encoding *TAS* loci [142, 157, 158]. Transcripts from each *TAS* locus are targets of miRNA-directed regulation and encode multiple distinct 21nt tasiRNAs that target and regulate the expression of genes in *trans* [142, 157-159]. Following miRNA-directed cleavage of primary TAS transcripts, DCL4 processes individual, non-overlapping tasiRNAs in phase from the miRNA cleavage site in an RDR6 and SGS3-dependent manner (Fig. 1.3). While the precise function of SGS3 is not clear, it is epistatic to RDR6 and is proposed to stabilize the primary transcript cleavage product to enable conversion by RDR6 into double-stranded RNA [159].

This pathway represents an example of a novel application of miRNA regulation, i.e. as a specificity component of the small RNA processing machinery employed to ensure accurate processing of a small RNA precursor. Production of siRNA from miRNA-regulated transcripts in plants is surprisingly uncommon on a genome-wide scale and interestingly, is almost entirely specific to transcripts targeted by more than one small RNA [142, 155, 156]. Transcripts from members of the *PPR-P* gene family contain multiple small RNA target sites and spawn abundant phased siRNA that are predicted to actively regulate these genes [156]. This gene family is rapidly expanding in higher plants, and consequently the role of small RNA regulation in genome expansion or conversely, for repression of expanding gene families, is an interesting area of future study.

Regulators from inverted duplications: regulation of retroelements and endogenous genes

RDR-independent small RNA arises from loci containing inverted duplications and from some types of transposon duplications and rearrangements. RNA from these loci is predicted to fold into self-complementary hairpin-like structures or other structured containing regions of extensive base-pairing that is recognized by one or more Dicer proteins. In male flies (*Drosophila melanogaster*), fertility requires repression of the *Stellate* repeats via *Su(Ste)* (*Suppressor of Stellate*), a series of tandem repeats derived from *Stellate* that include the promoter region, a mutated *Ste* ORF, and a *hoppel* transposon insertion [160]. Transcription from *Su(Ste)* results in production of double-stranded RNA and siRNA [161] and silencing in *trans* of *Stellate* [160]. In maize, the *Muk* allele contains an inverted duplication of a portion of the *MuDR* gene, and is the result of a rearrangement of a transposon from the Mutator family [162, 163]. Transcription from *Muk* results in a hairpin-forming RNA transcript and accumulation of *Muk*-derived small RNA. Presence of the *Muk* (Mu killer) locus in maize represses the accumulation of transcripts from Mutator transposons and results in heritable

silencing of this family of transposons. In each of these above cases, the silencing response mounted against a repeat element confers repression of additional loci, and in both cases, since *trans* regulation is advantageous, a basal level of RNA silencing may also be advantageous.

RNA silencing also represses expression of transposon-independent endogenous inverted duplications. The *I* (inhibitor) locus in soybean (*Glycine max*) represses expression of chalcone synthase (CHS) in the seed coat. The *I* locus contains a tandem repeat of *CHS* genes and an inverted duplication of the *CHS3* gene, and CHS silencing conferred by the *I* locus is associated with siRNA production [164]. CHS silencing is inhibited in potyvirus- and cucumovirus-infected plants, likely due to the activities of virus-encoded suppressors of RNA silencing [165]. *Arabidopsis thaliana,* unlike other members of the *Arabidopsis* genus, is self-compatible due to deletions and rearrangements of genes at the S (self-incompatibility)-locus [166]. In the *A. thaliana* accession Col-0, the S-locus gene *SKR* contains an inverted duplication that is transcribed into a hairpin-forming RNA and spawns 21nt siRNA [147]. It was initially proposed that self-compatibility in Col-0 is due to dominant negative activities of truncated proteins translated from nonsense *SKR* mRNA [167]. However, the distribution of small RNA from this inverted duplication also supports a model for self-compatibility in which the *SKR* gene is repressed by RNA silencing [147].

Evolution of MIRNA *genes*

The previous two examples demonstrate how gene duplications and rearrangements can trigger biogenesis of small RNA regulators that repress gene expression. In these cases, a heterogenous pool of siRNA arises from these loci, and these siRNAs act in *cis* to regulate the trigger locus, or in *trans* to regulate closely related loci. While siRNA-generating loci contribute significantly to the small RNA pool, miRNA also represents a significant portion of total small RNA populations [59, 142, 147-149] and *MIRNA* genes encode only a single small RNA that is frequently expressed at high levels and targets unrelated genes in *trans*. How have these specialized *MIRNA* loci evolved? We have proposed an inverted duplication model for the origin of miRNA regulators (Fig. 1.4; [168, 169]). In this model, aberrant transposition events or duplication and recombination of expressed gene sequences produces a small RNA-generating locus with the potential to be co-expressed with and regulate the progenitor gene. If this regulation were advantageous, the duplicated promoter element and region containing the regulating sequences would be selected for, and purifying selection will eliminate unnecessary and potentially deleterious cross-reacting regulators. Accumulated sequence drift will lead to

imperfect pairing in the foldback structure, requiring specialized miRNA processing factors for small RNA production.

This evolutionary model makes several predictions. Pre-miRNA transcripts from relatively new, non-conserved *MIRNA* loci will contain extensive similarity to their progenitor genes. New *MIRNA* loci will not have accumulated effects of purifying selection and thus will have longer foldback precursors. Finally, new miRNAs will not have evolved strict dependence upon miRNA biogenesis factors.

This model is supported by the observation that a number of small RNA-generating loci in *Arabidopsis* encode hairpin-forming RNA transcripts that have extensive similarity to protein-coding genes [142, 168, 169]. Some of these loci (i.e. *MIR161* and *MIR824*) encode discreet miRNAs rather than a heterogenous small RNA population [168, 169] and some are not DCL1-dependent [142]. Interestingly, non-conserved *Arabidopsis thaliana* miRNAs were identified that target the gene of origin, target unrelated genes, or that have lost targeting abilities. A three-fate model was proposed for small RNA-encoding genes that are products of gene duplication (Fig. 1.4; [169]). In the first fate, a regulator is selected with the capacity to regulate the progenitor sequence or progenitor family. In the second, targeting specificity for an unrelated gene or gene family is acquired. The third, and perhaps most common fate, involves elimination by mutation and sequence drift of the promoter and/or targeting elements, and the regulatory potential of the locus is eliminated.

Whether or not this model explains the number and diversity of *MIRNA* genes identified in eucaryotes awaits further study. Hundreds of *MIRNA* genes are now known [170]. The diversity of *MIRNA* families, target genes, and the roles of miRNA in the development of plants and animals, as well as in virus infections in animals, have been extensively reviewed (i.e. [21, 22, 171, 172].

Conclusion

Evolution of RNA silencing as a regulatory mechanism in eucaryotes has involved proliferation and specialization of all core components of RNA silencing pathways. Dicer proteins and processing cofactors such as double-stranded RNA binding proteins have specialized to produce ~21 nt miRNA and 21-24 nt siRNA for a variety of downstream processes. Argonaute proteins have proliferated significantly, and have evolved a range of small RNA-guided functions including amplification of RNA silencing and biogenesis of secondary small RNA, enzymatic cleavage of RNA targets, translational repression, and recruitment of

heterochromatin modification machinery. Transposon-associated DNA rearrangements and endogenous inverted duplications have generated novel small RNA regulators, and RNA-dependent RNA polymerases, and in plants, a fourth DNA-dependent RNA polymerase, have been deployed to convert primary transcripts into double-stranded RNA precursors for Dicer processing. Finally, numerous specialized MIRNA genes have evolved that encode potent trans-acting regulators of gene expression. The products of these specialization events now comprise an array of diverse small RNA-based regulatory pathways that control gene expression, the impact of which, particularly in the evolution of complexity in eucaryotic organisms, we are just now beginning to comprehend.

Table 1.1. Classes of small RNA identified in eucaryotes.

Class	Description	Biogenesis and Genomic Origin	Function
miRNA	microRNA	processing of foldback *MIRNA* gene transcripts by members of the Dicer family	post-transcriptional regulation of genes required for developmental timing, signaling, and patterning
siRNA	small interfering RNA	processing of RDR-derived long double-stranded or long foldback RNA by members of the Dicer family	suppression of transposons/retroelements; post-transcriptional regulation of transcript of origin; formation and maintenance of heterochromatin
tasiRNA	trans-acting siRNA	miRNA-dependent cleavage and RDR-dependent conversion of *TAS* gene transcripts to double-stranded RNA, followed by Dicer processing	post-transcriptional regulation of genes involved in vegetative phase change in plants; repression of rapidly expanding gene families in plants
nat-siRNA	natural antisense transcript-derived siRNA	Dicer processing of double-stranded RNA arising from sense- and antisense-transcripts from neighboring genes	post-transcriptional regulation of genes involved in pathogen defense and stress responses in plants
rasiRNA, piRNA	repeat-associated siRNA, piwi-interacting RNA	biogenesis mechanism unclear, may involve processing of transcripts from transposable elements	suppression of transposons/retroelements in the germline of flies and mammals

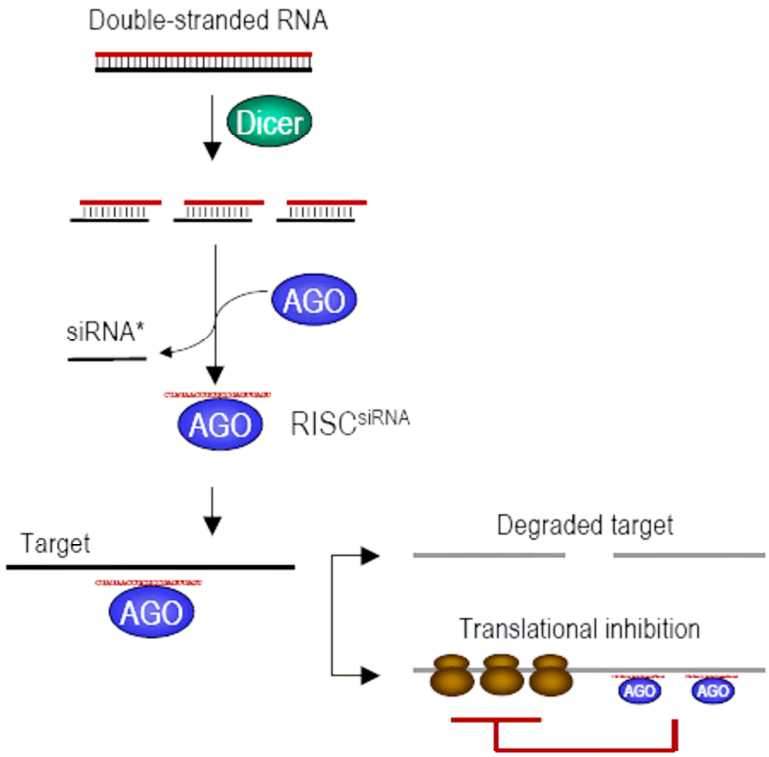

Fig. 1.1. Generalized core RNA silencing module. RNA silencing is triggered by double-stranded RNA, which is processed by a member of Dicer family of proteins into small RNA duplex intermediates. Small RNA duplexes are unwound and one strand is degraded (siRNA* is shown here as an example) while the guide strand (siRNA) incorporates into an Argonaute protein component of RISC (RNA-induced silencing complex). RISC programmed with a small RNA guide cleaves target RNA with sequence complementarity to the guide RNA.

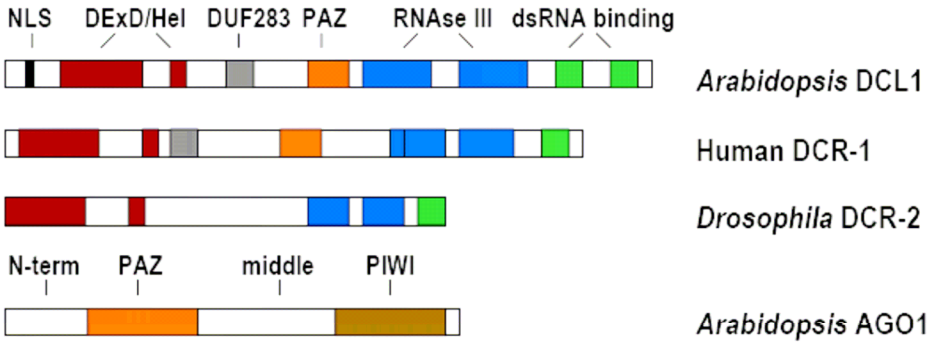

Fig. 1.2. Domain organization of representative members of the Dicer and Argonaute families. Domains found in members of the Dicer and Argonaute families include nuclear localization sequences (NLS), DExD and C-terminal domain helicase (DExD/Hel), a domain of unknown function (DUF283) and PAZ (for Piwi/Argonaute/Zwille), as well as domains involved in RNA cleavage and binding activities.

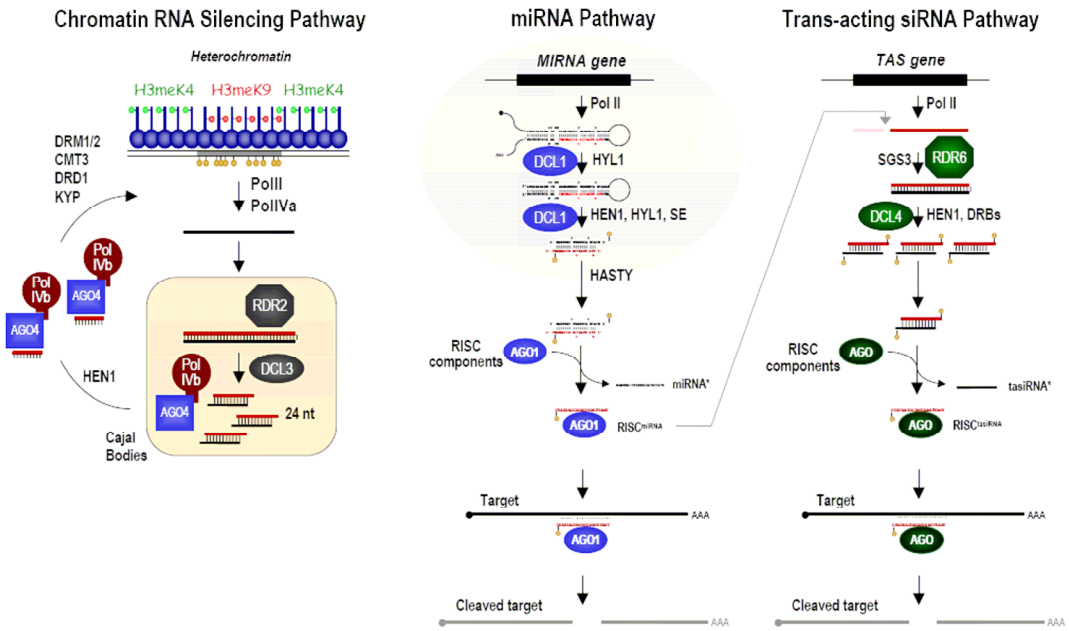

Fig. 1.3. RNA silencing pathways in plants. RNA-directed chromatin silencing is proposed to involve double-stranded RNA production through PolIVa activity on transcripts from methylated loci, followed by processing into siRNA by RDR2 and DCL3. siRNA interacts with AGO4 and PolIVb in Cajal bodies, and this complex relocalizes to repeat loci where it recruits DNA methylation machinery such as DRM1/2, CMT3, DRD1 and KYP. Processing of miRNA occurs via DCL1 and HYL1 cleavage of pre-miRNA transcripts from MIRNA loci, followed by additional processing by DCL1, SERRATE (SE), the double-stranded RNA binding protein (DRB) HYL1 and the HEN1 methyltransferase. miRNA duplexes are exported from the nucleus by HASTY, an Exportin-5-like protein. Mature miRNA assembles into the AGO1 component of the RISC (RNA-induced silencing complex) which targets RNA complementary to the guide RNA for cleavage. MiRNA-guided cleavage of TAS transcripts is followed by conversion of the cleavage product into double-stranded RNA in an SGS3- and RDR6-dependent manner. This double-stranded precursor is processed into trans-acting siRNA (tasiRNA) by DCL4, which requires HEN1 and DRBs. TasiRNAs regulate mRNAs in trans in a manner similar to miRNA regulation. Antiviral RNA silencing and nat-siRNA pathways are not shown.

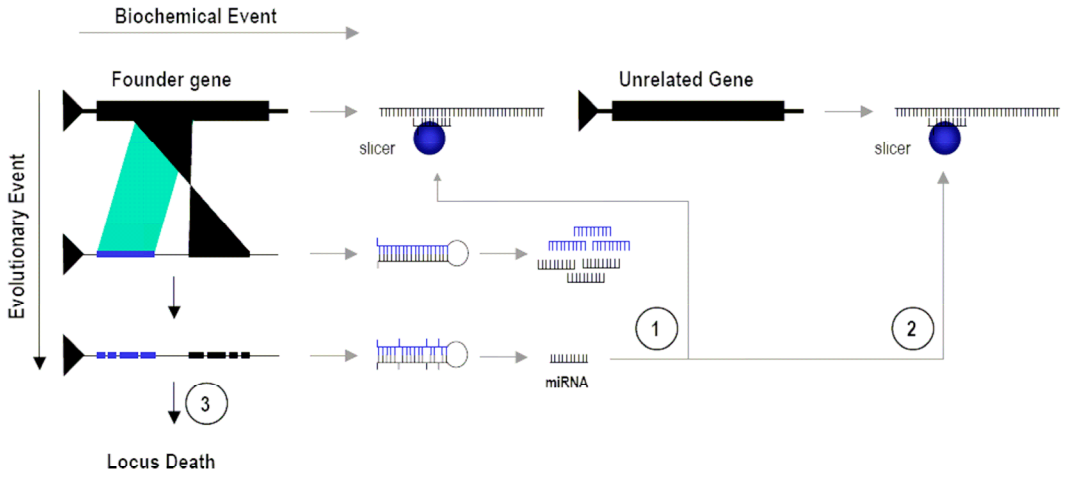

Fig. 1.4. Inverted duplication model for evolution of *MIRNA* genes. *MIRNA* genes born frequently by duplication events that form inverted repeats and produce siRNAs. Young *MIRNA* loci may experience three evolutionary fates: 1) Stabilization of a miRNA-generating sequence resulting in specificity for the gene or gene family of origin; 2) Stabilization of a miRNA-generating sequence that acquires specificity for a target distinct from the gene of origin; 3) Drift and death of the locus. Figure prepared with assistance from Noah Fahlgren.

CHAPTER 2

VIRAL RNA SILENCING SUPPRESSORS INHIBIT THE microRNA PATHWAY AT AN INTERMEDIATE STEP

Elisabeth J. Chapman, Alexey I. Prokhnevsky, Kodetham Gopinath

Genes & Development

Cold Spring Harbor Laboratory Press, 500 Sunnyside Boulevard, Woodbury, New York 11797

Volume 18 pages 1179-1186

Summary

RNA silencing suppressors from different plant viruses are structurally diverse. Besides inhibiting the antiviral silencing response to condition susceptibility, many suppressors are pathogenicity factors that cause disease or developmental abnormalities. Here, unrelated suppressors from multiple viruses were shown to inhibit microRNA (miRNA) activities and trigger an overlapping series of severe developmental defects in transgenic *Arabidopsis thaliana*. This suggests that interference with miRNA-directed processes may be a general feature contributing to pathogenicity of many viruses. A normally labile intermediate in the miRNA biogenesis/RISC assembly pathway, miRNA*, accumulated specifically in the presence of suppressors (P1/HC-Pro, p21 or p19) that inhibited miRNA-guided cleavage of target mRNAs. On a genome-wide level, miRNA target transcripts were generally up in miRNA biogenesis mutants and in plants expressing one of these suppressors, but were unaffected in the tasiRNA and chromatin siRNA mutants. Both p21 and p19, but not P1/HC-Pro, interacted with miRNA/miRNA* complexes and hairpin RNA-derived siRNAs *in vivo*. In addition, p21 bound to synthetic miRNA/miRNA* and siRNA duplexes *in vitro.*

Introduction

Systemic infection by plant viruses frequently results in disease symptoms that resemble developmental defects, including loss of leaf polarity, loss of proper control of cell division, and loss of reproductive functions [173]. These and other phenotypes are frequently associated with virus-encoded pathogenicity factors, many of which are suppressors of RNA silencing [174]. RNA silencing functions as an adaptive immune response that restricts accumulation or spread of inducing viruses [175]. Suppressor proteins encoded by members of different virus families are distinct, suggesting that plant viruses evolved this counterdefensive mechanism independently on many occasions [176, 177].

RNA silencing during virus infection is triggered by double-stranded RNA (dsRNA) generated during the course of virus replication or by the activity of a cellular RNA-dependent RNA polymerase [178]. Processing of dsRNA by DICER or DICER-LIKE enzymes results in heterogeneous short interfering RNAs (siRNAs) of 21 – 25 nucleotides [179]. siRNAs incorporate into RNA-induced silencing complexes (RISCs) [10, 11] and provide guide functions for sequence-specific ribonucleolytic activity. Protein components of RISC include ARGONAUTE family members, nucleases, and other factors [180].

microRNAs (miRNAs, ~21 nucleotides) are chemically similar to siRNAs, but they arise from processing of imperfect hairpin-forming RNA precursors transcribed from miRNA genes [181]. miRNA processing occurs by a multistep mechanism involving DICER [DICER-LIKE1 (DCL1) in *Arabidopsis*] activity to excise the miRNA from the hairpin stem [15]. miRNAs function as negative regulators of target mRNAs through directing either site-specific cleavage by RISC complexes or translational repression [15]. In plants, miRNAs target a wide range of mRNAs encoding transcription factors required for development [71, 182, 183]. These include factors required for meristem identity and maintenance, patterning, cell division, hormone signaling and developmental timing. Additionally, plant miRNAs also target mRNAs encoding miRNA metabolic factors and factors of unknown function [182, 184]. Loss of miRNA biogenesis or activity in *Arabidopsis* results in pleiotropic defects during embryonic, vegetative and reproductive development [42, 71, 185].

Despite differences in precursor structure for siRNAs (perfectly paired dsRNA) and miRNAs (imperfect hairpins), functional siRNA and miRNA molecules are incorporated into RISC through an asymmetric strand selection process. Precursor processing yields a duplex intermediate containing perfectly (siRNA) or imperfectly (miRNA) paired strands with two unpaired bases at each 3' end [10]. This duplex intermediate is unwound prior to, or during, incorporation of one strand into RISC complexes. Strand asymmetry depends on the strength of base-pair interactions at each end of the duplex, with the molecule containing the 5' end participating in the weakest interaction preferentially utilized [23, 28]. The non-selected strand (siRNA* or miRNA*) is rapidly degraded [15].

We previously showed that the *Turnip mosaic virus* (TuMV) silencing suppressor, P1/HC-Pro, interferes with miRNA-guided regulation of at least 10 target mRNAs in infected or transgenic *Arabidopsis* plants. The suppressor also caused multiple developmental defects, including some resembling those associated with *dcl1* mutants [185]. Here, we show that interference with miRNA-guided target cleavage/degradation and development in *Arabidopsis* is a general property of several, unrelated silencing suppressors encoded by evolutionarily distinct viruses. Through analysis of miRNA processing intermediates and suppressor-containing complexes *in vivo*, three of these suppressors were shown to inhibit the RISC assembly pathway after DCL1-catalyzed formation of miRNA/miRNA* duplexes.

Results

Interference with miRNA-guided target cleavage and development in *Arabidopsis* by P1/HC-Pro was shown previously using virus-infected and transgenic plants. To determine if miRNA interference is a general property of RNA silencing suppressors, the *Beet yellows virus* p21 [186], *Tomato bushy stunt virus* p19 [174, 187], *Turnip crinkle virus* coat protein (CP) [188, 189], and *Cucumber mosaic virus* 2b [190, 191] silencing suppressors were analyzed and compared to P1/HC-Pro [190, 192, 193]. These suppressors belong to evolutionarily and structurally unrelated protein families [186, 194-196]. To facilitate detection and experimental manipulation, constructs were expressed using *Cauliflower mosaic virus* 35S promoter and terminator sequences and engineered such that an influenza hemaglutinin (HA) epitope tag was added to the carboxy-terminus of each protein (Fig. 2A). Parallel sets of constructs were generated in which the coding sequence for soluble-modified GFP preceeds the carboxy-terminal HA epitope to generate GFP fusions to β-glucuronidase, wild-type suppressors P1/HC-Pro, p21, and p19, or mutant forms of these three suppressors. Mutant constructs HC-Pro (AS9), p21(81-21), and p19(W39/42R) were designed to encode proteins reported elsewhere to be compromised in suppression function [196-198]. To confirm that the epitope-tagged viral proteins used in this study were functional RNA silencing suppressors, we analyzed each construct in *Agrobacterium*-mediated transient assays in *Nicotiana benthamiana* leaves [199]. Transient assays involved co-expression of constructs encoding a reporter (GFP or SCL6, a target of miR171-mediated cleavage [100, 185]), an inducer of RNA silencing [a hairpin-generating construct (dsGFP) or miR171] and a silencing suppressor. In each assay, leaves were infiltrated with mixtures that either contained (experimental sample) or lacked (control sample) the suppressor construct. Suppression of GFP silencing was analyzed by inspection of visual GFP fluorescence, detection of GFP protein and mRNA, and detection of GFP-specific siRNAs. Suppression of miR171-directed cleavage was detected by determining the abundance of full-length SCL6 mRNA relative to the cleavage product.

Coexpression of GFP and dsGFP constructs without a silencing suppressor resulted in no or low levels of GFP fluorescence (Fig. 2.1A) and low or non-detectable levels of GFP mRNA and protein (Fig. 2.1B and 2.1C, lanes 5 and 6). Tissue expressing only GFP and dsGFP constructs also accumulated GFP-specific siRNAs of 21-24 nucleotides in length (Fig. 2.1D, lanes 5 and 6). Each of these features is characteristic of efficient RNA silencing of the GFP construct in the infiltration zone.

In contrast to the controls, RNA silencing was suppressed effectively in leaves that co-expressed each of the modified silencing suppressors P1/HC-Pro, p21, p19 and CP (Fig. 2.1A-C). Fluorescence was detected, and GFP protein and GFP mRNA accumulated to relatively high levels. P1/HC-Pro, p21 and p19 each failed to inhibit siRNA formation using dsGFP as the silencing inducer (Fig. 2.1D), indicating that these suppressors likely function to inhibit one or more steps after siRNA formation. Suppression of siRNA accumulation by p19 and P1/HC-Pro was previously reported by several groups using sense transgenes as a silencing locus [47, 151, 187, 200-202]. In these cases, efficient silencing of the inducer locus may require RNA-dependent RNA polymerase-mediated production of a double-stranded RNA intermediate and subsequent amplification steps [150, 151, 202, 203]. Loss of siRNA accumulation in these cases may be an indirect effect of suppression of amplification [151, 202].

The TCV CP suppressor inhibited siRNA formation, which is consistent with previous experiments using non-tagged versions in other assays [188]. The CP suppressor likely interferes with the dsRNA processing step catalyzed by a DICER-LIKE activity. CMV 2b was less effective than the other suppressors in this assay. GFP fluorescence was detected, but protein and mRNA accumulated only to low levels (Fig. 2.1A-C). Non-tagged forms of CMV 2b protein were shown previously to have incomplete suppressor activity in infiltrated leaves, and to inhibit intercellular spread of a mobile silencing signal [191].

Proteins of the predicted size were detected by immunoblot assay in tissues infiltrated with constructs encoding wild-type or mutant suppressor-GFP fusion proteins (Fig. 2.2A). GFP fluorescence was observed in tissues expressing a wild-type suppressor-GFP fusion protein, while coexpression of mutant suppressor-GFP fusion constructs and dsGFP constructs or resulted in no or low levels of GFP fluorescence (Table 2.1). Coexpression of miR171 and SCL6 constructs without a silencing suppressor or with a mutant suppressor resulted in cleavage of the *SCL6* mRNA (Fig. 2.2B). In tissues coexpressing miR171, SCL6, and a wild-type suppressor, cleavage was inhibited (Fig. 2.2B).

Functionality of the HA-tagged proteins was also tested in side-by-side comparisons to non-tagged forms of each suppressor. Inhibition of silencing by each HA-tagged protein was comparable to inhibition by the corresponding non-tagged form (data not shown).

Developmental abnormalities in Arabidopsis *expressing silencing suppressors*

To determine if induction of developmental phenotypes is a general property of silencing suppressors, the five HA-tagged constructs, as well as empty vector, were introduced

as transgenes into *Arabidopsis* Col-0 plants (Fig. 2.3B). Growth and development parameters were analyzed using a minimum of 58 primary transformants expressing each suppressor. Plants expressing P1/HC-Pro, p21, p19 and CP exhibited moderate to severe defects in leaf and rosette development (Fig. 2.3C). Rosette leaves [growth stage 5.1 according to the scale described by Boyes *et al.* [204]] were narrow, lobed or serrated, or curled (Fig. 2.3C and Table 2.2). Rosette diameter and leaf area were reduced, as were the weight of total aerial tissue and the length of the primary bolt (Fig. 2.4). Plants expressing CP displayed leaf and rosette phenotypes that were generally mild compared to those in plants expressing P1/HC-Pro, p21 and p19. Plants expressing 2b were indistinguishable from vector-transformed plants with respect to leaf morphology (Fig. 2.3C), although modest reductions in aerial tissue weight, leaf area and rosette diameter were detected (Fig. 2.4A).

Plants expressing P1/HC-Pro, p21, p19 and CP also had obvious flower phenotypes and were generally infertile (Figs. 2.3D and 2.4B). They failed to release pollen and, in the case of P1/HC-Pro-expressing plants, had split or non-fused carpels (Table 2.2 and Fig. 2.4B). P1/HC-Pro, p21 and CP expressing plants had narrow and unusually long sepals, while plants expressing p19 had short sepals (Figs. 2.3D and 2.4B, Table 2.2). In all four cases, however, organs in the internal whorls were exposed prior to opening (Fig. 2.3D). At a low frequency, plants expressing P1/HC-Pro, p19 or 2b contained additional trichomes on abaxial or adaxial sepals (Table 2.2). Plants expressing 2b were fertile, although they contained fewer flowers per plant compared to control plants (Fig. 2.4B).

Therefore, the correlation between strong RNA silencing suppressor activity (as measured in the hairpin dsRNA silencing assay) and strong developmental phenotypes in *Arabidopsis* is relatively high. The protein with weak silencing suppressor activity in the transient assay (2b) caused only mild developmental abnormalities in transgenic plants. The data also support the hypothesis that pathogenicity associated with these proteins involves, at least partly, interference with growth and development during virus infection.

Interference with miRNA-guided mRNA cleavage by three silencing suppressors

Developmental abnormalities in *Arabidopsis* plants expressing P1/HC-Pro correlate with inhibition of miRNA-guided target regulation [185]. To determine if this is a general property of strong RNA silencing suppressors, the levels of three mRNAs (*ARF8*, *ARF10* and *SCL6*) that are normally under negative regulation by miRNAs (miR167, miR160 and miR171, respectively) were measured in transgenic plants expressing each of the suppressors and

compared relative to the levels in vector-transformed plants. The *ARF8* and *ARF10* mRNAs accumulated to higher levels in plants expressing P1/HC-Pro, p21, and p19 (Fig. 2.5A and 2.5B, lanes 3-8). Also, the ratio of *SCL6* full-length to 3' cleavage product detected by blot assay was higher in plants expressing P1/HC-Pro, p21 and p19 (Fig. 2.5C, lanes 3-8). In plants expressing CP, only slight increases in *ARF8* and *ARF10* mRNA levels, and no change in the ratio of *SCL6* RNAs, were measured (Fig. 2.5, lanes 9 and 10). Plants expressing 2b showed no increases in the levels or ratios of miRNA targets (Fig. 2.5, lanes 11 and 12).The *ARF8*, *ARF10*, or *SCL6* mRNAs are targeted by three different miRNAs [100, 182, 185]. Increased accumulation of these target mRNAs in the presence of silencing suppressors could conceivably be the result of lower levels of miRNAs, although this would be inconsistent with previous studies [185, 201]. Indeed, there was no consistent decrease among the three miRNAs caused by any of the suppressors (Fig. 2.6).

To assay the genome-wide effects of silencing suppressors, Affymetrix Ath1 expression arrays were used to analyze *Arabidopsis* transcript profiles in plants expressing HC-Pro, p21 and p19. These were compared to expression profiles generated in parallel for plants infected by *Turnip mosaic virus* and plants with defects in the miRNA biogenesis pathway (*dcl1, hen1, hyl1, hst* mutants), tasiRNA pathway (*dcl4, rdr6* mutants), chromatin siRNA pathway (*dcl3, rdr2* mutants), and other RNAi pathways (*dcl2, rdr1* mutants). Methods for statistical analysis of these data were previously described [39, 48, 52, 129]. miRNA target transcripts were generally co-affected in the suppressor lines and miRNA biogenesis mutants (Fig. 2.7A). miRNA target transcripts were generally up in suppressor-expressing and miRNA mutants, but were unaffected in the tasiRNA and chromatin siRNA mutants (Fig. 2.7B). Not all miRNA targets represented on the array were up in the miRNA biogenesis mutant plants. However, those that were unaffected in the miRNA biogenesis mutants were generally unaffected in the suppressor lines. Importantly, most of the up-regulated miRNA target transcripts seen in the miRNA biogenesis mutants and suppressor lines were also up in TuMV-infected plants (Fig. 2.7B). These data strongly support the idea that HC-Pro, p21 and p19 disrupt miRNA-mediated regulation on a genome-wide scale.

Strong RNA silencing suppressors stabilize an intermediate in the miRNA pathway

The P1/HC-Pro, p21 and p19 silencing suppressors interfered with siRNA-guided and miRNA-guided target cleavage but not siRNA or miRNA formation, suggesting that inhibition is likely associated with RISC assembly or RISC activity (Fig. 2.3B-D, Fig. 2.5, Fig. 2.6).

Assembly of RISC complexes containing miRNA requires unwinding of the miRNA/miRNA* duplex intermediate, followed by (or concurrent with) incorporation of miRNA into RISC and degradation of miRNA* [23, 28]. If the strong suppressors inhibit RISC assembly at the point of unwinding miRNA/miRNA*, we predicted that miRNA* species would accumulate specifically in plants expressing these proteins. The miR167b*, miR160c* and miR171* levels were below the detection limit in vector-transformed plants (Fig. 2.6, lanes 1 and 2). However, they were each detected in plants expressing P1/HC-Pro, p21 and p19 (Fig. 2.6, lanes 3-8), with the highest levels of each accumulating in P1/HC-Pro-expressing plants. In contrast, each miRNA* accumulated to relatively low or nondetectable levels in plants expressing CP or 2b proteins (Fig. 2.6, lanes 9-12). Therefore, strong suppression of miRNA-guided target cleavage correlated with accumulation of normally labile miRNA* species, suggesting that the strong suppressors interfered with unwinding of miRNA/miRNA* duplexes.

*Interaction of silencing suppressors and miRNA/miRNA** in vivo *and* in vitro

Tombusvirus p19 dimers bind duplex siRNAs [187, 196, 205]. This led to a competitive inhibition model in which p19 interferes with RISC assembly by sequestering siRNA intermediates [93]. We tested the hypothesis that each of the three strong suppressors interacted with miRNA/miRNA* complexes *in vivo* by co-immunoprecipitation (IP) assays using anti-HA monoclonal antibody. Precipitated complexes from inflorescence (P1/HC-Pro and p21) and total aerial (P1/HC-Pro, p21 and p19) tissues were analyzed for suppressor protein, miRNAs and miRNAs*. As controls, IP assays were done using extracts from vector-transformed plants and using a heterologous monoclonal antibody mixture specific for NIa and NIb proteins of *Tobacco etch virus*.

Each suppressor protein was detected in total extracts (IP inputs) and in immunoprecipitated material using HA antibody, but not in IP fractions using NIa/NIb antibody (Fig. 2.8A and 2.8B, top panel). The input extracts from the suppressor-expressing plants contained each of the miRNAs and miRNAs* (miR167, miR171, miR167b*, miR171*, and miR160c*) analyzed in the respective tissues (Fig. 2.8A, lanes 4 and 7; 2.8B, lanes 4, 7, 10). No miRNAs or miRNAs* were detected in any IP fractions from plants expressing P1/HC-Pro (Fig. 2.8A and 2.8B, lanes 5 and 6). In contrast, each miRNA and miRNA*, but no 5S rRNA, co-immunoprecipitated with p21 in both tissue types (Fig. 2.8A and 2.8B, lane 9). Similarly, miR167, miR167b*, miR171 and miR160c* each specifically co-immunoprecipitated with p19 in aerial tissue extracts (Fig. 2.8B, lane 12), although the proportion of small RNA in the IP

fractions relative to the input extract was less using p19-expressing plants compared to p21-expressing plants. In vector-transformed plant extracts, miR167, miR171 and 5S rRNA were detected (Fig. 2.8A and 2.8B, lane 1), but none were present in the IP fractions (lanes 2 and 3). Faster migrating forms of some miRNAs and miRNAs* were detected in total extracts or IP fractions in the presence of P1/HC-Pro or p19 (Fig. 2.8, lanes 4, 9 and 12). Truncated miRNAs in the presence of p19 were also observed previously by others [70].

To confirm that co-IP of p21 and p19 with miRNA and miRNA* was not a peculiar artefact of the transgenic system, IP assays were also done using extracts from *N. benthamiana* leaves expressing the GFP hairpin RNA construct and each suppressor. siRNAs related to the GFP sequence ranged in size between 21 and 24 nucleotides, as shown previously (Fig. 2.8C, lanes 1, 4 and 7). siRNAs specifically co-immunoprecipitated with both p21 and p19 (Fig. 2.8C, lanes 6 and 9). Only the small size-class (~21 nucleotides) co-immunoprecipitated with p19, as expected from the well-characterized binding properties (Fig. 2.8C, lane 9) [187, 196, 205]. No siRNA was detected in IP fractions from tissue expressing GFP siRNA and P1/HC-Pro, although the efficiency of IP of P1/HC-Pro after transient expression was relatively low (data not shown).

We tested the hypothesis that p21 binds small RNA duplexes directly by electrophoretic mobility shift assays using purified recombinant p21 and synthetic miR171, miR171*, miR171/miR171* duplex, or an siRNA duplex. The miR171/miR171* duplex contained two mismatched positions and two G:U base pairs. In the absence of p21 protein, each single-stranded and duplex RNA migrated to near the bottom of the gel (Fig. 2.8D, lanes 1, 3, 5, 7). In the presence of p21, slower-migrating complexes were detected using the siRNA duplex and miR171/miR171* duplex (Fig. 2.8D, lanes 2 and 8). No p21 complexes were detected using single-stranded miR171 or miR171* (Fig. 2.8D, lanes 4 and 6). In reactions containing 0.1 μM siRNA duplex and 1 μM purified p21, synthetic miR171 was a poor competitor for p21 binding (10.34 μM required for half-maximal inhibition) compared to synthetic miRNA/miRNA* or the siRNA duplex (half-maximal inhibition concentration = 1.28 μM and 0.63 μM, respectively). These data indicate that p21 interacts directly with small RNA duplexes, regardless of whether or not the duplex contains perfectly complementary (siRNA) or mismatched (miRNA) strands. Using DNA oligonucleotides corresponding to a duplex siRNA, low levels of p21 complex were detected [less than 10% of the level detected using the RNA duplexes (data not shown)].

Discussion

Among the five silencing suppressors analyzed, three (P1/HC-Pro, p21 and p19) were characterized as strong due to their effects on siRNA-guided cleavage, miRNA-guided cleavage of target mRNAs, and development. However, we propose that these three suppressors function by at least two distinct mechanisms to arrest the RISC assembly pathway. Although each suppressor inhibited turnover of miRNA* species, a feature we interpret is due to lack of unwinding of miRNA/miRNA* duplexes, only two (p21 and p19) could be detected in a complex with miRNAs and miRNAs* *in vivo*. For both p21 and p19, a direct binding model seems quite likely given the clear duplex siRNA and miRNA binding properties of p21 and p19. Sequestration of miRNA/miRNA* duplexes may occur in the cytoplasm after processing by DCL1 in the nucleus and subsequent nucleocytoplasmic transport [15]. Cytoplasmic localization of p21 is in accord with this hypothesis [186].

Although P1/HC-Pro was the most effective miRNA pathway suppressor tested, we obtained no evidence that it interacts with miRNAs or miRNAs* *in vivo*. However, it clearly inhibited miRNA* turnover, again suggesting that miRNA/miRNA* unwinding and RISC assembly was suppressed. One possibility is that the interaction between P1/HC-Pro and miRNA/miRNA* was not preserved in the experimental conditions used for these immunoprecipitation experiments. Additional experiments are necessary to determine whether P1/HC-Pro interacts directly with small RNA duplexes (see Chapter 3).

Among the other two silencing suppressors analyzed in this study, TCV CP functioned effectively as an inhibitor of siRNA formation but only weakly as a miRNA pathway suppressor. This suppressor affects DCL4 activities required for production of TCV-derived siRNA [73] and trans-acting siRNA [81, 159] but not DCL1 required for miRNA processing. Nevertheless, CP had small but measurable effects on miRNA-directed target cleavage, which may partly explain the developmental consequences of CP expression in *Arabidopsis*. The CMV 2b protein had very little effect on miRNA-guided functions and development in *Arabidopsis*, and may function by mechanisms that are quite distinct from the others analyzed [190, 191].

Finally, is there a physiologic advantage to viruses that interfere with miRNA-guided gene regulation? If miRNAs are required for defense responses or for expression of genes required for susceptibility to a broad range of viruses, then evolution of virus-encoded miRNA inhibition functions can be easily rationalized. However, in view of the functions of known miRNA target genes in plants, and the lack of effects of *dcl1* (miRNA-deficient) mutations on virus susceptibility (Z. Xie and J.C. Carrington, unpublished data), no evidence for such

requirements is available. It seems more likely that interference with the miRNA pathway by suppressors such as P1/HC-Pro, p21 and p19 is a consequence of inhibition of shared steps in the silencing pathways involving siRNAs for antiviral defense and miRNAs for development.

Experimental Procedures

Genes and constructs

Suppressors were derived from the following viruses: P1/HC-Pro, *Turnip mosaic virus*; p21, *Beet yellows virus*; p19, *Tomato bushy stunt virus*; CP, *Turnip crinkle virus*; 2b, *Cucumber mosaic virus* strain Q. Coding sequences for each suppressor were amplified by PCR using a 3' primer that added a carboxy-terminal HA epitope (primer sequences available upon request). Resulting DNA fragments were cloned into a modified pCB-302 plant transformation vector [206] using *Nco*I and *Xba*I sites. Suppressor constructs contained a *Cauliflower mosaic virus* 35S promoter and a 5' nontranslated leader sequence from *Tobacco etch virus* [207]. Resulting constructs were introduced into *Agrobacterium tumefaciens* strains GV2260 and GV3101.

Agrobacterium-infiltration assays

Transient silencing and cleavage assays in *Nicotiana benthamiana* were done as described [199]. Cultures of *A. tumefaciens* GV2260 were injected at the following concentrations: *35S:GFP*, OD_{600} = 0.5; *35S:dsGFP-FAD2*, OD_{600} = 0.1; *35S:P1/HC-ProHA* or other suppressor constructs, OD_{600} = 0.5 (silencing assays); *35S:SCL6,* OD_{600} = 0.1; *35S:MIR171A,* OD_{600} = 0.75; *35S:P1/HC-ProGFPHA* or other suppressor constructs, OD_{600} = 0.15 (cleavage assays). In assays lacking one or more of these components, strains containing *35S:GUS* or *35S:vector* were used to normalize each injection to a constant OD_{600} = 1.0. Infiltrated tissues were harvested 48 hours post-injection for photography and protein and RNA analyses. GFP fluorescence was visualized using long-wavelength ultraviolet light.

Protein and RNA blot analysis

Protein extracts were prepared and normalized for SDS-PAGE using the Bradford assay (Bio-Rad). Immunoblot analysis of total protein samples (10 μg) was done using anti-HA-peroxidase conjugate (Roche). Total RNA was extracted from independent pools of leaf or inflorescence tissues using Trizol reagent [199]. Low-molecular weight RNA was isolated with RNA/DNA Midi Kits (Qiagen). Blot hybridization of normalized total or low-molecular weight

RNA (5 µg) was done as described [100] and hybridization intensities were quantified using a Phosphorimager or Scanning Densitometer (Molecular Dynamics). ^{32}P-Radiolabeled probes for mRNAs were synthesized by random-priming of cloned smGFP or genomic sequences [208]. ^{32}P-Radiolabeled miRNA probes were produced by end-labeling of complementary oligonucleotides. Sequences for miRNA* probes were predicted from miRNA precursor structures [72, 100]. Accumulation of miRNA-targeted mRNAs was normalized to levels of mRNA from the control gene *TyrAT* (At2g20610). Accumulation of *SCL6* full-length mRNA ('a' form) relative to the 3' cleavage product ('b' form) was represented as a ratio as described [100].

Transgenic plants

Arabidopsis thaliana Col-0 plants were transformed by the vacuum-infiltration method [209] using *A. tumefaciens* GV3101 carrying constructs for expression of epitope-tagged P1/HC-Pro, p21, p19, CP, 2b, or the empty expression vector. Seed from primary transformants was grown under selection for phosphoinothricin resistance in a standard greenhouse.

Phenotypic analysis of transgenic Arabidopsis

Phenotypic effects of suppressor expression were quantified at defined growth stages: vegetative parameters, stages 5.00 – 5.10; reproductive defects, stages 6.10 – 6.90; fertility, stages 8.00 – 9.70 [204]. Mean area of third and fourth rosette leaves was determined using a Licor Model 3100 area meter. Sepals and other floral organs were measured with a Mitutoyo Model CD-6"CS digital calipers. All data were subjected to statistical analyses using the Student's t-test.

Microarray analysis

Inflorescence tissue was collected and RNA was extracted and labeled as described previously [158]. Raw intensity values from ATH1 GeneChip arrays (Affymetrix) were normalized using RMA Express [210] and imported into Genespring v7 (Silicon Genetics) for analysis. Two sets were selected from genes represented on the ATH1 array: 1) validated or high-quality predicted miRNA target genes (see Appendix [142, 169, 182, 211]) that were significantly different in *dcl1-7* mutants than in wild-type plants (p<0.01); 2) validated or high-quality predicted miRNA target genes that were not significantly different in *dcl1-7* than in wild-type (0.8>p<0.2). Gene set 1 contained 21 genes; gene set 2 contained 33 genes. Significance

analysis for selection of each gene set incorporated a false discovery rate of 0.05 [212]. Hierarchical clustering for each gene set was done using standard clustering within GeneSpring.

Immunoprecipitation

Aerial or inflorescence tissue from transgenic *Arabidopsis* plants (5 weeks old), and leaf tissue from *Agrobacterium*-infiltrated *N. benthamiana*, were ground under liquid nitrogen and homogenized in 5 ml/g lysis buffer (50 mM Tris-HCl, pH 7.4, 100 mM KCl, 2.5 mM MgCl2, 0.1% NP-40, 2X Complete protease inhibitor cocktail [Roche]). Cell debris was pelleted by centrifugation for 15 min at 9500 x g. The clarified lysate was precleared for 20 min at 4°C with 10 µl bed volume protein A-agarose (30 µg protein A) per ml. Precleared lysates were reacted with 4 µg anti-HA (Roche) or anti-NIa/NIb [213] per ml for 1 hour at 4°C, then with 50 µl bed volume protein A-agarose (150 µg protein A) per ml for 3 hours at 4°C. Precipitates were washed 3 times in lysis buffer and divided for protein and RNA analysis. Nucleic acid was recovered by treatment with 3 vol proteinase K solution (100 mM Tris-HCl pH 7.4, 10 mM EDTA, 150 mM NaCl, 2% SDS, 0.2 µg/µl proteinase K) for 15 minutes at 65°C, extraction with saturated phenol and phenol:chloroform, and ethanol precipitation. For miRNA and miRNA* blot assays, 5 µg of RNA recovered from the input extract, or RNA from immunoprecipitation fractions representing 150 mg tissue, was used. Five µg of RNA recovered from input extracts corresponded to the equivalent of approximately 5 mg tissue. As a control, blots were stripped and rehybridized with an oligonucleotide probe specific to 5S rRNA.

For IP assays using *Agrobacterium*-infiltrated leaves, GFP silencing was induced as described [199]. *A. tumefaciens* cultures were injected at the following concentrations: *35S:dsGFP-FAD2*, OD_{600} = 0.1; *35S:vector*, *35S:P1/HC-Pro*, *35S:p21* or *35S:p19*, OD_{600} = 0.9. Infiltrated tissue was harvested and processed 72 hrs post-injection. ^{32}P-Radiolabeled probes for siRNAs were synthesized by random-priming of cloned smGFP sequence [199].

Electrophoretic mobility shift assays

The p21 coding sequence was amplified by PCR using a 5' primer that added an amino-terminal hexahistidine tag (primer sequences available upon request). Resulting DNA fragments were cloned into pET16b (Novagen) using *Nco*I and *BamH*I sites. The resulting construct was introduced into *E. coli* strain BL21(DE3) (Novagen), and protein was expressed and purified under native conditions using Ni-NTA resin (Qiagen) following the recommendations of the manufacturer.

Probes were synthesized by end-labeling RNA oligonucleotides miR171 (5' UGAUUGAGCCGCGCCAAUAUC 3'), miR171* (5' UAUUGGCCUGGUUCACUCAGA 3'), siRNA (5' CGUACGCGGAAUACUUCGAUU 3') or siRNA* (5' UCGAAGUAUUCCGCGUACGUG 3') (Dharmacon) using [^{32}P]ATP. Duplexes were formed during annealing reactions similar to those described by others [187]. Formation of duplexes was confirmed by electrophoresis mobility assays. p21 complex formation reactions contained 1 μM p21 and 0.1 μM oligonucleotide in 10 μl binding buffer [0.1 M KCL, 25 mM HEPES, 10 mM DTT, pH 7.6 [205]] and were done at 23˚C for 15 min. Competition reactions contained 0.1 μM labeled oligonucleotide and 0.5 μM, 25 μM and 125 μM unlabeled oligonucleotide.

Table 2.1. Suppression of dsGFP-induced silencing by suppressor-GFP fusion proteins.

	GFP fluorescence score	
	leaf 1	leaf 2
vector	-	-
vector + GFP	+	+
vector + dsGFP	-	-
vector + dsGFP + GFPHA	-	-
vector + dsGFP + GUSGFPHA	-	-
vector + dsGFP + P1/HC-ProGFPHA	+	+
vector + dsGFP + P1/HC-Pro(AS9)GFPHA	-	-
vector + dsGFP + p21GFPHA	++	++
vector + dsGFP + p21(8A)GFPHA	-	-
vector + dsGFP + p19GFPHA	+	+
vector + dsGFP + p19(W39/42R)GFPHA	-	-

*Rating scale:

- no visible fluorescence

+ visible GFP fluorescence

++ strong fluorescence

Table 2.2. Frequency of vegetative and reproductive defects in populations of transgenic *Arabidopsis* plants.

	# plants showing leaf defects				# flowers showing sepal defects					# flowers showing carpel defects	
genotype	n[a]	narrow	lobing	curling	n	narrow sepals	short sepals	long sepals	additional trichomes	n	unfused carpels
vector	79	0	1	0	20	0	0	0	0	20	0
35S:P1/HC-ProHA	83	75	75	10	20	20	0	20	6	20	20
35S:p21HA	199	50	88	115	20	19	0	20	0	20	0
35S:p19HA	129	0	117	0	30	8	27	0	4	30	0
35S:CPHA	131	71	71	5	20	8	0	18	0	20	0
35S:2bHA	58	0	1	0	30	0	0	0	8	30	0

[a]total number of plants or flowers analyzed

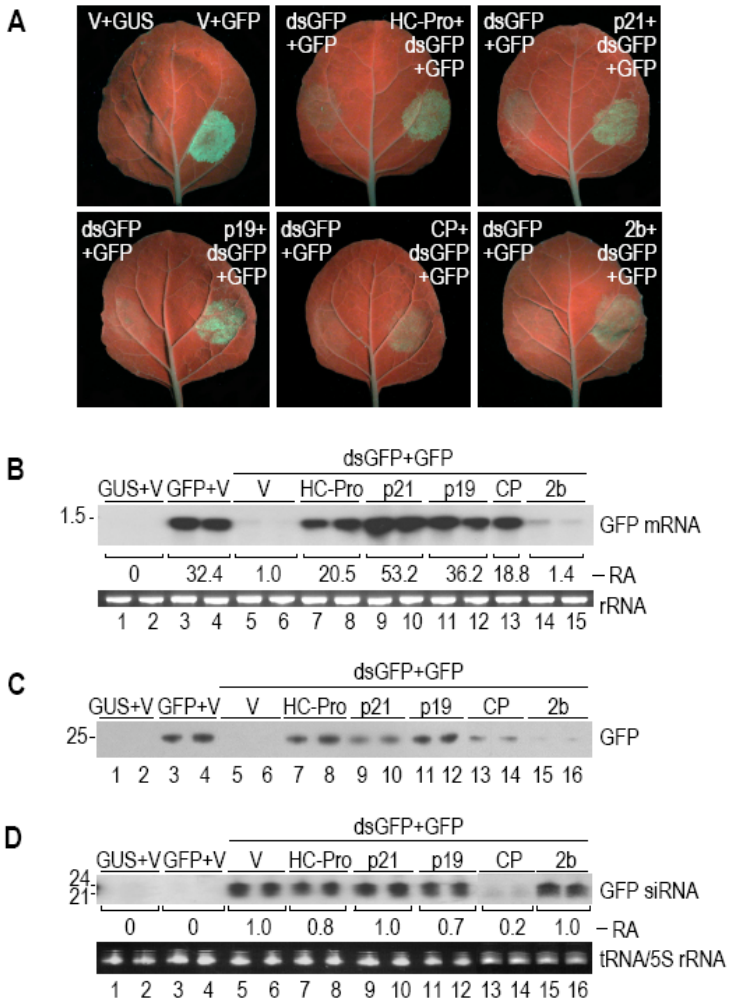

Fig. 2.1. Activity of epitope-tagged silencing suppressors. (A) Leaves of N. benthamiana infiltrated with cultures of Agrobacterium carrying empty vector (V) or genes for expression of ß-glucuronidase (GUS), GFP, dsGFP, and HA-tagged silencing suppressors. (B) GFP mRNA accumulation in co-Agro-infiltrated tissues. Blot hybridization was done on total RNA samples prepared from leaf tissues that were infiltrated with the Agrobacterium strains shown in (A). Mobility position of the 1.5 kb RNA standard is shown. Average accumulation (RA) of GFP signal in duplicate samples, relative to that in lanes 5 and 6 (dsGFP-FAD2 + GFP) is shown. Accumulation of GFP mRNA in a single sample from CP-infiltrated tissues is indicated (lane 13). (C) GFP protein accumulation in co-Agro-infiltrated tissues. Immunoblot analysis was done on total protein samples prepared from infiltrated tissues described in (A). (D) GFP-specific siRNA accumulation in co-Agro-infiltrated tissues. Blot hybridization was done using duplicate low molecular weight RNA samples prepared from the infiltrated tissues described in (A). Mean accumulation (RA) of signal in duplicate samples, relative to that in the control (lanes 5 and 6) is shown.

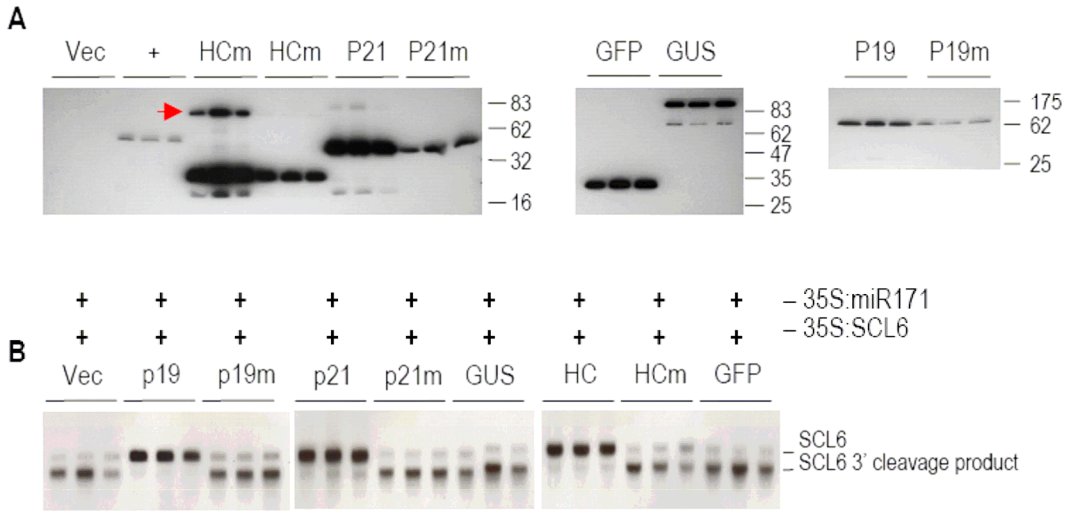

Fig. 2.2. Expression and activity of silencing suppressor-GFP fusion proteins. (A) Protein accumulation in Agrobacterium-infiltrated N. benthamiana leaves. Immunoblot analysis was done on total protein samples prepared from tissues infiltrated with cultures of Agrobacterium carrying empty vector (Vec), a positive control HA-tagged protein (+) or genes for expression of HA-tagged GFP, a ß-glucuronidase-GFPHA fusion (GUS), or silencing suppressor-GFP fusion proteins P1/HC-Pro (HC), p21, and p19 (GFP fusion variant for each) or mutant suppressors (m). Mobility positions of protein standards are shown. HC-Pro is subject to degradation under certain extraction conditions; the position of the predicted full-length protein is marked. For all other proteins, the predominant band migrated at the position predicted for each fusion protein. (B) P1/HC-Pro (HC), p21, and p19 (GFP fusion variant for each) inhibit miR171-guided activity on Scarecrow-like6 (SCL6) mRNA in the transient infiltration assay. These assays involve co-expression of miR171 precursor, SCL6 target mRNA, and suppressors. Active, but not mutant, suppressors inhibit miR171-guided cleavage of SCL6 mRNA.

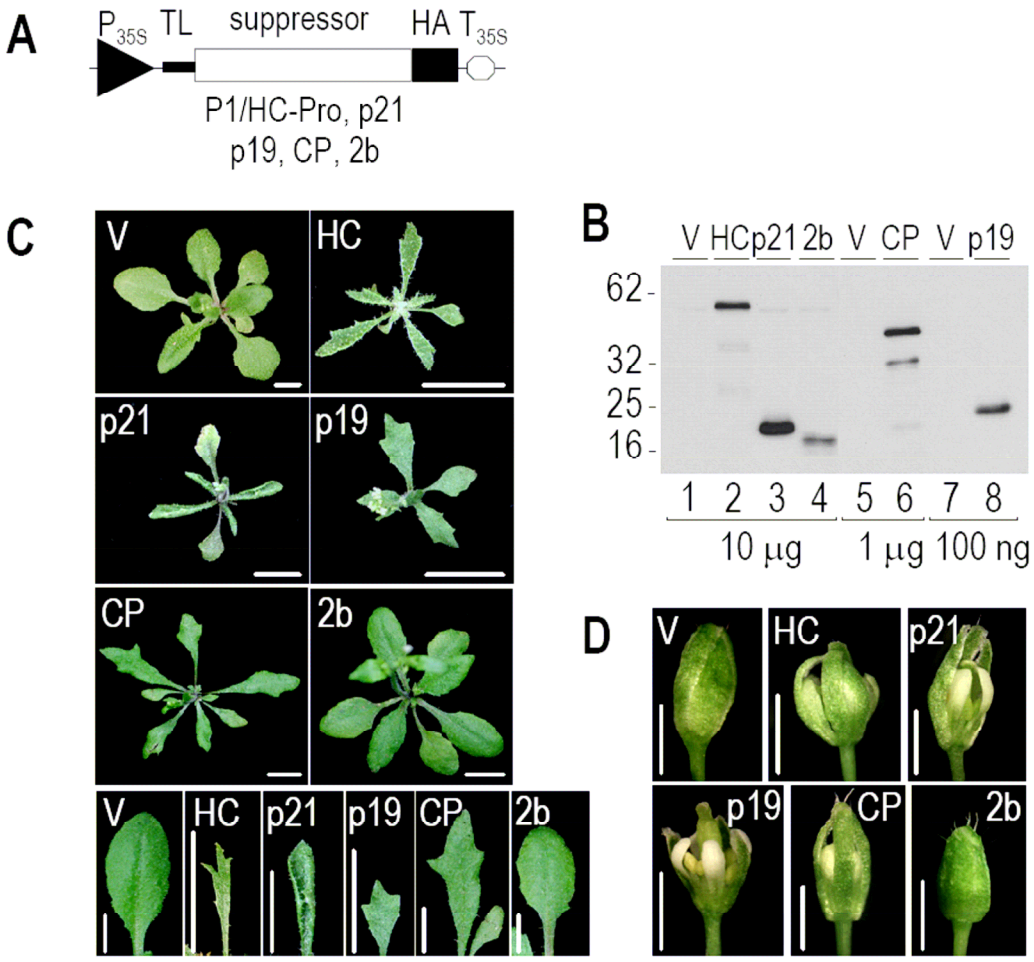

Fig. 2.3. Developmental defects induced by silencing suppressors from five viruses. (A) Diagram of constructs used for expression of HA epitope-tagged viral suppressor proteins. Each construct contained the gene for one of the suppressors shown. P35S = 35S promoter from Cauliflower mosaic virus; TL = translational leader element from Tobacco etch virus; T35S = 35S terminator from Cauliflower mosaic virus. (B) Expression of epitope-tagged silencing suppressors in transgenic *Arabidopsis* plants. Immunoblot analysis was done using total protein samples (amounts shown) from plants transformed with empty vector (V) or constructs encoding P1/HC-Pro (HC), p21, 2b, CP or p19. Mobility positions of 16 – 62 kDa protein standards are shown. Note that HC-Pro migrates as ~50 kDa protein, but that breakdown fragments migrate at ~25 kDa and ~35 kDa positions. (C) Effects of silencing suppressors on leaf and rosette morphogenesis. Scale bars represent 10 mm. (D) Effects of silencing suppressors on flowers (stage 11-12). Scale bars represent 1 mm.

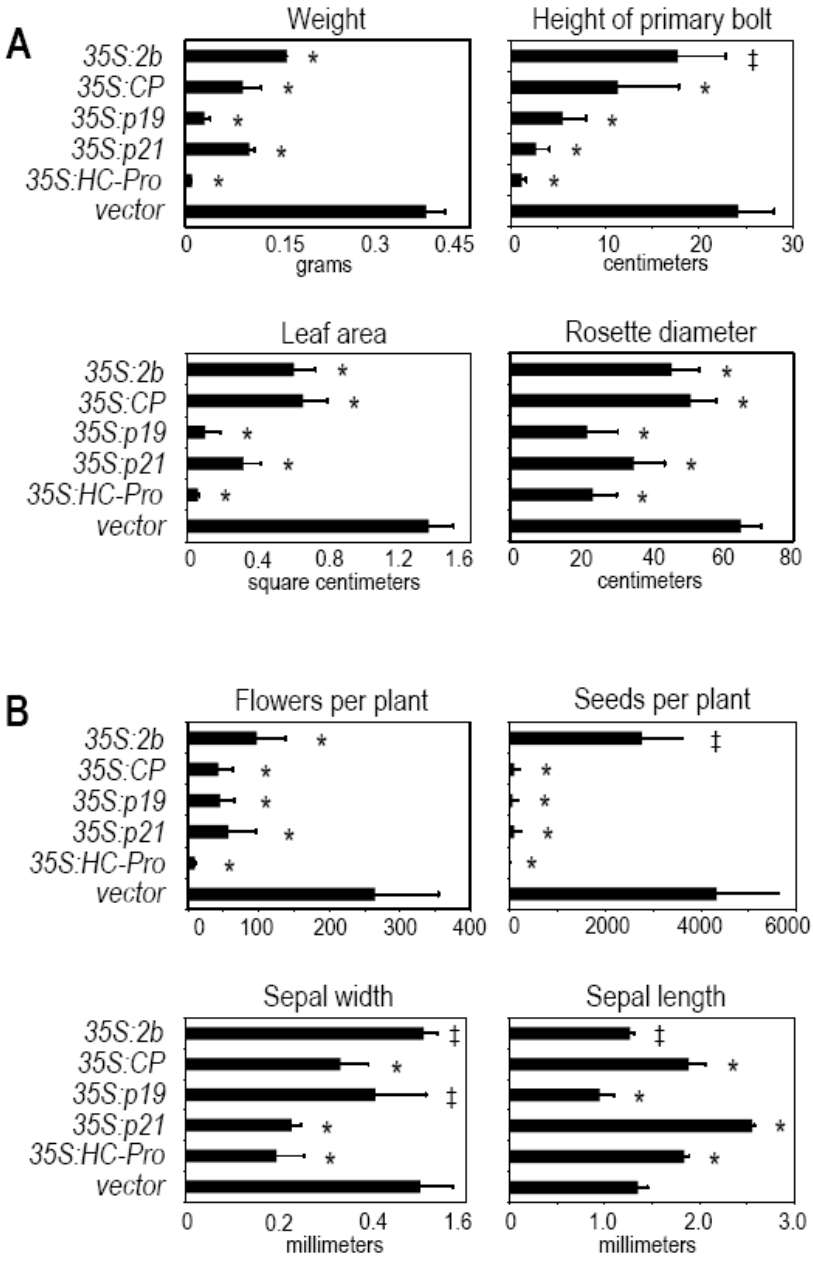

Fig. 2.4. Quantitation of defects in transgenic *Arabidopsis*. (A) Vegetative defects in vector-transformed and suppressor-expressing *Arabidopsis* plants. Mean values for all parameters are represented on each graph with standard deviations shown as error bars. p-values are represented as follows: ‡, $p > 0.05$; *, $p \leq 0.05$. (B) Reproductive defects in vector-transformed and suppressor-expressing *Arabidopsis* plants. Standard deviations and p-values are represented as in (A).

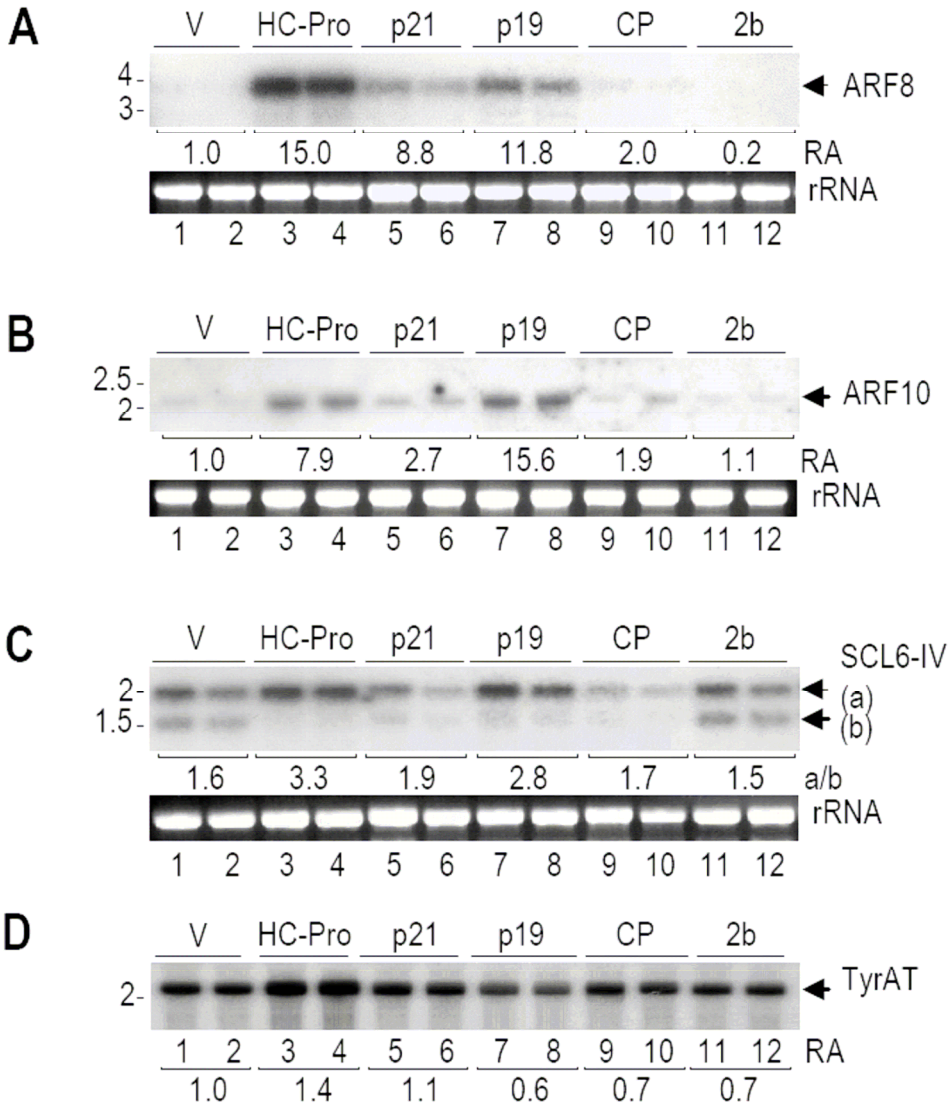

Fig. 2.5. Blot analysis of miRNA targets in transgenic *Arabidopsis* plants. RNA samples from vector-transformed (lanes 1 and 2) and suppressor-expressing transgenic *Arabidopsis* plants (lanes 3 - 12) were analyzed in duplicate by hybridization with DNA probes. Ethidium bromide-stained 28S rRNA is shown for each blot. (A) Accumulation of ARF8 (At5g37020) mRNA. Mean relative accumulation (RA) of ARF8 mRNA in suppressor-expressing plants was calculated relative to that in vector-transformed plants, normalized against the accumulation of the control TyrAT (At2g20610) mRNA. (B) Accumulation of ARF10 (At2g28350) mRNA. Mean relative accumulation (RA) of ARF10 mRNA was calculated as in (A). (C) Expression of full-length SCL6 (At4g00150) mRNA. The mean ratio of full-length SCL6 mRNA (a) to the cleavage product (b) is shown. (D) Expression of TyrAT (At2g20610) mRNA. This blot was first used to analyze ARF10 expression, then stripped and reprobed.

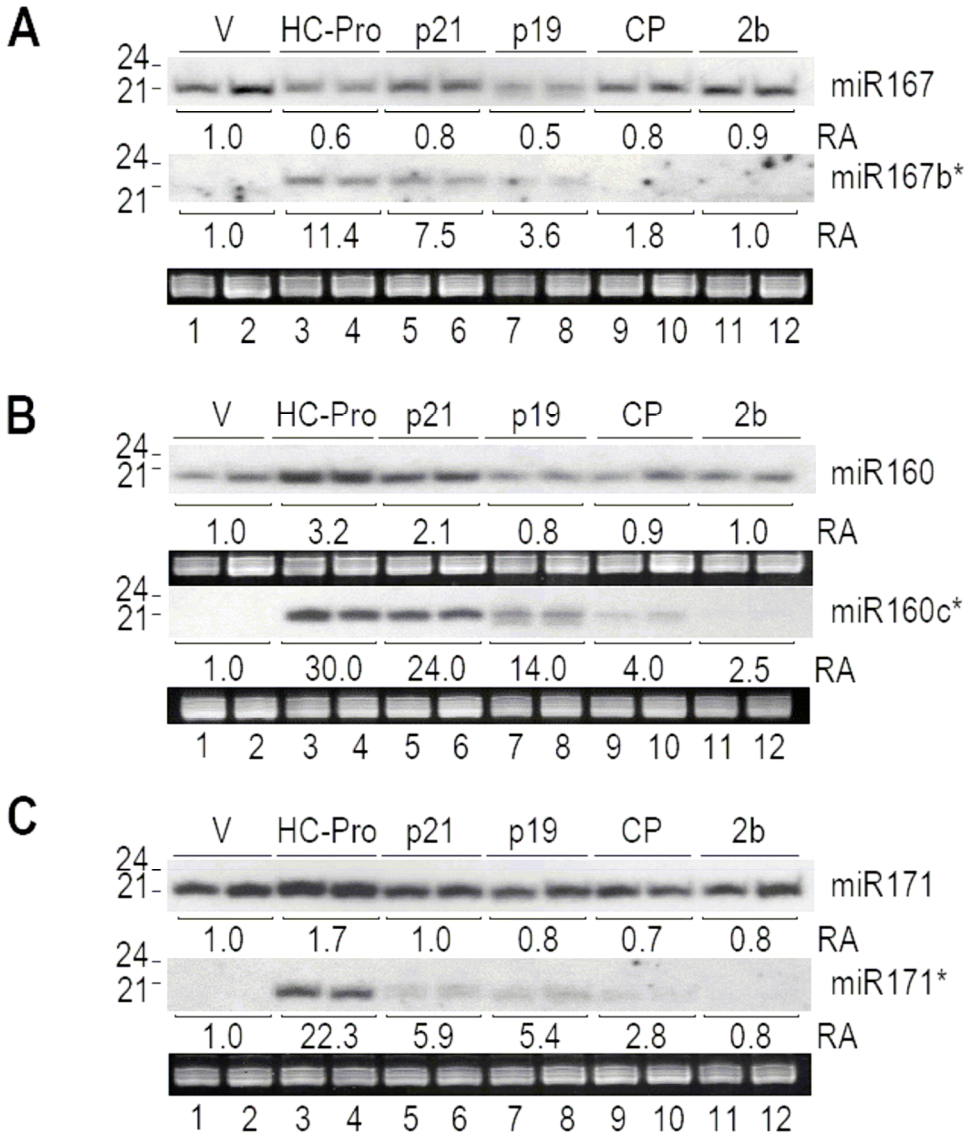

Fig. 2.6. miRNA and miRNA* accumulation in transgenic *Arabidopsis* plants. Small RNA samples from vector-transformed (lanes 1 and 2) and suppressor-expressing transgenic *Arabidopsis* plants (lanes 3 - 12) were analyzed in duplicate by hybridization with oligonucleotide probes. Ethidium bromide-stained tRNA and 5S rRNA is shown below each blot, and mobility positions of 24- and 21-nucleotide RNA size standards are shown. Mean relative accumulation (RA) of miRNA or miRNA* signal relative to that in vector-transformed plants (lanes 1 and 2) is shown. (A) Accumulation of miR167 and miR167* from the MIR167B locus. (B) Accumulation of miR160 and miR160* from the *MIR160C* locus. (C) Accumulation of miR171 and miR171*.

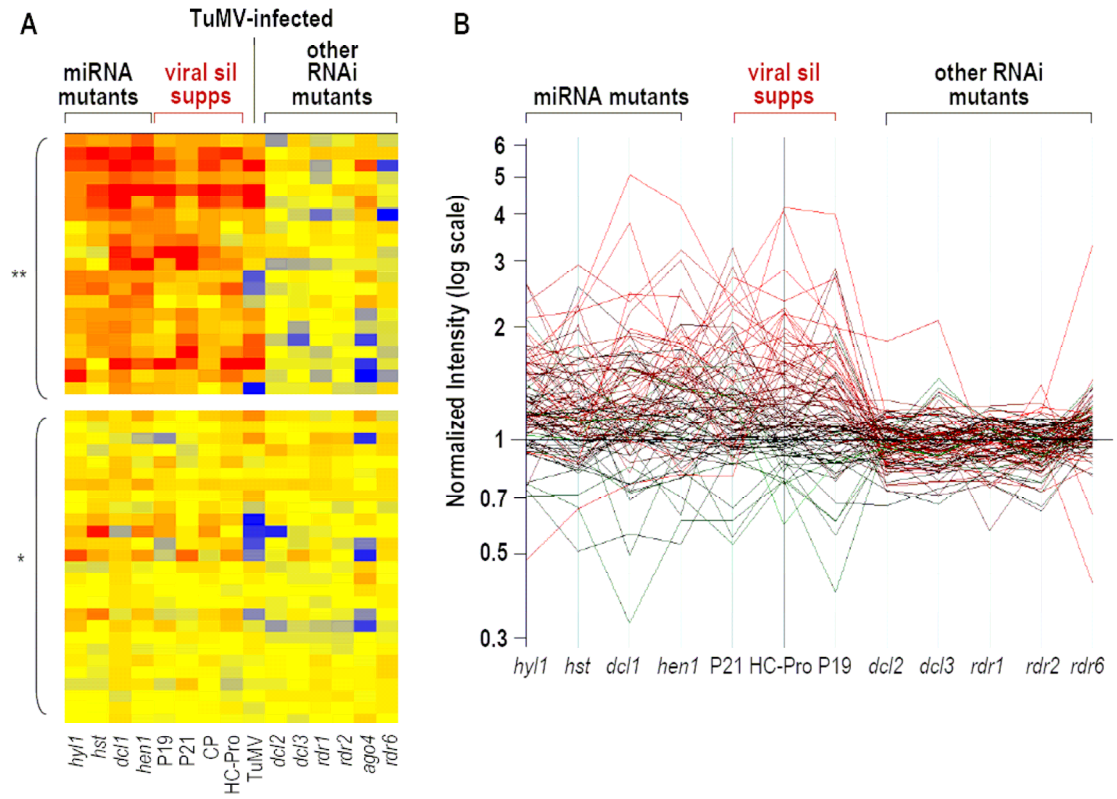

Fig. 2.7. MicroRNA target expression in transgenic *Arabidopsis* plants. (A) miRNA target transcripts are generally co-affected in suppressor-expressing plants (viral sil supps), TuMV-infected plants, and miRNA biogenesis mutants. **, miRNA targets differentially expressed in dcl1 and wild-type (P>0.01); *, miRNA targets not differentially expressed in dcl1 and wild-type. Red, up-regulated; Blue, down-regulated; Yellow, neutral. (B) Expression of all validated miRNA target transcripts in three suppressor-expressing lines and four miRNA-defective mutants, and other small RNA-defective mutants.

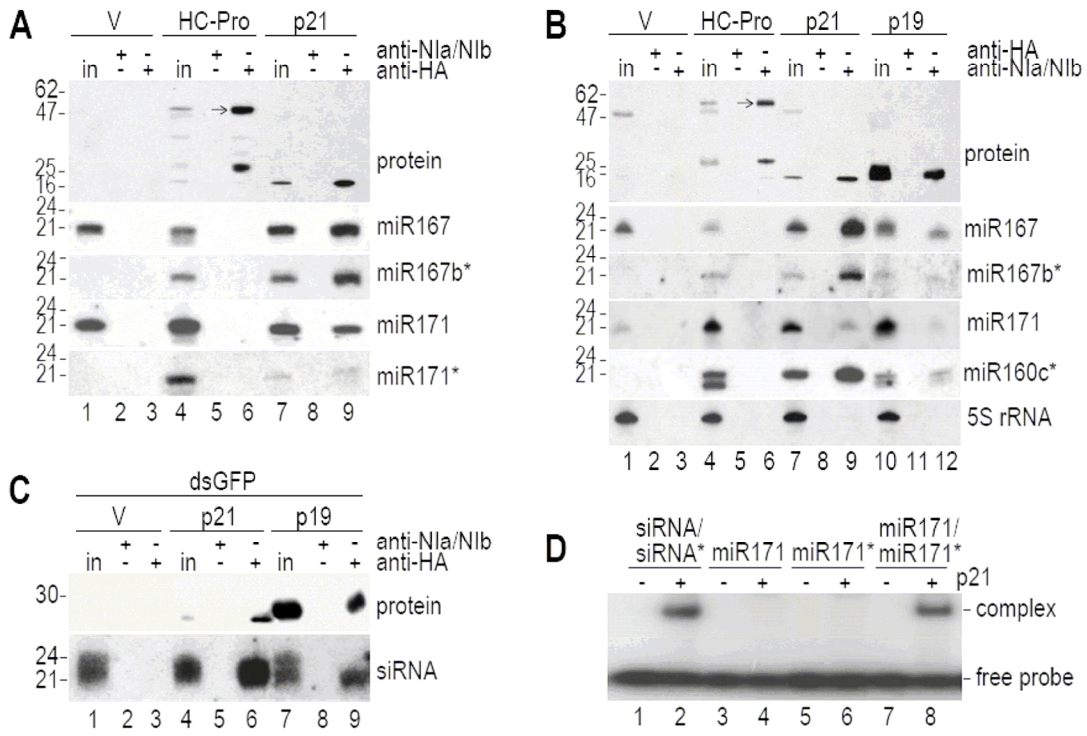

Fig. 2.8. Interactions between suppressor proteins and small RNAs. (A) Co-IP of suppressor proteins, miRNAs, and miRNAs* from transgenic *Arabidopsis* inflorescence tissue. Immunoblot analysis was done using IP input (in) samples and immunoprecipitated fractions using HA or NIa/NIb monoclonal antibodies. Mobility positions of 16 – 62 kDa protein size standards are shown. Blot hybridization assays were done using RNA recovered from immunoprecipitates. Mobility positions of 21- and 24-nucleotide RNA size standards are shown. (B) Co-IP of suppressor proteins, miRNAs, and miRNAs* from transgenic *Arabidopsis* total aerial tissue. (C) Co-IP of suppressor proteins and GFP-specific siRNAs from Agrobacterium-infiltrated *N. benthamiana* leaf tissue. (D) Electrophoretic mobility shift assays using purified p21 and siRNA duplex, miR171, miR171*, and miR171/miR171* duplex. Complexes were analyzed by native PAGE. The positions of p21:RNA complexes and free probes are shown.

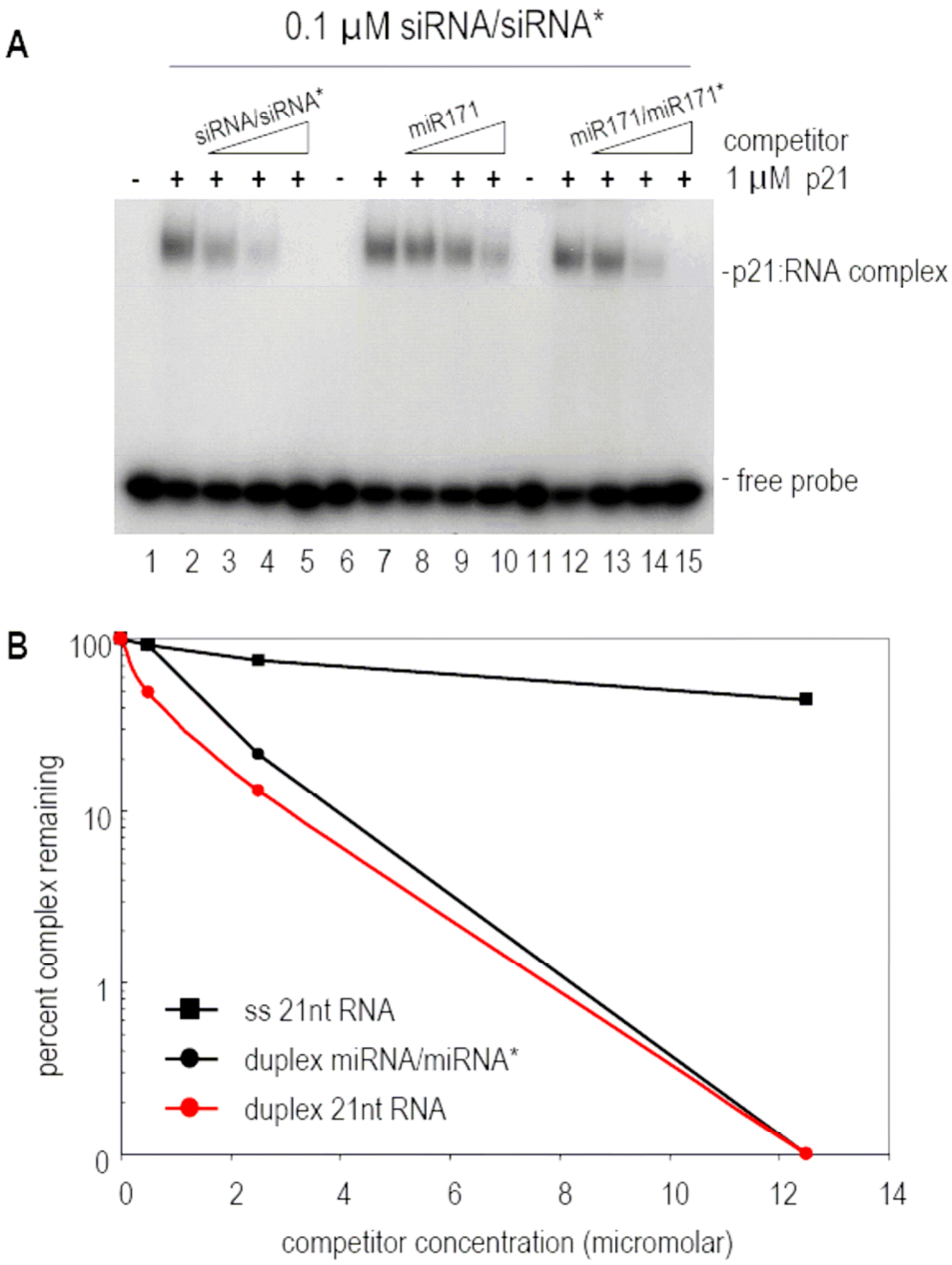

Fig. 2.9. Interactions between p21 protein and small RNAs. Electrophoretic mobility shift assays used purified p21 and siRNA duplex, and unlabeled competitors miR171, miR171/miR171* duplex, siR171/siR171* duplex, a DNA version of miR171, and a DNA version of siR171/siR171* duplex. (A) Complexes were analyzed by native PAGE. The positions of p21:siRNA complexes and free probes are shown. (B) Percent of p21:siRNA complex remaining in the presence of the unlabeled competitors shown in (A).

CHAPTER 3

VIRAL RNA SILENCING SUPPRESSORS INHIBIT ENDOGENOUS AND ANTIVIRAL SILENCING PATHWAYS BY SMALL RNA DUPLEX SEQUESTRATION

Elisabeth J. Chapman, Lorant Lakatos, Bin Yu, Zhiyong Yang, Yu-Ping Liu

The EMBO Journal
Meyerhofstrasse 1, 69117 Heidelberg, Germany

FEBS Letters
Heidelberg University Biochemistry Centre (BZH), Im Neuenheimer Feld 328, D-69120 Heidelberg, Germany
Reprint permission requested
Volume 580 pages 3117-3120

Summary

RNA silencing (RNAi) is an evolutionarily conserved system that functions as an antiviral mechanism in higher plants and insects. To counteract antiviral RNAi, viruses express silencing suppressors that interfere with siRNA- and microRNA-guided silencing pathways. Here we used comparative *in vitro* and *in vivo* approaches to further characterize the suppression mechanisms of three evolutionarily distinct suppressors. Each suppressor inhibited RISC-mediated target cleavage *in vitro* by preventing RNA silencing initiator complex formation. HC-Pro, like p21 and p19, interacted *in vivo* with virus-derived and endogenous small RNA duplexes and *in vitro* with synthetic small RNA duplexes. Each suppressor inhibited HEN1-dependent miRNA/miRNA* duplex methylation and interacted with miRNA/miRNA* duplexes *in vivo*. These findings strongly suggest that viral RNA silencing suppressors inhibit endogenous and antiviral RNAi pathways by sequestration of small RNA duplexes.

Introduction

Although RNA silencing operates through diverse pathways in plants and animals, canonical RNA silencing depends on small RNA that are typically 21-24 nucleotides in length, that arise from defined genetic loci or from invasive agents, and that typically function as sequence-specific regulators through well-established base-pairing rules [15, 17, 214, 215]. microRNAs (miRNAs) or short-interfering RNAs (siRNAs) arise from imperfectly base-paired foldback structures or from dsRNA precursors, respectively. Multidomain RNaseIII-type enzymes, termed Drosha and Dicer (DCR, animals) or DICER-LIKE (DCL, plants), catalyze processing of precursors to small RNA duplexes, containing 2-base overhangs at each 3' end, through multistep processes [15, 216]. Precursor processing also requires additional dsRNA-binding proteins [30-33, 217, 218]. Small RNAs associate with ARGONAUTE (AGO) proteins in effector complexes (also called RISC) to guide target RNA cleavage, translational repression, or chromatin modification [219, 220]. Loading of effector complexes involves molecular sensing mechanisms that assess the base-paired properties of the duplex [23, 28, 221]. In extracts from *Drosophila* embryos, DICER2-R2D2 (DCR2-R2D2) complexed with duplex siRNA serves as an initiator of RISC assembly. This complex then interacts with an AGO2-containing multi-protein complex and cleaves the passenger strand of the siRNA duplex to form the single-stranded (ss) siRNA-containing 80S holo-RISC, which catalyzes sequence-specific cleavage of target RNA [29, 129, 221]. ARGONAUTE proteins contain an RNaseH-related domain that catalyzes

cleavage of target RNAs, provided that the small RNA-target duplex contains sufficient complementarity [110, 222-224].

RNA silencing is active in many biological processes, including chromatin remodeling, development, stress responses and antiviral defense. While gene families encoding common components (AGO, DCR/DCL) of small RNA biogenesis and effector complex machineries are conserved across eukaryotic kingdoms, small RNA pathways have diversified and specialized between and within kingdoms. This is most evident in plants such as *Arabidopsis thaliana*, which contains four DCL, ten AGO, and at least three functional RDR genes [216, 225, 226]. For example, *Arabidopsis* contains four DCLs yielding miRNA (DCL1, 21 nt) and several siRNA classes (DCL4, 21 nt tasiRNA; DCL2, 22 nt siRNA; DCL3, 24 nt siRNA).

Heterochromatin-associated siRNAs (24 nt) form through the activities of RDR2, RNA polymerase IV (Pol IVa and Pol IVb), and DCL3, and require AGO4 for activity to direct or reinforce transcriptionally silent chromatin [47, 48, 59, 60, 73, 125, 143, 144, 227]. This system functions preferentially on loci with repeat sequences, such as transposons, retroelements, and 5S rRNA arrays. A frequent consequence of the RDR2/Pol IV/DCL3/AGO4 system is cytosine methylation of DNA (CNG and CNN contexts) and histone H3 methylation at Lys9 [47, 48, 60, 73, 143, 144, 227], both of which are associated with repressive chromatin states. Although the nature of chromatin-associated factors that interface with the RDR2/Pol IV/DCL3/ AGO4 system have been identified through genetic approaches [228], the mechanisms whereby this system triggers or reinforces repressive marks are poorly understood.

MicroRNAs (miRNAs) are 20- to 24 nucleotide (nt) RNAs that function as sequence-specific regulators of gene expression through translational repression and/or transcript cleavage [15]. miRNA/miRNA* duplex intermediates are formed by DICER or DICER-LIKE enzymes, which generate products with 2 nt 3’ overhangs containing 5’ monophosphate and 3’-OH groups. The miRNA strand is incorporated into the RNA-induced silencing complex (RISC), while the passenger miRNA* strand is ejected and degraded [15]. The 2’-OH of the 3’-terminal nucleotide of miRNAs and siRNAs is methylated in *Arabidopsis* by a methyltransferase, HUA ENHANCER1 (HEN1) [25-27, 229]. The reduced accumulation and heterogeneity in size of miRNAs, as well as loss of miRNA function, in *hen1* mutants [33, 71, 73, 230] reflect the importance of miRNA methylation in plants. The size increase of small RNAs in the *hen1-1* mutant is due to the addition of one to five U residues to the 3' ends of the small RNAs by a novel uridylation activity targeting the 3' ends of unmethylated miRNAs and siRNAs [26]. Therefore, one role of small RNA methylation is likely protecting 3' ends of the small RNAs

from uridylation activity. HEN1 uses miRNA/miRNA* duplexes as substrates *in vitro* [27, 229], but it is not known whether it also acts on miRNA/miRNA* duplexes *in vivo*.

Most plant miRNA associate with AGO1 to guide cleavage or non-degradative repression of target mRNAs in *trans* [21, 72, 100, 182, 185, 231]. *Arabidopsis* contains over 150 well-documented *MIRNA* genes in at least 85 families [21, 142, 147, 169]. Approximately one-half of the conserved miRNA families target of *Arabidopsis* mRNAs encoding transcription factors, such as AUXIN RESPONSE FACTORS (ARFs), involved in growth and development [21]. Other target classes include mRNAs encoding several miRNA and siRNA metabolic factors (DCL1, AGO1, AGO2), metabolic enzymes and factors (e.g. ATP sulfurylase and sulfur transporters), RNA binding proteins, and other proteins [21, 126, 158, 169, 182, 211]. Additionally, three miRNA families target primary transcripts for tasiRNA [142, 158].

tasiRNA function like miRNA to negatively regulate mRNA target transcripts through site-specific cleavage [84, 142, 157]. Eight tasiRNA-generating loci belonging to four families (*TAS1, TAS2, TAS3, TAS4*) are currently known, and collectively they target mRNAs encoding transcription factors, RNA binding proteins, and proteins of unknown function [84, 157-159, 232]. tasiRNA originate from primary Pol II transcripts that are site-specifically processed by miRNA-guided cleavage. In one family (*TAS3*), the primary transcript interacts with two miRNA-AGO complexes, both of which are required for tasiRNA biogenesis [155, 233]. The processed pre-tasiRNA transcript is recruited to the RNA-dependent RNA polymerase6 (RDR6) amplification loop, which also requires the accessory factor SGS3, for dsRNA synthesis [84, 150, 152, 157], resulting in dsRNA that is processed by DCL4 [81, 85, 159]. The DCL4 reactions occur in an end-dependent, successive manner to yield 21 nt siRNA in phased arrays.

RNA silencing was first shown to have biological relevance when it was discovered to be an antiviral response in plants, and is now firmly established as an antiviral mechanism in some animals (flies and worms) [234]. RNA silencing functions to dampen accumulation of virus, and in many cases, to limit spread and pathogenicity. Antiviral silencing in plants results from modular processes involving multiple DCLs. This was revealed using *Arabidopsis* mutants with defects in silencing components, as well as from analysis of virus-encoded suppressors. Full antiviral responses against many (likely most) viruses require both siRNA-generating DCL2 and DCL4, although examples of responses involving DCL1 and DCL3 are also known [57, 73, 80, 235]. Many antiviral responses that require DCL2 and DCL4 also require RDR6/SGS3 to amplify silencing signals through de novo generation of dsRNA [87, 152, 236-238]. Thus, antiviral silencing shares features with the tasiRNA pathway. Additionally, some viruses, such

as DNA-containing geminiviruses, the pararetrovirus Cauliflower mosaic caulimovirus (CaMV) and some (+)-strand RNA viruses also involve DCL1 and DCL3 in viral siRNA biogenesis [57, 58, 80, 239]. The involvement or requirement of DCL1 is intriguing, as it suggests some viral RNAs resemble miRNA precursors with foldback structure. There is no direct evidence to indicate that dsRNA from replicative forms or replicative intermediates triggers antiviral silencing.

How is silencing of (+)-strand RNA viruses initiated? One possibility is that siRNA from self-complementary foldbacks interact with and guide cleavage of viral RNA at positions corresponding to the foldback-complementary sequence in the genome [240]. This would be similar to the tasiRNA initiation mechanism, in which miRNA-guided cleavage triggers RDR6-dependent dsRNA synthesis and sets the phase for DCL4 processing. For the few viruses that have been analyzed, siRNA from the (+)-strand is more abundant than (-)-strand siRNA, which is consistent with a major population arising directly from viral RNA in a pre-amplification step [58, 80, 241].

After the initial and amplification phases, DCL4 has a unique function in generation of a cell-nonautonomous signal that functions at least 10-15 cells away the cell of origin [58, 86, 151]. This may have particular relevance in conferring immunity in tissues adjacent to vascular cells, and interfering with virus establishment after vascular (phloem)-dependent long-distance movement. Therefore, systemic virus infection of plants requires effective mechanisms to suppress RNA silencing.

To counteract antiviral RNA silencing, most plant and many animal viruses have evolved silencing suppressor proteins which condition susceptibility by interfering with the host antiviral silencing response [234, 242, 243]. Many viral RNA silencing suppressors are also pathogenicity factors that cause developmental defects in host plants [242]. Collectively, viral suppressors are diverse in sequence, structure, and function. The 19 kDa protein of tombusviruses (p19) and the p21 kDa protein of *Beet yellows virus* directly bind and sequester small RNA duplexes, thus trapping an essential intermediate in the silencing pathway [92, 94, 196, 205, 244]. Other suppressors, such as the nodavirus B2 and influenza A NS1 proteins, bind longer dsRNA and inhibit DICER processing [98, 245]. The *Cucumber mosaic virus* (CMV) 2b suppressor inhibits AGO1, the only one of 10 AGO proteins in plants with known functions in antiviral silencing, by direct interaction [128]. The TCV capsid protein, CP (a.k.a. P38), functions by inhibition of DCL4, although by mechanisms that are unclear [57, 58]. Yet another suppressor, P0 from poleroviruses, functions as an F-box-like proteins and may direct

degradation of silencing components [246, 247]. The diversity of structure and function of silencing suppressors suggests that this counterdefense strategy evolved independently on many occasions in plants and animals.

HC-Pro was one of the first viral proteins identified as a suppressor of transgene- and virus-induced RNA silencing. Analyses of data from variant experimental systems led to the development of several different models for the mechanism of HC-Pro silencing suppression. In one model, HC-Pro was proposed to reverse established RNA silencing [174, 190, 192]. Another model involved the enlistment of a cellular negative regulator of RNA silencing, such as rgs-CaM, a calmodulin-related protein [248]. A third model proposed that HC-Pro acts downstream of an RNA-dependent RNA polymerase but inhibits accumulation of siRNAs, suggesting that DICER activity was impaired [91, 200]. A fourth model predicted that RISC activation was suppressed through interaction between HC-Pro and a protein or complex required for siRNA duplex unwinding [92]. Importantly, most comparative studies concluded that the possible mechanism by which HC-Pro suppresses RNA silencing differs from the mechanism of other suppressor proteins, including p19 of tombusviruses and p21 of BYV [91, 92, 242].

Previously, we showed that P1/HC-Pro, like p19 and p21, was an effective miRNA pathway suppressor that inhibited miRNA* turnover [92]. We proposed that miRNA/miRNA* unwinding and RISC assembly was suppressed in the presence of P1/HC-Pro. Recently it was shown that the expression of P1/HC-Pro weakly affects 5'end modification of miRNAs in tobacco [25], suggesting again that the small RNA population is sequestered from modification and silencing complexes in the presence of P1/HC-Pro. Here, we present evidence that P1/HC-Pro expressed from a transgene or a replicating virus interacts *in vivo* with miRNA/miRNA* duplexes. We show that P1/HC-Pro, p21 and p19 inhibit methylation of endogenous miRNA/miRNA* duplexes *in vivo* and that P1/HC-Pro interacts with small RNA duplexes *in vitro* as well as *in vivo*. We also demonstrate that small RNA duplex-binding suppressors inhibit siRNA-directed target RNA cleavage by preventing RNA silencing initiator complex programming. These results clearly define a class of viral RNA silencing suppressors that, while distinct evolutionarily and structurally, each act by sequestering small RNA duplexes, a crucial and common intermediate in endogenous and antiviral RNAi pathways.

Results

Suppressors inhibit siRNA-directed RNA cleavage in Drosophila *embryo extracts*

The *Drosophila* embryo extract-based *in vitro* RNA silencing system allows quantitative analysis of RISC programming and RISC-mediated target cleavage [29, 129] and was used successfully to analyze silencing suppression mediated by p19 [93]. We used the *Drosophila* system to analyze the silencing mechanisms of *Carnation Italian ringspot virus* p19, BYV p21 and TEV HC-Pro of TEV. P19 and a GST-p21 fusion protein were expressed and purified from bacteria, while HC-Pro was purified from virus-infected plants. We tested the effects of each suppressor on siRNA-guided RNA cleavage in target cleavage assays in which inducer siRNA, target RNA (containing sequences complementary to the inducer) and purified suppressor proteins were added simultaneously to *Drosophila* embryo extracts. RISC-mediated target cleavage was measured by quantification of the 5'-end product of siRNA-directed cleavage of target RNA generated over a dilution series of the suppressor protein. To analyze the effect of silencing suppressors on preassembled RISC activity, siRNA was preincubated with extracts to allow RISC formation, then target RNA and suppressor proteins were added and RNA cleavage was measured as described above. We tested suppressor effects on RISC programming in assays in which inducer siRNA and purified suppressor proteins were added to *Drosophila* embryo extracts, and formation of siRNA-programmed RISC was measured by electrophoresis mobility shift assays.

Addition of p19 resulted in complete inhibition of target cleavage with a half-maximal inhibition concentration (IC_{50}) of 15.24±2.3 nM (Figure 3.1A, upper and right panels). In reactions containing preassembled RISC, target RNA cleavage occurred with similar efficiency regardless of the concentration of p19 (Figure 3.1A, lower and right panels). Inhibition of target cleavage was inefficient with the p19 W39/42R mutant (IC_{50} = 75.9 +/- 3.4 nM; data not shown). p19 W39/42R does not effectively suppress transgene silencing, and introduction of this allele in a CIRV mutant results in reduced symptom severity during infection [196]. Due to participation of Trp39 and Trp42 in end-capping interactions with small RNA duplexes, the W39/42R mutant is likely compromised in siRNA binding [196]. Programming of RISC with inducer siRNA was compromised in the presence of p19 (data not shown; [94]). These results support the direct siRNA-binding model for p19 suppressor function [93, 196, 205] and suggest that the inhibition of target cleavage depends on the siRNA binding activity of p19.

p21 inhibited siRNA-mediated target RNA cleavage, although complete suppression of cleavage was not achieved with the GST-p21 preparation (Figure 3.1B, upper and right panels). Activity of preassembled RISC was refractory to this suppressor at all concentrations tested (Figure 1B, lower and right panels). RISC programming with inducer siRNA was reduced in a dose-dependent manner in the presence of p21 (data not shown; [94]).

HC-Pro also inhibited target RNA cleavage in the *Drosophila* system, and this effect was dose dependent (IC_{50}=118.22±5.36 nM) (Figure 3.1C, upper and right panels). Surprisingly, HC-Pro did not inhibit the activity of preassembled RISC at any concentration tested (Figure 1C, lower and right panels) but inhibited RISC programming (data not shown; [94]). These observations lead us to speculate that HC-Pro may possess siRNA duplex binding activity.

HC-Pro interacts with 21 nt siRNA duplexes in vitro

We analysed the RNA binding properties of purified HC-Pro with gel mobility shift assays. Synthetic 21 nt or 24 nt siRNA duplexes and 19 nt or 21 nt blunt-ended RNA duplexes were used to test the small RNA binding specificity of HC-Pro. HC-Pro bound 21 nt siRNA duplexes with the highest affinity among the RNAs tested (Figure 3.2A). Complexes of HC-Pro with 24 nt siRNA duplexes and blunt duplexes were detected, although in much lower quantities than with 21 nt siRNA duplexes (Figure 3.2B-D). HC-Pro failed to bind single-stranded RNAs (not shown). HC-Pro is proposed to interact with cellular proteins to mediate silencing suppression [92, 248]. To address the contribution of cellular proteins to the HC-Pro suppression mechanism, HC-Pro-small RNA binding assays were repeated in the presence of *Arabidopsis thaliana* plant extracts. Strikingly, an unidentified plant factor significantly enhanced the affinity of HC-Pro for all duplexes tested (Figure 3.2A-D). Plant extracts did not similarly affect the binding of p19 or p21 to siRNA duplexes (not shown).

Suppressors bind virus-derived siRNA duplexes and miRNA/miRNA intermediates* in vivo

To determine whether p21 and HC-Pro interact with virus-derived siRNA *in vivo,* we analysed siRNA accumulation and *in vivo* interactions between these two suppressors and virus-derived siRNA in BYV- and TEV-infected *Nicotiana benthamiana* plants. In systemically-infected leaves, viral genomic RNA and virus-derived 21 nt siRNA increased during an infection time-course, suggesting that siRNA production is not inhibited *in vivo* (Figure 3.3A,B).

BYV-derived siRNAs were analysed in α-HA immunoprecipitates (IP eluates) from BYV-infected plants expressing HA-tagged p21. BYV-derived siRNAs and p21 were detected in total extracts (inputs) and α-HA IP eluates from virus-infected plants (Figure 3.4B). Neither p21 nor BYV-derived siRNAs were detected in the eluate of the α -His control IP (Figure 3.4B).

In previous experiments, *in vivo* interactions between small RNA duplexes and HC-Pro were not detected. However, miRNA* intermediates and two of three cognate miRNAs accumulated to higher levels in *Arabidopsis* plants expressing an HC-Pro transgene than in wild-type plants [92]. To re-investigate potential interactions between HC-Pro and small RNA duplexes, HC-Pro containing an N-terminal 6xHis tag was expressed from TEV [249] in infected *N. benthamiana* plants and immunoprecipitated using conditions less stringent than those used by Chapman et al. [92]. miR171, miR171* and virus-derived siRNA were analysed in total extracts and α -His IP eluates from mock-inoculated plants and TEV-infected plants. miR171 and miR171* were detected in total extracts from infected plants (IP inputs) and in α-His IP eluates (Figure 3.4A). miR171 was detected in mock-inoculated plant extracts, but at lower levels than in extracts from TEV-infected plants (Figure 4A, bottom panel). miR171* was not detected in total extracts from mock- inoculated plants, consistent with previous findings [91, 92, 201]. Neither miR171 nor miR171* was detected in α -His IP eluates from mock-inoculated plants. TEV-derived siRNA also co-immunoprecipitated with tagged HC-Pro. The immunoprecipitated TEV-derived siRNA co-migrated during electrophoresis with a size standard corresponding to 21 nt duplex siRNA, but not a single-stranded 21 nt siRNA (Figure 3.4C).

Suppressors inhibit miRNA methylation

Recently it was shown that the expression of P1/HC-Pro affects 5'end modification of miRNAs in tobacco [25], which is likely a result of miRNA/miRNA* duplex sequestration by P1/HC-Pro. We hypothesized that HEN1-mediated methylation of miRNA/miRNA* duplexes would be inhibited in the presence of the RNA-binding suppressors P1/HC-Pro, p21 and p19. To determine if these suppressors affect methylation of the 3'-terminal nucleotide of miRNAs, total RNA from *Arabidopsis* transgenic lines that constitutively expressed these viral silencing suppressors [92] was treated with sodium periodate followed by β elimination. Specific miRNAs were analyzed by RNA gel blot and filter hybridization. The reactions eliminate a 3' nucleotide containing both 2' and 3' OH groups on the ribose, resulting in an RNA product that is one

nucleotide shorter than the substrate RNA and that contains a 3' phosphate group [250]. Therefore, miRNAs that contain a 3' terminal ribose with both 2' and 3' OH groups will migrate faster during electrophoresis by between one and two nucleotides after the treatment. Methylated miRNAs whose 3' terminal ribose contains a 2'-O-methyl group do not participate in the reactions and remain unchanged in mobility. This method was previously used to study miRNA 3' end ribose methylation [27, 251]. Although this method does not distinguish a 2'-O-methyl modification from other potential modifications on the 3' terminal ribose, previous mass spectrometry studies demonstrated that methylation is the only modification of an *Arabidopsis* miRNA [27].

The methylation status of miR159, miR167, miR171 and miR172 was analyzed in pools of primary transformants containing each of the three viral silencing suppressors or the vector DNA alone [92]. No mobility change was detected in the vector–transformed plants, indicating that the miRNA pools in control plants were fully or near fully methylated (gel blots not shown; [252] and Fig. 3.5). A portion of the miRNAs from the suppressor-expressing plants migrated more quickly after β-elimination reactions. The proportion of unmethylated miRNAs ranged from 17.5% (miR171) to 42.7% (miR159) in p19-expressing plants, 19.5% (miR159) to 68% (miR172) in P1/HC-Pro-expressing plants and from undetectable (miR171) to 88.5% (miR172) in p21-expressing plants (Figure 3.5). Therefore, the viral silencing suppressors partially affected miRNA methylation, but to different levels depending on the miRNA.

Co-immunoprecipitation of methylated and unmethylated miRNAs with suppressors

To determine if these viral silencing suppressors have a preference for unmethylated or methylated miRNAs, the methylation status of miRNAs associated with co-immunoprecipitated (co-IP) suppressor complexes was analyzed. Modified suppressors were immunoprecipitated using an anti-hemagglutinin (HA) monoclonal antibody, which recognizes an influenza HA epitope tag fused to the C terminus of each suppressor [92].

miR167, miR171 and miR172 co-immunoprecipitated differentially with the three suppressors. Each miRNA co-immunoprecipitated with p21, whereas only miR167 and miR171 were detected in co-IP fractions with p19 (Figure 3.6, lanes 15-16, 7-8). Only miR172 co-immunoprecipitated with HA-tagged HC-Pro (Figure 6, lanes 11-12). However, in all cases of miRNAs that co-immunoprecipitated with suppressor, the methylation status was similar to that

detected in the input samples (Figure 3.6), indicating that each suppressor binds both methylated and unmethylated miRNAs.

Discussion

A multi-pronged approach was used to compare the activities of three distinct RNA silencing suppressors. Results obtained in the *Drosophila* embryo extracts revealed that p19, p21 and HC-Pro each inhibited siRNA programming of RISC and target cleavage. This occurred when the suppressors were introduced into embryo extracts simultaneously with effector siRNA duplexes, but not when suppressors were added after siRNA duplexes. Time-course analysis of RISC assembly indicated that each suppressor acted primarily to block formation of initial siRNA-DICER2-R2D2 intermediate complexes (data not shown; [94, 129]). These data support the hypothesis that suppressor-mediated reduction of target cleavage was primarily due to inhibition of RISC programming rather than inhibition of activity of assembled RISC.

The basis for inhibition of RISC assembly and target cleavage by p19, p21 and HC-Pro was clearly sequestration of siRNA duplexes. p19 and p21 bound siRNA with high affinity in the presence or absence of *Drosophila* extract, as predicted from the known properties of both proteins. Effective binding of HC-Pro to siRNA duplexes HC-Pro required a cellular factor that was detected in *Arabidopsis* extracts. However, the question of whether HC-Pro requires a cellular factor for efficient silencing suppression *in planta* requires further study. The finding that p19, p21 and HC-Pro each bind siRNA duplexes in a dose-dependent manner extends the sequestration model to three evolutionarily distinct virus-encoded suppressors [93, 187].

Mechanisms for suppression of RNA silencing by HC-Pro, p19 and p21 in vivo

Both structural properties and biochemical activities reveal clearly that p19 binds siRNA duplexes in a size-selective (~19 bp) manner regardless of the presence of 3' 2 nt overhangs [196]. Here, p21 is reported to bind small RNAs *in vitro* and *in vivo*, possess a higher affinity for 21 nt siRNA duplexes relative to 24 nt siRNA duplexes, bind preferentially to duplexes with 3' 2 nt overhangs, and to lack ss siRNA-binding activity. These data are fully consistent with our previous *in vitro* and *in vivo* analyses of p21-small RNA interaction [92]. However, these data are not easy to reconcile with a recent study suggesting that p21 is a general nucleic acid binding protein [253]. One possible explanation for the disparate results is the concentration of protein, and protein/RNA ratios, used in different experiments. p21-siRNA duplex complexes form specifically at protein/siRNA ratios between 10:1 and 10,000:1 and protein concentrations up to

1 μM [92]; see Chapter 2). At high protein concentrations (1 – 10 μM) that exceed nucleic acid concentrations by 10,000-fold [253], we detected primarily nonspecific complexes with very slow mobility in gel shift assays (E.J.C. and J.C.C., unpublished data). It is possible that these slow mobility complexes correspond to the p21 octameric structures analyzed at atomic resolution by Ye and Patel [253]. We suggest that the octameric complexes may represent a structural form of p21 with functions that are distinct from RNA silencing suppressor functions.

Quantitative data showed that HC-Pro binds to 21 nt siRNA duplexes containing 3' 2 nt overhangs with higher affinity than to 19 nt duplexes lacking overhangs or to 24 nt siRNA duplexes. Thus, both p21 and HC-Pro differ from p19 by a requirement for 2 nt 3' end overhangs within the 21 nt siRNA duplex. TEV HC-Pro was also shown here to co-immunoprecipitate with virus-derived siRNAs and miR171/miR171* duplexes in extracts from infected plants. Infection by TEV promoted accumulation of miR171*, which normally accumulates to very low levels in noninfected plants. Previously, HC-Pro from *Turnip mosaic virus*, as well as p19 and p21, expressed in transgenic *Arabidopsis* were each shown promote accumulation of miR171*, miR167b* and miR160c*, leading to the suggestion that the three suppressors inhibit miRNA/miRNA* duplex unwinding [92]. However, unlike p19 and p21, which co-immunoprecipitated with siRNAs from long dsRNA and miRNA/miRNA* duplexes, no direct interaction between siRNAs or duplexes were detected with HC-Pro [92]. The failure to detect direct interactions between small RNAs and HC-Pro in the previous study may have been due to several reasons. First, the immunoprecipitation conditions used in the previous study were more stringent than that used in the current experiments. And second, the N-terminal 6xHis epitope tag used here differed from the C-terminal HA tag used previously. And third, the tagged HC-Pro used here was recovered from virus-infected plants rather than transgenic plants, the latter of which may have resulted in protein that was less accessible for immunoprecipitation.

None of the three suppressors affected siRNA- and miRNA-programmed RISC activity *in planta* as analyzed using sensor constructs in transient assays (not reproduced here; [94]). This is consistent with experiments using the *Drosophila* extract system. Previous reports suggested that HC-Pro could reverse established RNA silencing in a transgene-based system [174, 190, 192]. It was also reported that in HC-Pro-expressing transgenic plants, transgene-derived siRNAs are downregulated, suggesting that HC-Pro interferes with siRNA production [91, 200]. However, recent results [94] clearly showed that co-expression of HC-Pro, p19, or p21 with GFP-IR in *N. benthamiana* GFP16c/RDR6i did not compromise dsRNA processing to siRNAs, which agrees with previous data obtained with HC-Pro using a different GFP inverted

repeat construct [199]. In plants, initiator-dependent maintenance of RNA silencing requires the RNA-dependent RNA polymerase, RDR6 [150]. Sequestration of siRNAs by suppressors in RDR6-dependent systems would inhibit the maintenance/amplification step and lower the levels of siRNAs produced. This would have the same effect as losing DICER activity. Therefore, the proposed sequestration mechanism for HC-Pro provides a reasonable explanation for many observations derived from different systems. Of course, HC-Pro may also possess other activities or mechanisms related to RNA silencing. The association of HC-Pro with one or more cellular factors, as indicated by the *in vitro* assays (Fig. 2) and other analyses [248], means this remains an open possibility.

How do viral silencing suppressors affect miRNA methylation? First, they may compete with HEN1 for substrate miRNA/miRNA* duplexes. Sequestration by the suppressors could exclude HEN1 from interacting with duplexes or prevent HEN1 access to the 2' OH of the 3' terminal nucleotide. Alternatively, viral silencing suppressors may bind directly to HEN1 and inhibit its activity. However, we were unable to detect HEN1 as a co-IP component with viral silencing suppressor complexes using the anti-HA serum (data not shown), arguing against this hypothesis. It is still possible that these suppressors interact with other factors required for HEN1 function. Third, viral suppressors may affect the subcellular localization of HEN1. It was previously reported that p19 relocates the ALY protein from the nucleus to the cytoplasm [254], indicating the potential of suppressors to affect subcellular protein trafficking. Whether or not these suppressors affect HEN1 localization remains to be determined. It should be noted that our studies were done with transgenic *Arabidopsis* expressing the viral RNA silencing suppressors with the strong 35S promoter. It remains to be determined whether or not these proteins inhibit miRNA methylation during viral infections, although it is clear that miRNA functions are inhibited in TuMV-infected plants [185]. It was recently shown that infection of *Arabidopsis* with *Oilseed rape mosaic tobamovirus* (ORMV) leads to reduced methylation of endogenous miRNAs and siRNAs, suggesting that ORMV also inhibits HEN1 [239].

Given the fact that many viral silencing suppressors do not affect miRNA/miRNA* duplex formation [91, 92] and that they interfere with the methylation of miRNA in *Arabidopsis*, they could be excellent tools to study when and where the methylation process occurs. For instance, p19 and p21 probably function in the cytoplasm [70, 254, 255] after miRNA/miRNA* duplexes are generated by DCL1 in the nucleus and subsequently exported to the cytoplasm [92, 256, 257]. This would mean that at least a portion of miRNA/miRNA* duplexes is methylated in the cytoplasm.

Is siRNA duplex sequestration a widely used strategy to suppress RNA silencing?

Inhibition of antiviral RNA silencing is critical prerequisite for the successful systemic invasion by many or most plant viruses. Silencing inhibition through siRNA sequestration seems advantageous, as production of siRNAs is a conserved element of the antiviral silencing in any host. p19, p21, and HC-Pro are structurally and evolutionarily unrelated proteins, each representing a small protein family specific to a respective viral taxon [196, 253, 255, 258, 259]. Although only a limited number of silencing suppressors have been shown experimentally to bind small RNA duplexes [91-93], there are several additional suppressors for which a similar mechanism is predicted. For example, p14 of PoLV is an siRNA-interacting and long dsRNA binding protein [244], and there are several other viruses that trigger production of siRNA-binding proteins during infection (D. Silhavy, personal communication). However, there are also other known or suggested mechanisms of silencing suppression, including F-box-like activity that leads to destruction of silencing components [247], direct inhibition of one or more DICER activities [188] or sequestering of mature small RNAs after duplex unwinding [260]. The siRNA duplex-binding mechanism, however, represents a recurring mechanism that has evolved independently in several families (Tombusviridae, Potyviridae, and Closteroviridae) within the positive-strand RNA viruses.

Experimental Procedures

Transgenic plants

Arabidopsis thaliana Col-0 transgenic plants expressing influenza HA epitope-tagged P1/HC-Pro, p21, p19 or the empty expression vector were constructed as described [92]. Seeds were grown under selection for phosphoinothricin resistance and pools of primary transformants were analyzed.

Analysis of the methylation status of small RNAs

Sodium periodate treatments were done as described [27, 250, 251]. RNA was dissolved in borax/boric acid buffer (0.06 M, pH 8.6) and sodium periodate (200 mM in water) was added to a final concentration of 25 mM. The RNA was then incubated in darkness at room temperature. After 1 h of incubation, 1/10 volume of glycerol was added to the RNA and the incubation was continued for an additional 30 minutes. The RNA was then precipitated in the

presence of ethanol. For β elimination, the sodium periodate-treated RNA was dissolved in NaOH/borax/boric acid buffer (0.055 M, pH 9.5), incubated at 45°C for 90 min, and precipitated with ethanol.

Immunoprecipitation

The immunoprecipitation of HA-tagged viral RNA silencing suppressors and RNA isolation from precipitates were performed as described [92]. Briefly, tissues from transgenic *Arabidopsis* were ground in liquid nitrogen and homogenized in 5 ml/g lysis buffer (50 mM Tris-HCl at pH 7.4, 100 mM KCl, 2.5 mM $MgCl_2$, 0.1% NP-40, and 2x complete protease inhibitor cocktail; Roche), and centrifuged for 15 min at 9500 x g. Clarified lysates were pre-cleared with protein A-agarose, reacted with anti-HA (Roche) for 1 h at 4°C, and then incubated with protein A-agarose beads for 3 h at 4°C. Precipitates were washed three times in lysis buffer and divided for protein and RNA analysis. Nucleic acid was recovered by treatment with 3 volumes of proteinase K solution (100 mM Tris-HCl, pH 7.4, 10 mM EDTA, 150 mM NaCl, 2% SDS, and 0.2 µg/µl proteinase K) for 15 min at 65°C followed by extraction with saturated phenol and phenol:chloroform and ethanol precipitation. For immunoprecipitation of HA-tagged p21, leaves of BYV-infected *N. benthamiana* plants were infiltrated with *Agrobacterium* carrying a construct for expression of HA-tagged p21 (OD_{600} = 1.0). Tissue was collected 48 hours post-infiltration for preparation of extracts in IP buffer, and immunoprecipitations using α-HA or α -His were done as described in [92]. For immunoprecipitation of 6xHis-tagged HC-Pro, α-His antibody-conjugated beads were added to total extracts, incubated for 1 hour at 4°C, and washed with IP buffer. Immunoprecipitated complexes were eluted as described previously [93], and IP eluates were divided for protein extraction and RNA isolation.

Detection of small RNAs by RNA filter hybridization

RNA isolation, gel electrophoresis, blotting, hybridization and washes were performed as described [71, 93, 199]. Denaturing RNA gel blot hybridisation and analyses were done as described in Silhavy et al. [187]. 5' end-labeled (^{32}P) antisense DNA or LNA oligonucleotides were used as probes. Radioactive signals were detected and quantified with a phosphorimager. Total RNA from leaves of TEV-infected and mock-inoculated *N. benthamiana* plants was isolated in 2x PK buffer with Proteinase K as described previously [93]. Total RNA from leaves

of BYV-infected and mock-inoculated *N. benthamiana* plants was isolated using Trizol reagent [199].

Target cleavage assays

Drosophila embryo extract preparation, target RNA labelling, and siRNA annealing were described previously [261]. Reactions used 2 µl of embryo extract, 5 nM siRNA (final concentration) and 1x lysis buffer containing 10% v/v of glycerol in a total volume of 10 µl. Cap-labelled GFP target RNA was used at 0.5 nM final concentration. In direct competition assays, reactions were incubated for 1 hr. In preassembled RISC assays, siRNA and embryo extracts were pre-incubated for 30 min to allow RISC assembly prior to addition of target RNA and suppressor proteins. Samples were deproteinised and RNA was analyzed on an 8% denaturing gel.

Statistical analysis

All *in vitro* target experiments were done three times. The curves were best fitted to the indicated sets of data with the computer program Microcal Origin 5.00. The average of three trials ± standard deviation is shown.

Electrophoretic mobility shift assays

Labelling and annealing of RNA duplexes was done as described previously [93]. Purified proteins and labelled RNAs were incubated for 30 min at room temperature in lysis buffer [11] supplemented with 0.02% Tween 20. Complexes were resolved on 6% polyacrylamide 0.5x TBE gels. Gels were dried and exposed to a storage phosphor screen (Molecular Dynamics Typhoon Phosphorimager, Amersham Biosciences), and bands were quantified with a Genius Image Analyser (Syngene).

Fig. 3.1. Silencing suppressors inhibit siRNA-guided target RNA cleavage *in vitro*. Target cleavage assays used embryo extracts, target RNA (0.5 nM), 32P-labeled siRNA (5 nM), and a concentration series of purified p19 (A), GST-p21 (B), or HC-Pro (C). siRNA and target RNA was added to embryo extracts prior to purified suppressor protein (A-C: top panel) or siRNA was preassembled into RISC prior to addition of target RNA and purified suppressor protein (A-C: bottom panel). Effect of purified suppressor protein on target RNA cleavage by RISC (black squares) and preassembled RISC (red triangles) was plotted as a function of the concentration of suppressor used (A-C: right panel).

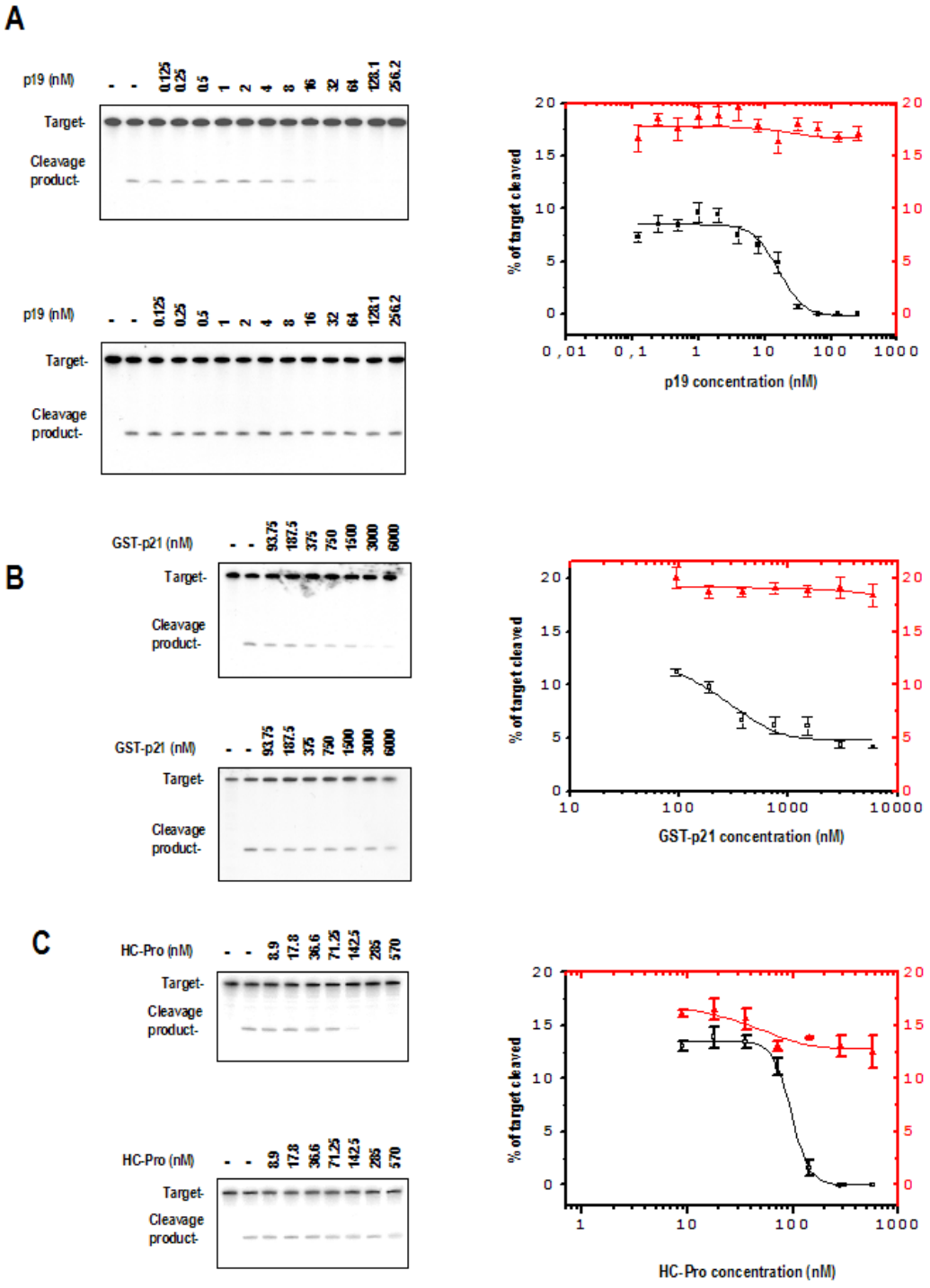

Fig. 3.1. Silencing suppressors inhibit siRNA-guided target RNA cleavage *in vitro*.

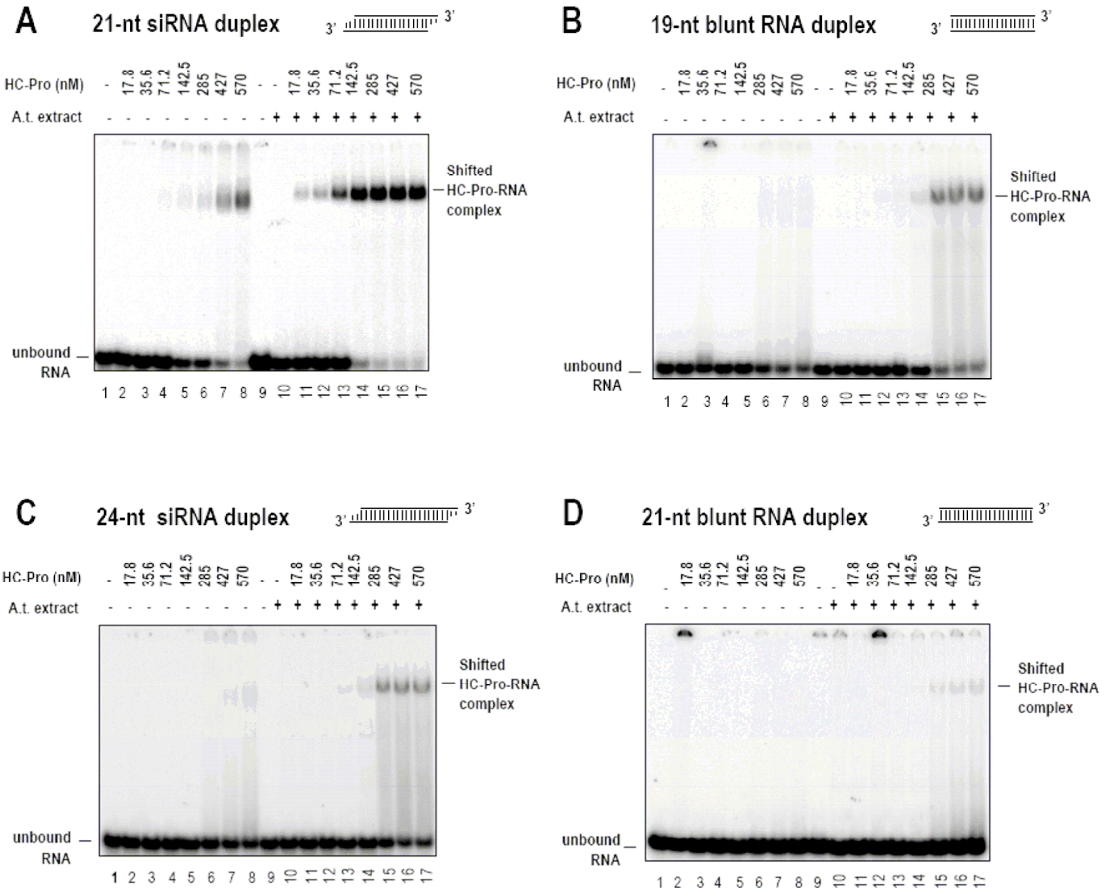

Fig. 3.2. Interactions between HC-Pro and small RNA duplexes. Electrophoretic mobility shift assays used purified HC-Pro and 21 nt siRNA duplex (A), blunt-ended 19 nt RNA duplexes (B), 24 nt siRNA duplex (C), and blunt-ended 21 nt RNA duplex (D) without (lanes 2-9) or with 4 μg of *Arabidopsis* thaliana cell culture extract (lanes 10-17).

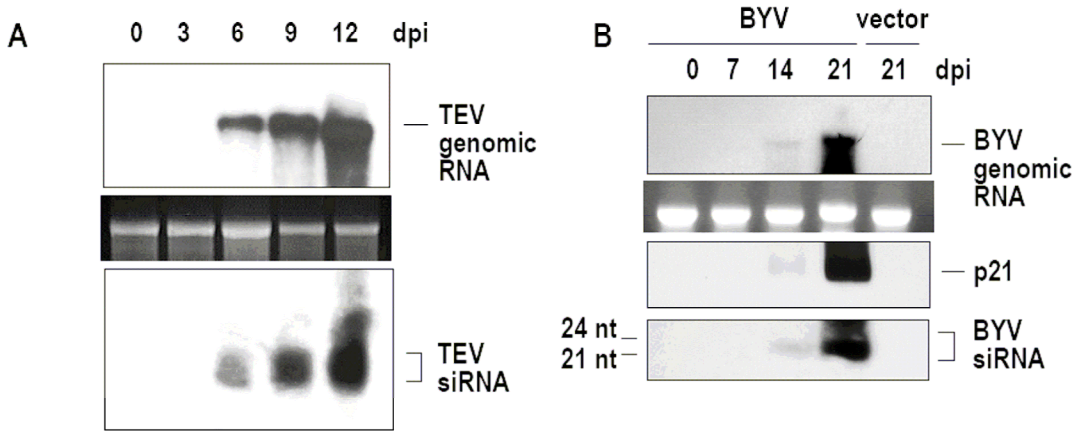

Fig. 3.3. Virus-derived siRNA accumulation *in vivo*. (A) Accumulation of viral genomic RNA and virus-derived siRNA during a TEV infection timecourse. (B) Accumulation of viral genomic RNA, p21 protein and virus-derived siRNA during a BYV infection timecourse.

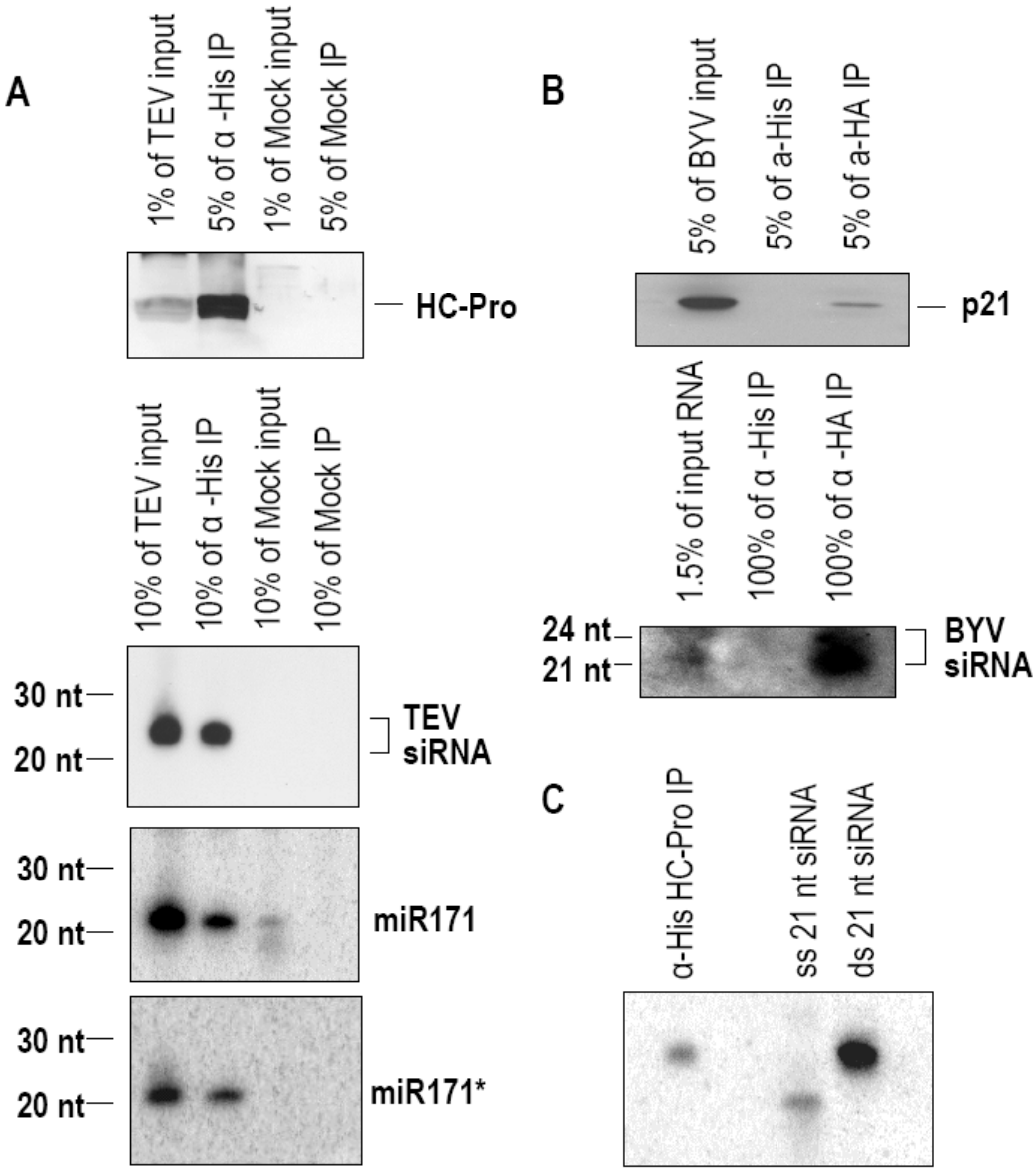

Fig. 3.4. Interactions between suppressor proteins and small RNAs in virus-infected plants. (A) Co-IP of HC-Pro, TEV-derived siRNAs, and miRNA171 and miRNA171* from TEV-infected N. benthamiana tissue. (B) Co-IP of p21 and BYV-derived siRNAs from BYV-infected N. benthamiana tissue. (C) TEV-derived siRNAs in HC-Pro IP eluates from TEV-infected plants were analyzed by non-denaturing gel electrophoresis; 21 nt siRNA duplexes and 21 nt ss RNAs were used as markers.

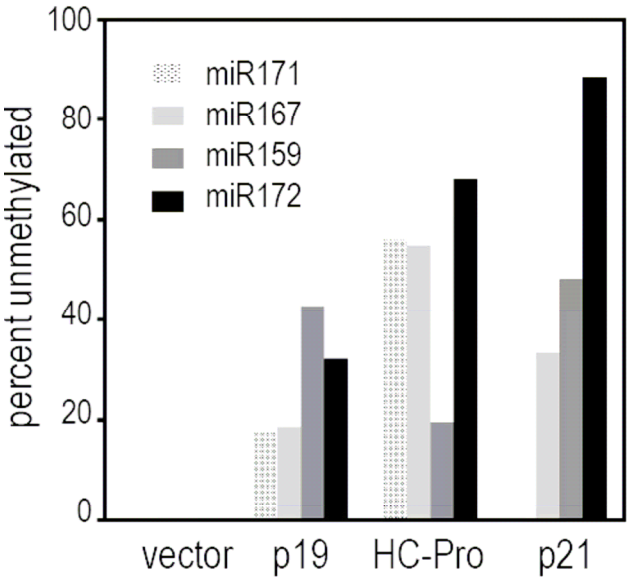

Fig. 3.5. Viral RNA silencing suppressor effects on miRNA methylation.
Methylation status of miRNAs in suppressor-expressing and control plants was calculated [unmethylated miRNA/(unmethylated miRNA+ methylated miRNA)*100%]..

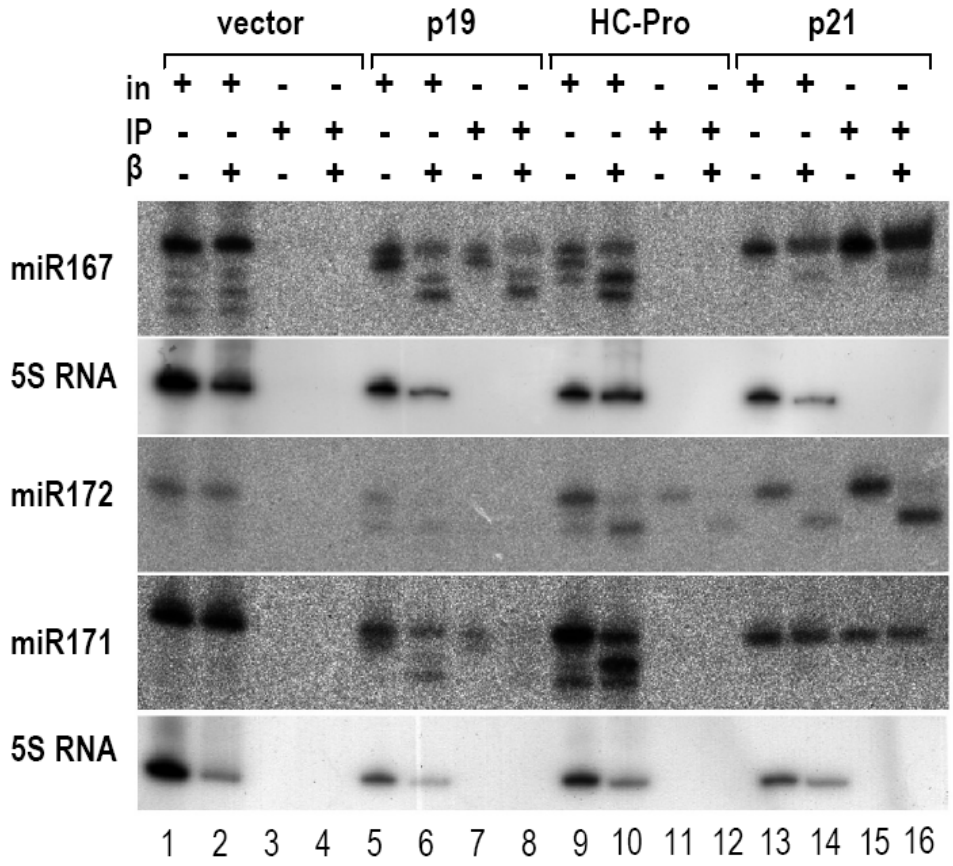

Fig. 3.6. Co-IP of methylated and unmethylated miRNAs with viral silencing suppressors. RNAs from the IP input (IN) sample or from IP eluates (IP) were either subjected (+) or not (–) to chemical modification reactions (β) and RNA species were detected by filter hybridization. The miR172 and miR171 panels were from the same blot and share the 5S rRNA control (bottom panel).

CHAPTER 4

DEVELOPMENT OF SMALL RNA PROFILING BY HIGH THROUGHPUT SEQUENCING IN *ARABIDOPSIS*

Elisabeth J. Chapman, Noah Fahlgren, Kristin D. Kasschau, Miya Howell, Christopher M. Sullivan, Jason Cumbie

PLoS ONE
185 Berry Street, Suite 3100, San Francisco, CA 94107 USA
Volume 2 Article e219 (Open Access)

PLoS Biology
185 Berry Street, Suite 3100, San Francisco, CA 94107 USA
Volume 5 Article e57 (Open Access)

Summary

Eukaryotes contain a diversified set of small RNA-guided pathways that control genes, repeated sequences and viruses at the transcriptional and posttranscriptional levels. Analysis of small RNA populations on a genome-wide level is facilitated by generation of large small RNA datasets using high-throughput sequencing (HTS) technologies. We developed procedures for small RNA profiling by HTS that maximize yield of high-quality small RNA sequences in multiplexed formats. We used this methodology to generate genome-wide profiles of small RNAs in wild-type (wt) *Arabidopsis thaliana* and silencing pathway mutants with defects in three *RNA-DEPENDENT RNA POLYMERASE* (*RDR*) and four *DICER-LIKE* (*DCL*) genes, as well as in a time course of *Arabidopsis* infected with *Pseudomonas syringae* pv. tomato DC3000*hrcC*. These small RNA profiles revealed dynamic changes in expression of some microRNA families as well as the genome-wide distribution of small RNA in each major size class. These results establish high-throughput sequencing as a small RNA profiling tool and provide a comprehensive description of the major small RNA pathways in *Arabidopsis*.

Introduction

Most eukaryotes contain RNA-based silencing pathways that function at the transcriptional or posttranscriptional level to negatively regulate or control genes, repetitive sequences, viruses and mobile elements [262, 263]. At the heart of these silencing pathways are small RNAs, which arise from double-stranded (ds) RNA or self-complementary foldback structures by the activity of Dicer or Dicer-like (DCL) ribonucleases [264]. MicroRNAs (miRNAs) derive from RNA polymerase II transcripts that adopt foldback structure, and generally function in *trans* as negative regulators through base-pair interactions with mRNAs [17]. Interaction of miRNAs with target transcripts results in either target degradation or non-degradative inhibition of translation, depending on the degree of sequence complementarity [17]. Trans-acting siRNAs (ta-siRNAs) occur in plants and function like miRNAs as posttranscriptional negative regulators of target transcripts, but they form through an RNA-dependent RNA polymerase (RDR)-based mechanism in which precursor transcripts are converted to dsRNA [238, 265]. SiRNAs originating from dsRNA by bidirectional transcription, extended foldbacks with perfect complementarity or RDR-based mechanisms can also guide transcriptional silencing in which a locus adopts heterochromatic features, including dense cytosine methylation and histone modifications associated with silent chromatin [262, 266]. In all forms of RNA silencing, small RNAs associate with effector proteins in the

Argonaute (AGO) family [220]. Most AGO proteins contain an RNaseH-like ribonucleolytic domain that is required for RNA silencing activity [224].

Plants, like other multicellular eukaryotes, have evolved diversified sets of DCL, RDR, AGO and other factors to provide specialized or preferential functions in RNA silencing pathways [238, 265]. At least four DCL enzymes, in combination with specific dsRNA-binding proteins (DRBs, including HYL1), catalyze formation of miRNAs or various classes of siRNAs in plants (see Chapter 1). MiRNAs generally require HYL1, which likely works with DCL1 during precursor processing, and function through AGO1 [31, 33, 71, 112, 126]. For the siRNA pathways, at least three RNA-dependent RNA polymerases (RDR1, RDR2 and RDR6) are functional in plants [73, 144, 150, 152, 267]. RDR2 works with DCL3 to form chromatin-associated siRNAs (24 nts) that function through AGO4 [48, 73, 125, 130, 268, 269]. Functional chromatin-associated siRNAs also require RNA PolIVa and PolIVb, as well as DNA methylation, in a pathway that may include a reinforcement loop [73, 143-145, 227]. Ta-siRNAs form through the activity of RDR6, which functions with the precursor-stabilizing factor SGS3, and DCL4 to yield RNAs in a phased 21 nt configuration [84, 157, 158]. However, there is clearly diversification of functions or requirements within the ta-siRNA class, as subsets of ta-siRNAs form by a DRB4-dependent mechanism and function through an AGO7 (ZIPPY)-dependent pathway [34, 270-272]. Additionally, unique combinations of DCL and RDR proteins function during formation of natural antisense siRNA (nat-siRNA) from a locus with bidirectional transcription [12]. Thus, plants use diverse combinations of factors in specialized pathways to silence different classes of endogenous genes and sequences. Further, many of these factors, including DCL2 and DCL4, are also allocated to antiviral silencing pathways involving both cell-autonomous and systemic defense [58, 73, 127, 152, 273].

New technology permits practical deep sequencing of small RNA populations [60, 274]. This allows both qualitative discovery of new small RNAs or small RNA classes, as well as quantitative profiling of small RNA populations through direct sequence analysis. Despite the increased size of small RNA datasets afforded by HTS, some small RNA species remain rare or absent from such datasets. This observation lead to speculation that biases exist in the small RNA preparative steps and that despite the quantum leap in sequencing depth, sample sets remain too small to be saturating [147].

Here, small RNA preparative steps preceeding HTS were developed for maximized yield of high-quality small RNA sequence reads, and a molecular barcoding method was developed for small RNA profiling by HTS in multiplexed formats. Using this methodology,

the small RNA profiles from wt and *P.syringae*-infected *Arabidopsis thaliana* and seven RNA silencing mutants were analyzed. This revealed genome-wide patterns of expression involving each of the major classes of small RNAs, as well as redundancy and non-redundancy among the DCL family. Finally, the utility of HTS for dynamic expression profiling of small RNA was examined.

Results

Development of Small RNA Preparation Procedures

Yield of large datasets from high-throughput sequencing of small RNA requires preparation of small RNA-derived DNA templates of sufficient quantity and quality to generate a maximum number of high-quality sequence reads. To generate DNA templates for HTS and analysis of *Arabidopsis* small RNAs, we developed a series of size fractionation, RNA ligation, and PCR amplification steps based on methods described elsewhere [10, 275] and designed to maximize sequence read yield (Figure 4.1). This series consists of nine steps: 1) selection of 18-25 nt RNA; 2) ligation of adaptors onto 3' ends of small RNA using T4 RNA ligase; 3) selection of 3'ligated RNA; 4) ligation of adaptors onto 5' ends of small RNA using T4 RNA ligase; 5) cDNA synthesis; 6) cDNA amplification; 7) selection of full-length library amplicons; 8) asymmetric PCR amplification of full-length library amplicons; 9) selection of longer single-strands of asymmetric duplexes. The end product of this procedure is a pool of single-stranded DNA molecules containing sequence tags for HTS. Preliminary experiments indicated that size selection of small RNA or DNA amplicons, ligation of adaptors onto small RNA 3' ends, and cDNA amplification steps were critical for end product quality. Results of experiments designed to improve these steps are described below.

Size Selection. Separation of short nucleic acid fragments from higher-molecular species by gel electrophoresis effectively removes the majority of slower-mobility contaminant nucleic acids with various efficiencies depending on the gel and sample composition and other factors [276]. Subsequent recovery of RNA and DNA from polyacrylamide gel fragments using passive diffusion methods is inefficient (this paper). We tested the recovery of a radiolabeled 21nt RNA from a 17% PAGE-urea gel using passive diffusion into NaCl solution or electrophoretic transfer onto DE81 ion-exchange paper and subsequent elution by ion-exchange. Recovery using the DE81 method was 10-fold higher than with the NaCl method (Figure 4.2). Subsequent experiments testing elution of DNA fragments had similar results (data not shown). We applied

the DE81 method to all size selection steps throughout the preparative series: step 1, selection of 18-25 nt RNA; step 3, selection of 3'ligated RNA; step 7, selection of full-length library amplicons; and step 9, selection of the longer single-strands of asymmetric duplexes.

Ligation of adaptors onto 3' ends of small RNA using T4 RNA ligase. In sealing reactions catalyzed by T4 RNA ligase, a 5'phosphate moiety participates in a ligation reaction to an available 3'hydroxyl. RNA molecules containing both a 5'phosphate and 3'hydroxyl, such as DICER-generated small RNA, undergo circularization in reactions catalyzed by T4 RNA ligase [25, 277]. Circularized small RNA is unable to form an intermolecular ligation product with adaptor molecules.

We hypothesized that circularization would be minimized in the presence of a molar excess of adaptor and the accumulation of intermolecular ligation products would be favored. We tested the efficiency of intermolecular ligation in reactions containing varied molar ratios of 3'adaptor to input RNA. The 21nt input RNA containined a radiolabeled 5'monophosphate. To facilitate identification of various ligation products, alkaline phosphatase digests were done to remove 5' radiolabel from input RNA and intermolecular ligation products. Circularized forms were not detected in reactions with a 1000:1 ratio of adaptor to input RNA (Figure 4.3, lanes 7-8). However, circularized RNA accumulated in reactions containing these components at a 100:1 ratio or less (Figure 4.3, lanes 9-14). We selected a 1000:1 ratio of adaptor to small RNA for all subsequent preparative RNA ligation steps.

In RNA ligase reactions lacking exogenous ATP, RNA circularization is minimized and sealing reactions involving intermolecular interactions between a 5'adenylated RNA and an available 3'hydroxyl are favored [277]. To determine the efficiency of small RNA ligation to adenylated adaptors, we measured the accumulation of intermolecular ligation products in reactions containing 5'adenylated adaptors and in ATP-supplemented reactions containing 5'phosphorylated adaptors. In eight independent pairs of ligations, intermolecular ligation products accumulated to higher levels in reactions containing 5'adenylated adaptors than in ATP-supplemented reactions containing 5'phosphorylated adaptors (Figure 4). We therefore selected 5'adenylated adaptors for ligations to small RNA.

cDNA amplification. PCR amplification of small RNA-derived cDNA templates using ExTaq DNA polymerase resulted in accumulation of primer-dimer artifacts that contaminated DNA preparations and resulted in loss of yield of high-quality small RNA sequences (Figures 4.5 and 4.6, and data not shown). ExTaq also generated products in the absence of template that comigrated with full-length library amplicons (Figure 4.6, right side of graph). To identify

cDNA amplification parameters that reduce the accumulation of spurious products, we compared JumpStart (an antibody-conjugated DNA polymerase) across a range of annealing temperatures and concentrations of magnesium chloride to ExTaq used with the standard conditions (50ºC annealing temperature, 2.0 mM $MgCl_2$). We used a synthetic cDNA template at 20 pg/µl, which allows significant accumulation of primer-dimer forms under the standard ExTaq conditions (Figure 4.6). Ratio of full-length library amplicons to primer-dimer forms was highest in reactions containing 2.0-2.5 mM $MgCl_2$ and annealing temperatures ranging from 50 – 60ºC (Figure 7). Full-length library amplicons were not detected in reactions containing JumpStart polymerase and 1.5 mM $MgCl_2$ (Figure 4.7). We selected JumpStart polymerase, a 55ºC annealing temperature, and 2.5 mM $MgCl_2$ for PCR amplification of small RNA-derived cDNA.

Molecular barcoding of small RNA populations.

Use of HTS for small RNA expression profiling necessitates analyses of multiple small RNA populations in parallel. We developed a molecular barcoding approach to facilitate computational distinction of replicate samples and samples from different experimental treatments. The barcoding system consists of a set of eight unique 5' adaptors, each containing a variation of a four-nucleotide identifier sequence (Figure 4.8). Overall GC content and terminal sequences were preserved. Following adaptor-small RNA ligation steps and cDNA synthesis, cDNA amplification and asymmetric PCR was done using a forward PCR primer homologous to the adaptor used for each sample. cDNA from each independent sample was amplified independently, such that the barcode was preserved throughout the DNA phase of library construction. Forward and reverse primers also contained HTS recognition sequences flanking the adaptor-homologous sequences. This molecular barcoding approach allowed independent preparation and quality assessment of samples, which were then pooled for HTS in multiplexed formats (Figure 4.8).

Sequence analysis of multiplexed small RNA populations

To determine the extent to which the major small RNA pathways function across the *Arabidopsis* genome, small RNA populations from several tissues of wt, infected, and mutant plants were amplified by RT-PCR and sequenced using high-throughput methods (Table 4.1). Amplicons were prepared by 5' and 3' adaptor ligation and RT-PCR using small RNA fractions from leaf tissue of wt Col-0 plants or plants infected for 0, 1 or 3 hours with *P.syringae* pv.

tomato DC3000*hrcC*. Amplicons from inflorescence tissue (containing stage 1-12 flowers) were prepared from wt Col-0 plants, mutants with defects in each *DCL* gene (*dcl1-7, dcl2-1, dcl3-1, dcl4-2*), and mutants with defects in each *RDR* gene for which a function has been established (*rdr1-1, rdr2-1, rdr6-15*). Amplicons from whole seedlings (3 day post-germination) were prepared from Col-0 and *rdr6-15* plants, and from leaves of Col-0 plants.

The Col-0 and mutant series samples were sequenced in a multiplexed format. Amplicons from three to six uniquely-barcoded samples were pooled and subjected to picoliter-scale pyrosequencing [274]. Small RNA sequences were parsed from data files and assigned to specific samples through identification of adaptor boundaries and barcode analysis, respectively. Sequences were mapped to the *Arabidopsis* genome by BLAST analysis and deposited into the *Arabidopsis* Small RNA Project (ASRP) database (http://asrp.cgrb.oregonstate.edu/db/) [278]. Given the base-call error rate using this method (~2%), a relatively high proportion of reads contained single-position mismatches relative to an *Arabidopsis* sequence. Positions containing such mismatches were sampled with all possible base substitutions, and the resulting sequences were queried by BLAST analyses. Repairs that led to unambiguous *Arabidopsis* sequences were incorporated in the database, resulting in recovery of 56,654 reads. For the analyses presented here, only sequences that were 20-25 nts in length were used. In total, each wt or mutant population was represented by between 13,688 and 78,583 reads (Table 4.1). For comparisons involving the mutants, population sizes were normalized using the smallest sequence set (from *dcl2* mutant inflorescence tissue).

In most samples, regardless of tissue type, small RNAs of 24 nts were the predominant size class (Figure 4.9B). The two exceptions were inflorescence samples from *rdr2* and *dcl3* mutants, in which the 24 nt size class was underrepresented and the 20-22 nt size classes were overrepresented. This is consistent with the known role of RDR2 and DCL3 in the biogenesis of 24 nt small RNAs [73, 82]. Given that the 21 nt miRNAs and ta-siRNAs do not increase in absolute abundance in *rdr2* and *dcl3* mutants [60, 73, 81, 279], the apparent increase in these mutants is likely due to overrepresentation in the sequenced population when the abundant 24 nt RNAs are lost.

Genome-wide distribution of small RNA-generating loci

The distribution of small RNA-generating loci from all samples was analyzed across each chromosome. A scrolling window analysis of all small RNA-homologous sequences (Figure 4.10A, Total loci) was done. These sequences include all loci, regardless of how

repetitive, that corresponded to a small RNA. For each chromosome, the density of small RNA-homologous loci was highest in the centromeric and pericentromeric regions (Figure 4.10). Pericentromeric and centromeric regions contain a high density of repeat sequence classes, such as transposons and retroelements. Indeed, the pericentromeric small RNA locus density roughly corresponded with the density of transposons and retroelements in each chromosome (Figure 4.10A). This could indicate that the repeat-rich centromeric/pericentromeric regions spawn relatively high numbers of small RNAs, or it could reflect the generation of relatively few small RNAs but from many, highly repetitive loci. Three additional scrolling-window analyses were done to distinguish between these possibilities. In the first, only small RNAs from unique loci (Figure 4.10A, Unique loci), to which a small RNA could be unambiguously assigned, were counted. In the second, the value of a prospective small RNA locus was divided by the number of loci corresponding to the small RNA sequence (Figure 4.10A, Repeat-normalized loci). In both of these analyses, the chromosome-wide counts of loci were decreased, although the density was still relatively high in pericentromeric regions. In the third, the abundance of small RNAs in the sequenced populations was plotted after repeat-normalization (Figure 4.10A, Repeat-normalized counts). This resulted in a small RNA abundance distribution that reflected the repeat-normalized locus density in the pericentromeric regions. This indicates that small RNA-generating loci from pericentromeric regions are abundant, highly active and diverse. Interestingly, repeat-normalized counts were relatively low at the centromeres, indicating a low small RNA sequence diversity in these regions (Figure 4.10A). Additionally, sharp isolated peaks of small RNA abundance were detected around the *Arabidopsis* genome. In nearly all cases, these corresponded to 21 nt miRNAs or ta-siRNAs, although exceptions included the small RNA population from inverted duplications at two positions in chromosome III, and the 24 nt miR163 (Figure 4.10B).

Genome-wide changes in small RNA abundance in each *dcl* and *rdr* mutant (inflorescence tissue) were analyzed by the scrolling window method and displayed in relation to chromosome position. Data were visualized using heat maps to show over- or under-representation of repeat-normalized small RNA levels. For most mutants, small RNA distribution changes were sporadic and localized. In the case of *dcl1*, these localized changes reflected loss of miRNAs and ta-siRNAs. In contrast, total small RNAs (20-25 nts) were systematically underrepresented around centromeric and pericentromeric regions of each chromosome in the *rdr2* mutant (Figure 4.10C). A comparable underrepresentation pattern was not detected in the *dcl3* mutant (Figure 4.10C), even though these two factors function

coordinately in the same pathway. To investigate this discrepancy between *rdr2* and *dcl3* mutants, each small RNA size class was analyzed independently. The 24 nt small RNAs were lost in the same chromosome-wide pattern in both *rdr2* and *dcl3* (Figure 4.10C), and in a pattern that was coincident with the loss of total small RNAs in *rdr2*. Interestingly, 21 and 22 nt small RNAs were increased in the *dcl3* mutant in a reciprocal chromosome-wide pattern relative to the 24 nt RNA loss pattern (Figure 4.10B, C). A corresponding increase in 21 and 22 nt RNAs was not detected in *rdr2* (Figure 4.10B, C).

The chromosome-wide changes in small RNA profiles were further analyzed by measuring and plotting the abundance of repeat-normalized 21+22 nt and 24 nt size class RNAs in Col-0, *rdr2* and *dcl3*. In Col-0, 24 nt small RNA accumulation correlated well with the distribution of total repeat-normalized small RNA loci, with broad peaks detected in pericentromeric regions (Figure 4.10B). The abundance of 21+22 nt RNAs was characterized by isolated peaks corresponding to miRNAs and ta-siRNAs, but also by broad, relatively low peaks that tracked with the 24 nt RNA abundance in the pericentromeric regions (Figure 4.10B). In the *rdr2* mutant, nearly all 24 nt small RNAs were absent (Figure 4.10B), although isolated peaks of 24 nt RNAs corresponding to miR163 and a few large inverted duplications were detected. Although the major miRNA and ta-siRNA peaks were unaffected in *rdr2*, other 21+22 nt small RNAs were generally less abundant from pericentromeric regions (Figure 4.10B). In *dcl3*, the major redistribution of size classes suggested by the heat maps was clear along each chromosome, particularly in the pericentromeric regions. Except for DCL1-dependent miR163, most 24 nt small RNA peaks and zones were lost and replaced by 21+22 nt small RNAs in *dcl3* (Figure 4.10B). Collectively, these data strongly implicate RDR2 and DCL3 as factors required for widespread biogenesis of 24 nt siRNAs. However, they also reveal genome-wide surrogate DCL functions that act on RDR2-dependent precursors to form 21 and 22 nt small RNAs in the absence of DCL3.

Effects of dcl3 and rdr2 mutations on small RNA populations from transposons and retroelements

The small RNA accumulation patterns in and around transposons, retroelements and pseudogenes were examined in Col-0 and each mutant plant. Small RNAs of all size classes from transposons/retroelements and pseudogenes were decreased specifically in the *rdr2* mutant (Figure 4.9C, E; Figure 3). In *dcl3*, however, the 24 nt size class was specifically lost and the 21+22 nt size classes were elevated from transposons/retroelements and pseudogenes (Figure

4.9C, E; Figure 4.11). These siRNAs generally accumulated in bidirectional clusters (Figure 4.11). This suggests that transposon/retroelements and pseudogenes are targeted by similar 24 nt siRNA generating systems.

The association of small RNAs of different size classes in and around transposons and retroelements in wt Col-0 plants was analyzed further. Specifically, the hypothesis that 24 nt, RDR2-dependent small RNAs are overrepresented from loci containing transposons/retroelements was tested. First, all annotated transposons and retroelements [280] were mapped to bins (250 nts) spanning each *Arabidopsis* chromosome (Figure 4.11A). Each bin overlapped with adjacent bins by 125 nts. A total of 81,328 bins contained transposons or retroelements. For statistical analysis, 10,000 random sets of 81,328 bins were assembled from the genome-wide pool. The randomized bin sets contained 31,860±701 total small RNA loci, with a size distribution similar to that of total small RNAs from Col-0 plants (Figure 4.9B). Bins containing transposons and retroelements contained 126,311 small RNA loci, indicating a highly significant enrichment (Z-score=135) over the randomized sets (Figure 4.12B). To normalize for potential bias from highly repeated loci, a similar analysis was done using only single-locus small RNAs that could be assigned unambiguously to one genome position (Figure 4.12C). Total small RNAs were highly overrepresented (Z-score=20.9) in transposon and retroelement bins using the unique-locus sequences. These analyses were then done for individual 21, 22, 23, and 24 nt size classes. Small RNAs of 23 and 24 nts were most highly overrepresented in the transposon and retroelement-containing bins using both total (Z-scores=138 and 156, respectively) and unique (Z-scores=23.7 and 22.6, respectively) sequences from Col-0 plants (Figure 4.12B, C). Although 21 and 22 nt small RNAs from transposon and retroelement bins were statistically overrepresented in each analysis using Col-0 data, they were overrepresented to a far lesser extent than were 23 and 24 nt RNAs (Figure 4.12B, C). Further, overrepresentation of all size classes in transposon and retroelement bins was lost or nearly lost in *rdr2* mutant plants (Figure 4.12B, C). In fact, unique 24 nt small RNAs in transposon and retroelement bins were represented at levels similar to those in the random bin sets, and 21 and 22 nt size classes were underrepresented in transposon and retroelement bins in *rdr2* plants (Figure 4.12C). These data indicate that transposon and retroelement loci are preferentially associated with the 24 nt-generating, RDR2-dependent pathway, but also that transposon and retroelement-derived small RNAs of all size classes depend on a functional RDR2 protein.

Direct sequencing as a miRNA expression profiling tool

Data from multiple mutants and treatments allowed us to explore the suitability of direct sequencing as a miRNA profiling tool. This was done in several ways. First, the reproducibility of the pyrosequencing method was analyzed using two normalized biological replicates [Col-0 inf. 1 (35,666 reads) and Col-0 inf. 2 (42,917 reads)]. miRNA families represented in both samples by at least three reads were compared and yielded an 83% correlation (Figure 4.13). This experimental approach is essentially a random sampling from a very large population of small RNAs and hence we anticipate that this correlation would increase with greater sequencing depth.

We next compared normalized expression profiles of miRNA families in Col-0 (inflorescence tissues) and *dcl1-7* (inflorescence tissues) mutant plants. As expected, based on the requirement for DCL1 in miRNA precursor processing, levels of nearly all miRNA families were decreased in *dcl1-7* mutant plants (Figure 4.13B). A few miRNAs were largely unaffected in *dcl1-7* plants. This likely reflects the known residual activity in, and the insensitivity of some miRNAs to, the partial-loss-of-function *dcl1-7* mutant allele [168].

To extend our small RNA profiling beyond different tissues and developmental stages, we generated expression profiles of miRNA families in leaves following infection by *P. syringae* pv. tomato (DC3000*hrcC*) at 0, 1 and 3 hr post-inoculation (p.i.) timepoints. This bacterium is unable to cause disease, because it is mutated in the key virulence factor delivery machine, the type III secretion system. This strain does trigger a robust basal defense response in *Arabidopsis* [281]. As predicted [282], miR393 was strongly up-regulated (10-fold at 3 hr p.i.) (Figure 4.13C). miR393 targets mRNAs encoding the auxin receptor, TRANSPORT INHIBITOR RESPONSE1 (TIR1), and related proteins [211, 282]. Two additional miRNA families, miR160 and miR167, were significantly elevated at 3 hr p.i. by 5-fold and 6-fold, respectively (Figure 4.13C). miR160 and miR167 each target mRNAs encoding members of the ARF family of transcription factors [182]. These data suggest that basal defense responses triggered by *P. syringae* pv. tomato (DC3000*hrcC*) include miRNA-mediated suppression of multiple components of auxin signaling pathways. Hence, our data extend the conclusions of Navarro et al. [282]. miR825, for which several targets were predicted but not validated (Table 2), was significantly down-regulated 3 hr p.i. (Figure 4.13C).

Thus, direct sequence-based profiling shows considerable promise in revealing dynamic changes in miRNA populations, although shallow sequencing depth will limit applicability of the method. While both open-ended platforms (such as direct sequencing) and closed-ended

platforms (such as solid-state microarrays) can be used to profile known miRNAs, the sequence-based profiling approach affords discovery of previously unknown miRNAs. It also allows discrete measurements of complex mixtures of small RNAs that arise in heterogeneous populations from loci with diffuse boundaries [59, 147, 148].

Discussion

454 Sequencing for small RNA profiling

High-throughput, parallel sequencing technology was used to profile small RNA populations in wt plants and mutants with defects in RNA silencing components. Application of 454 sequencing technology to small RNA profiling is appealing due to the read lengths, throughput, and capacity to multiplex [274]. The latter attribute was particularly useful given the number of samples analyzed in this study. Using a 4 nt molecular barcode on the 5' adapter, up to 6 samples were sequenced simultaneously and then resolved computationally. Of course, although multiplexing permits throughput of larger numbers of samples, it is done at the cost of sequencing depth. For analyses that require saturation sequencing or reads for very low-abundance small RNAs, multiplexing would introduce more serious limitations. Further limiting the number of reads in a multiplexed format is the relatively high error rate for base calls [274]. Unambiguous sequence identity was required to parse small RNAs from adapters, map sequences to the *Arabidopsis* genome, and assign sample IDs. The need for three steps to generate a successful read in a multiplexed format, therefore, lowers the yield of usable sequences from a single 454 run.

The value of 454 sequencing as a profiling method arises from the ability to count reads as a quantitative indicator of abundance. Thus, the method works across an open-ended range of values greater than zero, and without the need for prior knowledge of any specific sequence. Normalization to adjust for variable sequencing depth between samples was achieved using a linear proportion method. However, interpretation of representational differences in abundance of specific small RNAs or small RNA classes in different samples required careful interpretation. For example, increased representation of miRNAs in sequenced populations from the *rdr2* mutant, as shown here and documented previously [147, 279], was due to miRNA overrepresentation in a low-diversity pool of sequences that lack most 24 nt siRNAs. The absolute abundance of miRNAs in the *rdr2* mutants is relatively unaffected [60, 73, 81, 279]. In contrast, increased representation of 21 nt small RNAs from transposons and retroelements in the *dcl3* mutant was due to absolute increases in abundance of this size class.

Utilization and molecular redundancy of the RDR2-DCL3-dependent small RNA biogenesis pathway

By far, the most abundant and broadly utilized small RNA pathway in *Arabidopsis* depends on RDR2 and DCL3, resulting in accumulation of a highly diverse population of primarily 24 nt RNAs. Using a variety of measures, this pathway has a major affiliation with repeated sequences, including transposons and retroelements. The 24 nt RNAs are highly overrepresented from these feature classes, even when quantitative adjustments for repeat density are used (Figure 4.12). In other words, the repeated sequences analyzed here have disproportionately high activity as loci that spawn 24 nt small RNAs. Analysis of the unique sequence component of transposon/retroelement small RNA populations, as well as the repeat-normalized small RNA data from transposon/retroelement sequences, allows us to rule out the possibility that the high number of small RNAs from these feature classes is due simply to locus abundance.

In the *rdr2* mutant, all size classes of small RNAs from transposons, retroelements and pseudogenes were lost, suggesting that other RDR proteins (RDR1 and RDR6) are largely unable to compensate for loss of RDR2. This was in striking contrast to the *dcl3* mutant, in which only the 24 nt size class was lost. In fact, there was a major redistribution of sizes from 24 nt to 21/22 nts at repeat loci in the *dcl3* mutant. This almost certainly indicates that alternate DCLs that yield 21 and 22 nt small RNAs gain considerably more access to RDR2-dependent dsRNA in the absence of DCL3. DCL1 and DCL4 are known to catalyze formation of predominantly 21 nt RNAs [71, 72, 81, 85, 159], while DCL2 catalyzes formation of 22 nt RNAs [58, 73, 81]. We suggest that RDR2-dependent dsRNA has the potential to be processed by multiple DCLs, although preferentially by DCL3, in wt plants. Are these alternate-size small RNAs formed in *dcl3* mutant plants functional to guide events (DNA methylation and chromatin modifications) associated with transcriptional silencing? Functional analyses of transcriptional silencing at several loci, including *AtSN1* retroelements and the repeat sequences in the *FWA* gene promoter, reveal that *dcl3* mutants are generally weaker than *rdr2* mutants [73, 83, 283]. Partial redundancy or compensation in the absence of DCL3 may explain the limited transcriptional silencing in *dcl3* mutants.

Why might DCL3, but not RDR2, be subject to molecular redundancy or surrogacy? The events that route transcripts and small RNAs through the RDR2-DCL3-dependent pathway were found to be spatially localized in nuclear compartments [82, 146]. PolIVa, which is

necessary for most small RNAs generated by RDR2-DCL3 [47, 143-145, 227, 228], is spatially dispersed in the nucleus [82]. RDR2, DCL3, AGO4, Pol IVb, and siRNAs, on the other hand, accumulate in RNA processing centers (Cajal bodies) associated with the nucleolus [82, 146]. It is likely, therefore, that RNA trafficking mechanisms deliver RDR2 substrates to processing centers. Access to alternative RDR proteins (RDR1 and RDR6) may be prevented by distinct accumulation patterns outside of these centers, or by unique association of substrate transcripts with factors that interact specifically with RDR2. Although each of the DCL proteins have nuclear transport signals, only DCL3 is known to accumulate in processing centers [42, 70, 73, 82]. We suggest that there is relatively free access to RDR2-dependent dsRNA by alternate DCLs when DCL3 is absent. DsRNA synthesis, siRNA duplex formation and AGO4 loading with siRNAs likely occur in a spatially restricted complex [82, 146] in which functional DCL3 may occupy a preferential position (and thus normally exclude other DCLs) through unique associations with other factors.

Role of RDR2-DCL3-dependent small RNA pathway in regulation of genes and genetic buffering

The RDR2-DCL3-dependent pathway is necessary to maintain transcriptional silencing of some (though not all) transposons and repeat sequences over the course of single or multiple generations [60, 73, 83, 283, 284]. Loss of this pathway in *Arabidopsis* is associated with accumulation of defects over multiple generations [81]. Specific transposons in maize are progressively reactivated over multiple generations in the absence of MOP1, the ortholog of *Arabidopsis* RDR2 [154, 285-288]. The patterns of siRNA accumulation across the *Arabidopsis* genome strongly support the concept that RDR2-DCL3 pathway provides either maintenance or generational reinforcement of transposon and repeat sequence silencing.

While some of the developmental defects that accumulate in the absence of the RDR2-DCL3 pathway [81] may be due to mutagenesis associated with transposon activity, some may involve stochastic, epigenetic events that release silent chromatin to an active state through histone modifications and loss of DNA methylation. In these cases, the RDR2-DCL3-dependent siRNA pathway may have a role in silencing initiation, but not necessarily in maintenance of silencing in subsequent generations. Loss of RDR2 and DCL3, therefore, may affect some transcriptionally silent loci in unpredictable ways, or may have no effect at all, even if siRNAs from this pathway triggered initial silencing.

Experimental Procedures

Plant materials

A. thaliana (Col-0 background) mutants containing *dcl1-7, dcl2-1, dcl3-1, dcl4-2, rdr1-1, rdr2-1* and *rdr6-15* alleles were described previously [72, 73, 84, 157, 158]. Plants used for inflorescence tissue (containing stage 0-12 flowers) were grown under standard greenhouse conditions with supplemental light on a 16-hour photoperiod. Tissue was collected 4-6 hours into a light cycle. For seedling tissue, seeds were washed with 80% ethanol/0.1% Triton X-100 for ten minutes, then 30% Clorox bleach/0.1% Triton X-100 for 10 minutes, and were rinsed three times with sterile water. Seeds were imbibed and cold-treated for 72 hours in 0.1% agar prior to plating on MS agar medium overlaid with nylon mesh (Nitex 03-100/47; Sefar America). Agar plates were oriented vertically in a growth chamber with a 16-hour photoperiod. 3-day-old seedlings were collected 4-6 hours into a light cycle. Leaf tissue from *P. syringae*-infected plants was provided by Jeff Dangl and Sarah Grant (University of North Carolina).

Small RNA amplification and sequencing

Small RNA libraries were constructed as described previously [289], with the following modifications. The sequence of the 3'-adaptor was altered (5'ATTGATGGTGCCTACA3'), and the 5'-adaptor was replaced with a chimeric RNA/DNA oligonucleotide [275]. Preadenylated 3'adaptors (Modban) were obtained from Integrated DNA Technologies. Sequences within the RNA portion of the 5'-adaptor were varied to generate eight unique 5'-adaptors (1-1, ATCGTAG**GCAC**CUGAUA; 1-2, ATCGTAG**GCCA**CUGAUA; 1-3, ATCGTAG**GCUG**CUGAUA; 1-4, ATCGTAG**GCGU**CUGAUA; 2-1, ATCGTAG**CGAC**CUGAUA; 2-2, ATCGTAG**CGCA**CUGAUA; 2-3, ATCGTAG**CGUG**CUGAUA; 2-4, ATCGTAG**CGGU**CUGAUA). Throughout small RNA isolation and 3'-adaptor ligation steps, RNA was purified using 17% denaturing PAGE with trace amounts of ^{32}P-radiolabeled RNA transcripts as internal size standards [275]. Nucleic acids were eluted from gel fragments by 16 hour incubation at 4ºC in 0.3 M NaCl, or were electrophoretically transferred DE81 paper and recovered by 15-30 minute incubation at 70ºC in high salt buffer (10 mM Tris-HCl, pH 7.6; 1 mM EDTA, 1 M NaCl, 50 mM L-Arginine) followed by two ethanol precipitation steps. Reverse transcription was primed using a primer complementary to the 3'-adaptor (RT primer; [275]) and cDNA was amplified by PCR using the RT primer and a DNA form of the 5' adaptor, or 3'PCR FusionB (5'-GCCTTGCCAGCCCGCTCAGATTGATGGTGCCTACAG-3') and 5'PCR FusionA (5'-

GCCTCCCTCGCGCCATCAGATCGTAGGCACCTGATA-3') primers. Variant 1-1 of 5'PCR FusionA is shown; underlined bases were varied to match the 5'-adaptor used during RNA ligation steps. PCR primers contained the 'A' and 'B' tag sequences used by 454 Life Sciences during sample processing and sequencing [274]. Asymmetric PCR was done using 5'PCR FusionA and 3'PCR FusionB containing a 3' A_{20}/iSp18 modification. PCR amplicons were gel-purified using 4% Metaphor agarose, 8% PAGE, or 8% PAGE-90% formamide gels. PCR amplicons were eluted from gel fragments using the QIAEX II Gel Extraction Kit (Qiagen) followed by phenol and chloroform extractions and ethanol precipitation, or using the DE81 method (above). DNA amplicons from three to six source libraries were pooled and sequenced by 454 Life Sciences. Small RNA preparations from leaf samples of Col-0 that were either uninoculated or inoculated by *P. syringae* pv. tomato (DC3000*hrcC*) for 1 hr and 3 hr, were also analyzed.

Small RNA parsing and analysis

FASTA formatted files containing 706,567 reads from four 454 Life Sciences sequencing runs were parsed by Perl scripts using the barcoded 5' and 3' adapter sequences to identify the small RNA/adapter boundaries. Parsed small RNA sequences were associated to the correct source library, mapped to the *Arabidopsis* genome (TAIR version 6) by BLAST analysis, and deposited in a MySQL database. Sequences with single nucleotide mismatches were tested by substituting all other nucleotides at the mismatch position, followed by BLAST against the *Arabidopsis* genome. Those sequences that could be unambiguously repaired with a single nucleotide substitution were added to the database with an appropriate annotation. *Arabidopsis* small RNA sequence data were added to the ASRP database (http://asrp.cgrb.oregonstate.edu/db/) [278]. For all comparisons between mutants, small RNA counts were normalized to the library containing the fewest reads for each tissue type.

A second MySQL database was created using a "bin" format in which each chromosome was divided into 250 nt segments containing all corresponding annotation and small RNA data. These data were used to analyze chromosome-wide distribution of small RNA loci and small RNA abundance. Repeat-normalization was done by dividing each small RNA-homologous locus count by the total number of loci with identity to that small RNA. For data shown in Figure 2, locus distribution plots were generated by summing numbers of all or repeat-normalized small RNA loci in 200 bin (50,000 nt) windows with a scroll of 20 bins (10,000 nt) for consecutive data points. For each library, the total, read normalized isolations were

calculated in each window, the data was normalized to Col-0 and the natural log was plotted on heat maps using R v2.3.0 [290].

Database annotation for all protein coding genes was from version 6 of the *Arabidopsis* genome annotation (www.arabidopsis.org). Genes annotated as transposons and retrotransposons, pseudogenes, miRNA and ta-siRNA-generated genes, and all structural RNA genes were analyzed independently from protein-coding genes. The pseudogene set consisted of the 648 pseudogenes that did not overlap with transposons and retroelements. Transposon and retroelement sequences and annotation were from RepBase v10.0.1 [280].

Table 4.1. Small RNA sequence reads parsed from each library.

Genotype	Tissue/Treatment	Total Reads
Col-0	inflorescence	78583
rdr6-15	inflorescence	46892
rdr1-1	inflorescence	19761
rdr2-1	inflorescence	19626
dcl1-7	inflorescence	17839
dcl2-1	inflorescence	13688
dcl3-1	inflorescence	30318
dcl4-2	inflorescence	29497
Col-0	leaf	15826
Col-0	leaf; 1 hpi[a]	18368
Col-0	leaf; 3 hpi[a]	10363
Col-0	seedling	22462
rdr6-15	seedling	23980

[a]hours post-infiltration with *P. syringae* pv. tomato DC3000*hrcC*

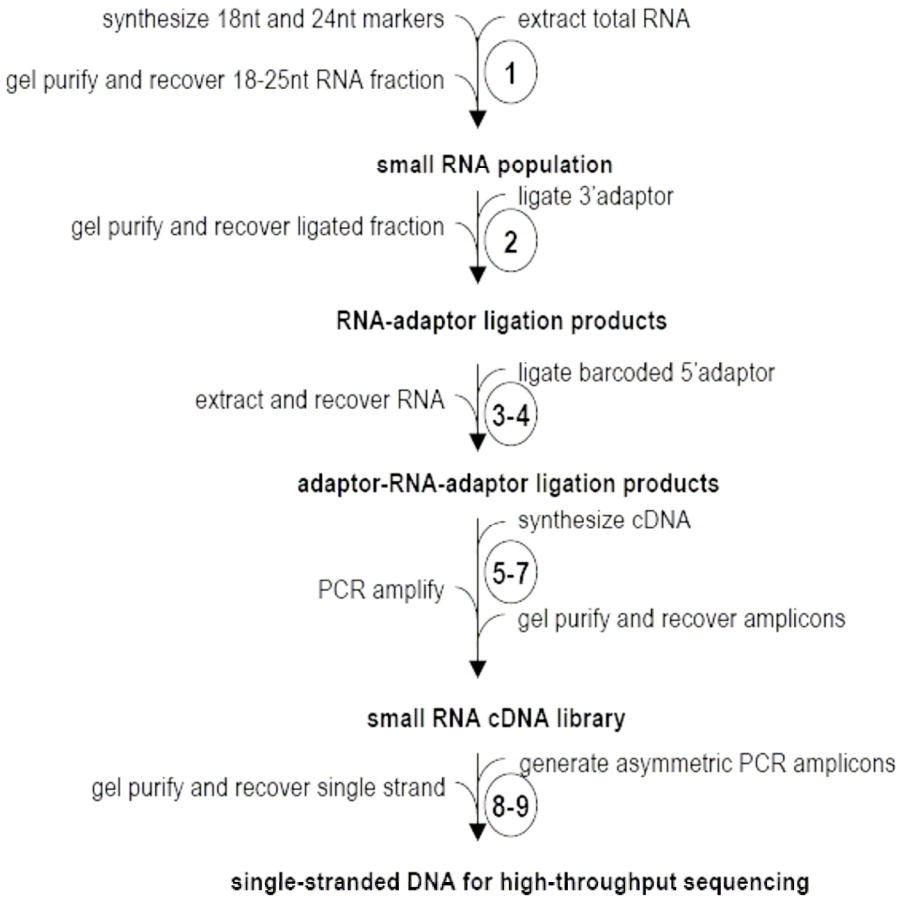

Fig. 4.1. Preparation of template DNA for high throughput sequencing of small RNA libraries. Sequencing template DNA is prepared by RT-PCR amplification of adaptor-ligated size-fractionated RNA. Steps are numbered as indicated in the text.

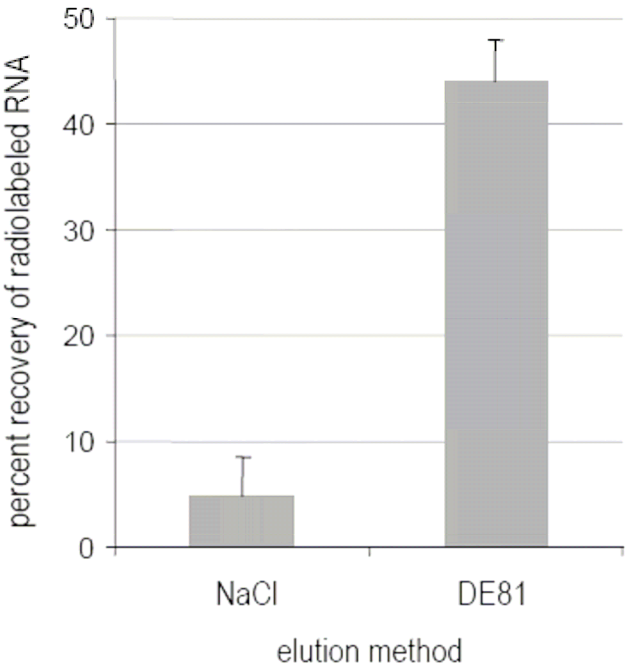

Fig. 4.2. Recovery of RNA from PAGE gels. Percent of input RNA recovered from PAGE gels by diffusion into NaCl solution (NaCl) or by electrophoretic transfer and ion-exchange (DE81) was quantified.

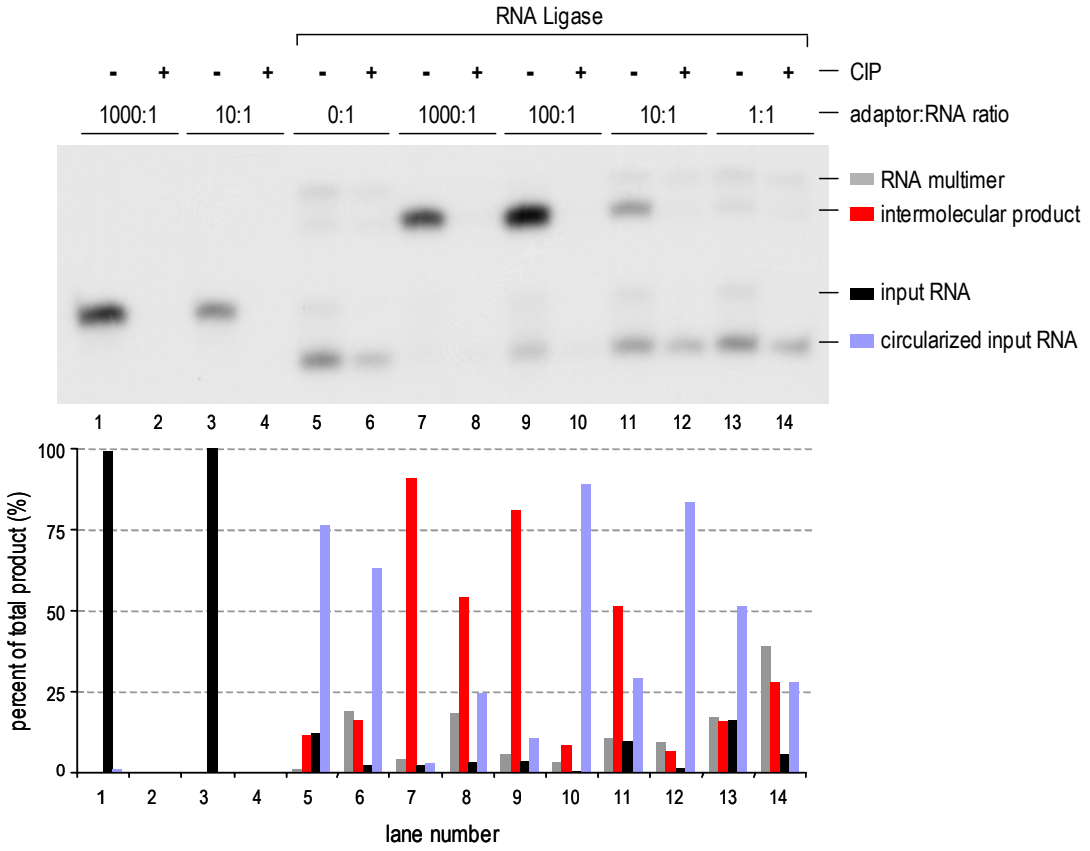

Fig. 4.3. Accumulation of circularized and intermolecular ligation products in reactions containing ATP and T4 RNA ligase. (A) Electrophoretic separation of ligation products and products treated with phosphatase. Radiolabeled 5' phosphate is protected from phosphatase activity in circularized products and multimers. (B) Quantification of RNA distribution among input and circularized, multimerized, and intermolecular ligation product forms.

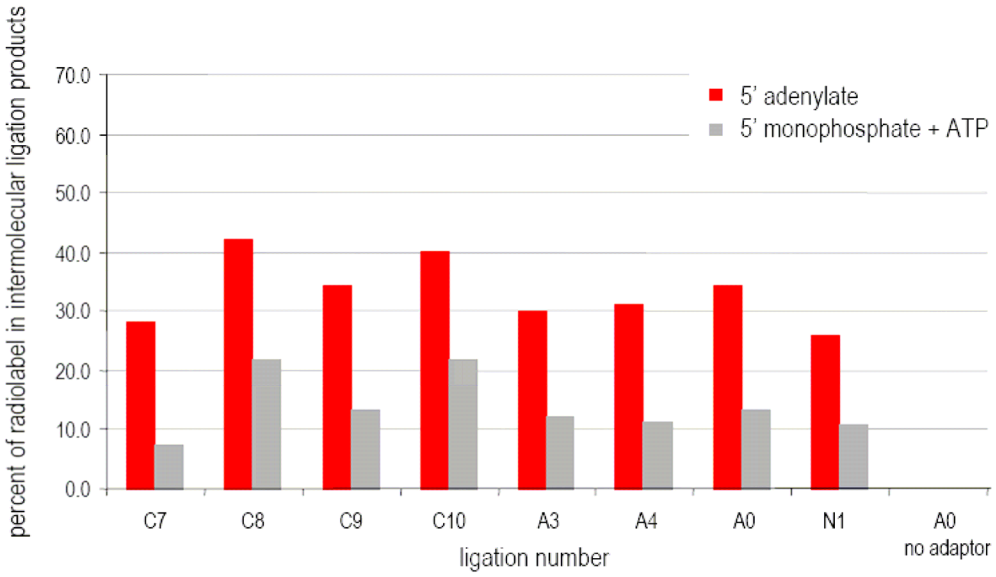

Fig. 4.4. Efficiency of intermolecular ligation. Reactions contained ATP and 5' monophosphorylated adaptors, or pre-adenylated adaptors (5' adenylate) and no exogenous ATP.

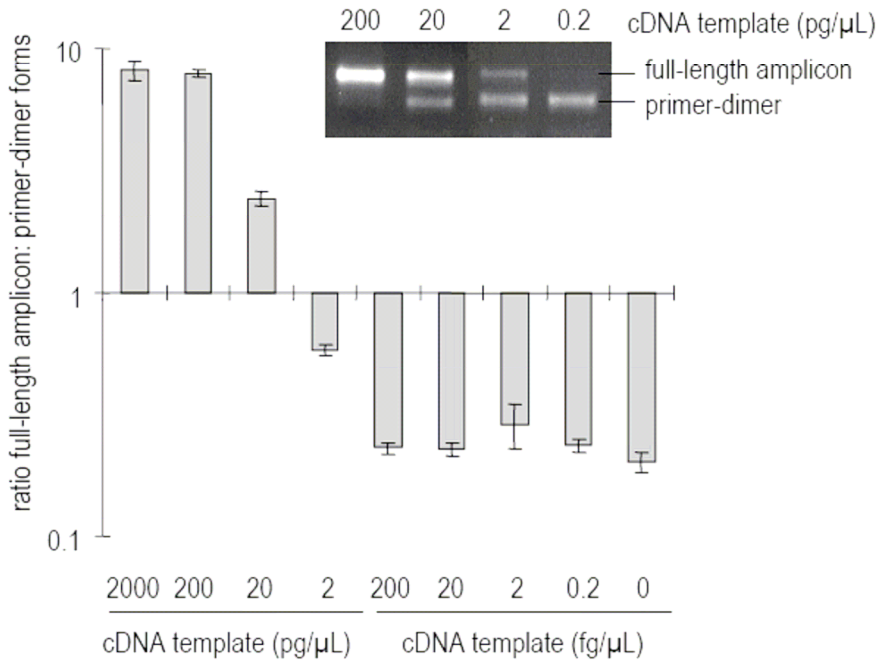

Fig. 4.5. Effect of input cDNA template concentration on accumulation of full-length amplicons and primer-dimer products. Distribution of product among full-length amplicons and primer-dimer forms was analyzed on agarose gels stained with ethidium bromide. Accumulation of full-length amplicons relative to primer-dimer forms is expressed as a ratio.

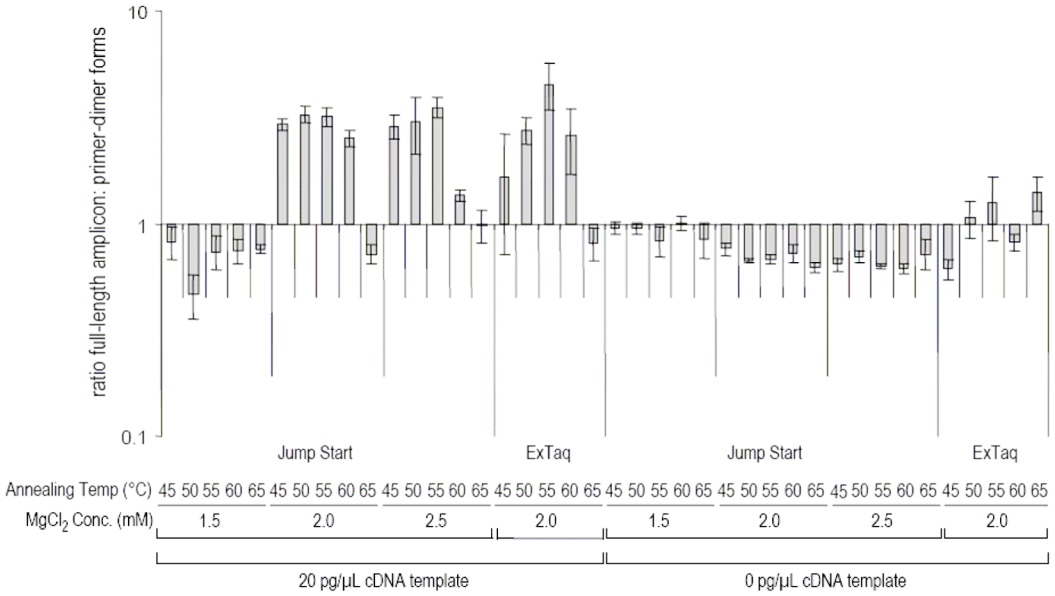

Fig. 4.6. Optimization of annealing temperature and magnesium chloride concentration for PCR amplification of cDNA template. Accumulation of full-length amplicons relative to primer-dimer forms is expressed as a ratio. Magnesium chloride concentration, $MgCl_2$ Conc.

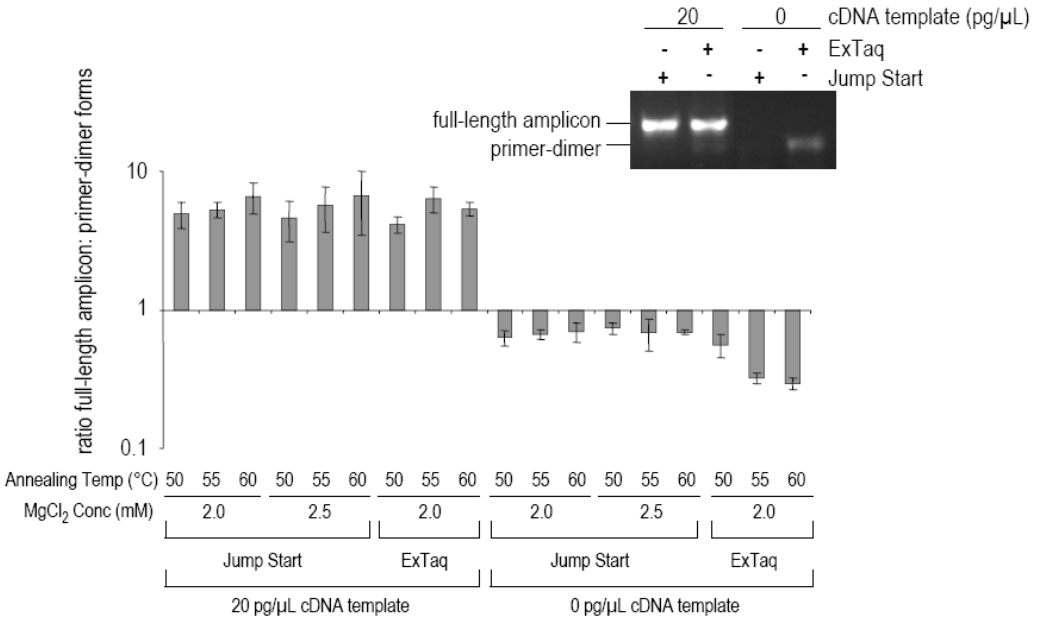

Fig. 4.7. Selection of thermostable polymerase for asymmetric PCR reactions. Distribution of product among full-length amplicons and primer-dimer forms was analyzed on agarose gels stained with ethidium bromide. Accumulation of full-length amplicons relative to primer-dimer forms is expressed as a ratio.

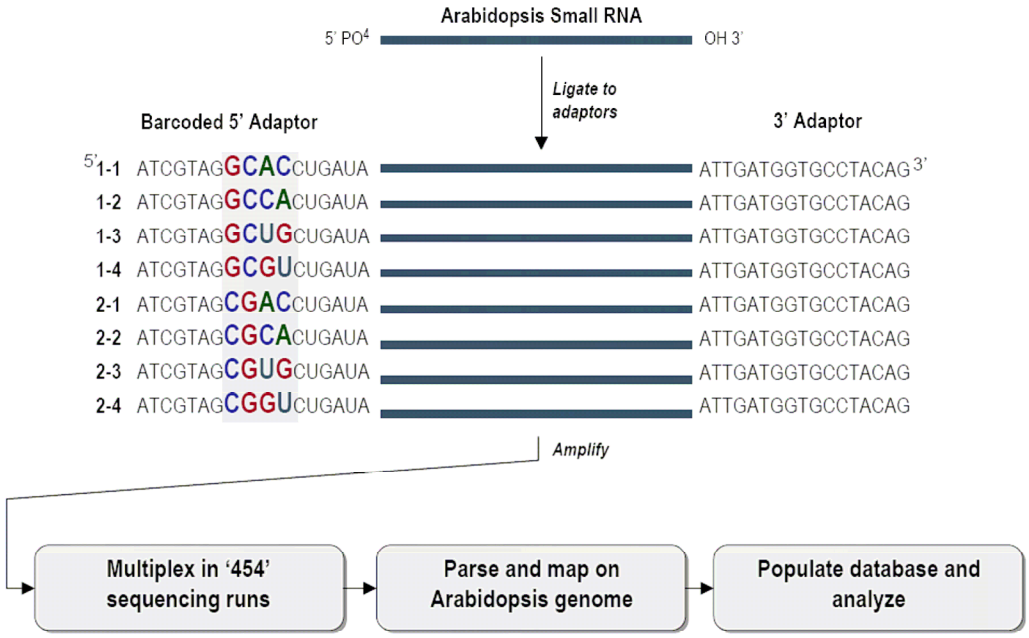

Fig. 4.8. High throughput sequencing and computational analysis of small RNAs.
5' Adaptors contain four-nucleotide molecular barcodes for computational identification of reads originating from each source tissue.

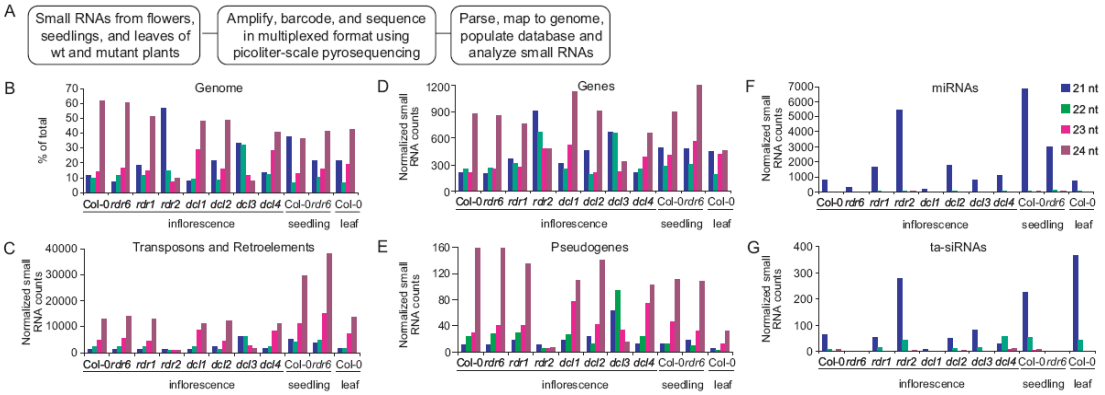

Fig. 4.9. Small RNA sequencing and the distribution of small RNA loci in feature classes. (A) Flowchart for high-throughput sequencing and analysis of small RNAs. (B) to (G) Distribution of small RNAs from wild-type Col-0 plants and rdr6-15, rdr1-1, rdr2-1, dcl1-7, dcl2-1, dcl3-1, and dcl4-2 mutants in the genome (B), transposons and retroelements (C), genes (D), pseudogenes (E), miRNAs (F), and ta-siRNAs (G). In (B), the percentages of small RNAs in each of four size classes within each library are presented. In all other panels, normalized small RNA levels in each feature class are presented.

Fig. 4.10. Distribution of small RNA-generating loci from each chromosome. (A) Scrolling window analysis (50,000 nt window and 10,000 nt scroll) of small RNA loci. Total, repeat-normalized and unique small RNA loci, as well as transposon/retroelement loci, are shown. Abundance of repeat-normalized, library size-normalized counts (individual sequencing reads) are also shown. (B) Scrolling window analysis of repeat-normalized, library size-normalized small RNA abundance in Col-0, rdr2 and dcl3. The summed, 21 and 22 nt size classes (blue, above x-axis) and 24 nt size class (red, below x-axis) were plotted independently. Note that in both (A) and (B), maximum values plotted were capped at the value corresponding to the maximum y-axis value. (C) Scrolling window analysis of relative increase (red) or decrease (green) in repeat-normalized, library size-normalized small RNA abundance in each mutant. Col-0 inflorescence was used as the reference library.

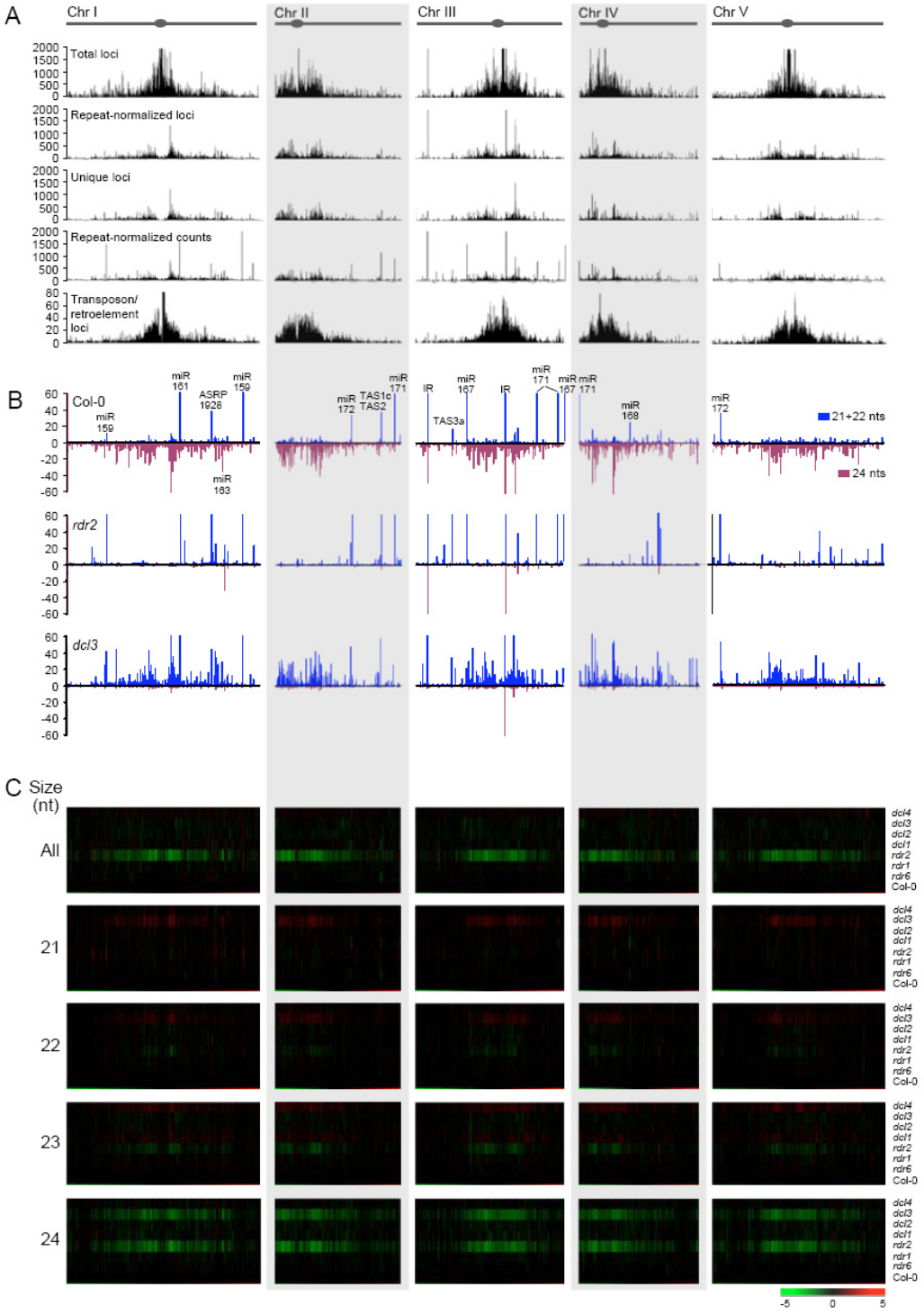

Fig. 4.10. Distribution of small RNA-generating loci from each chromosome.

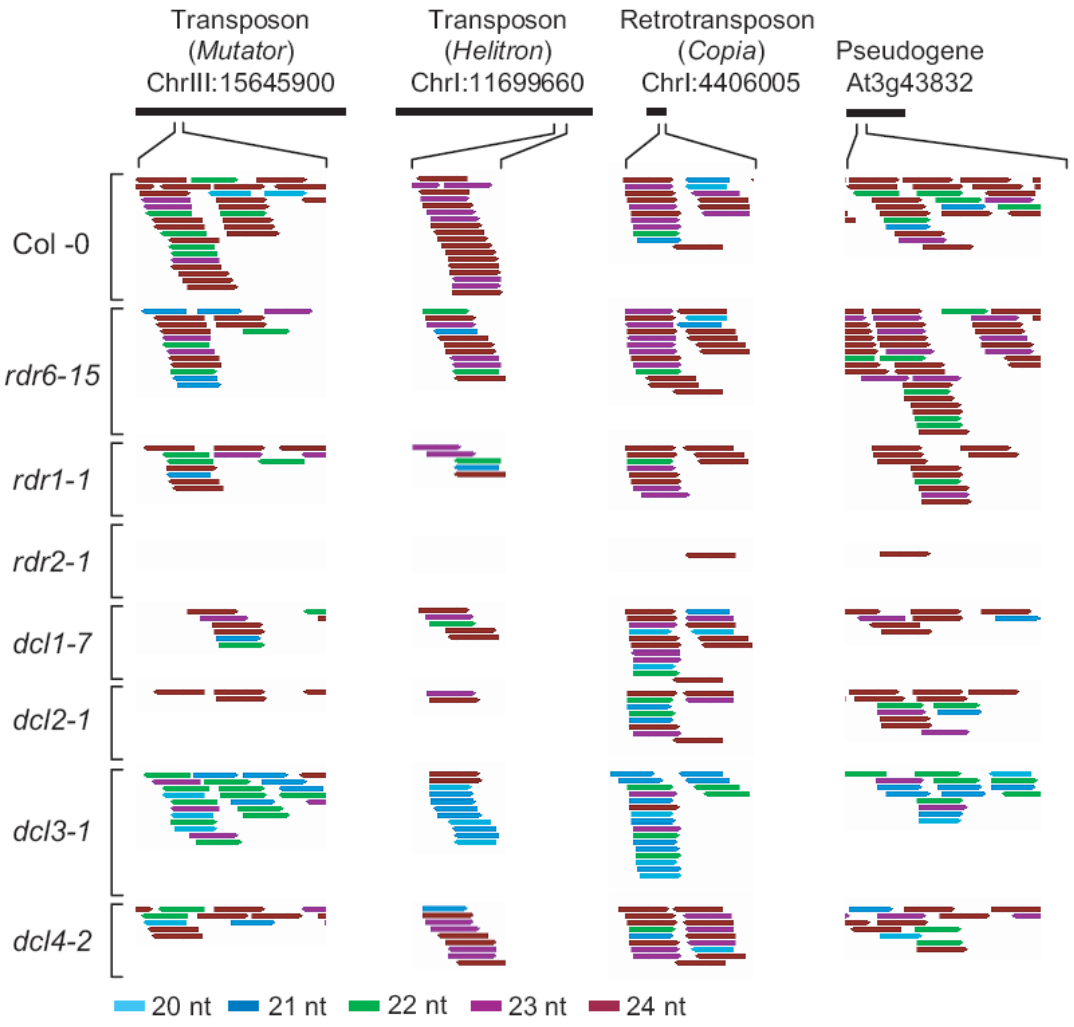

Fig. 4.11. Small RNAs from segments of selected transposons, retroelement and pseudogene. Each unique small RNA is indicated and color-coded based on size.

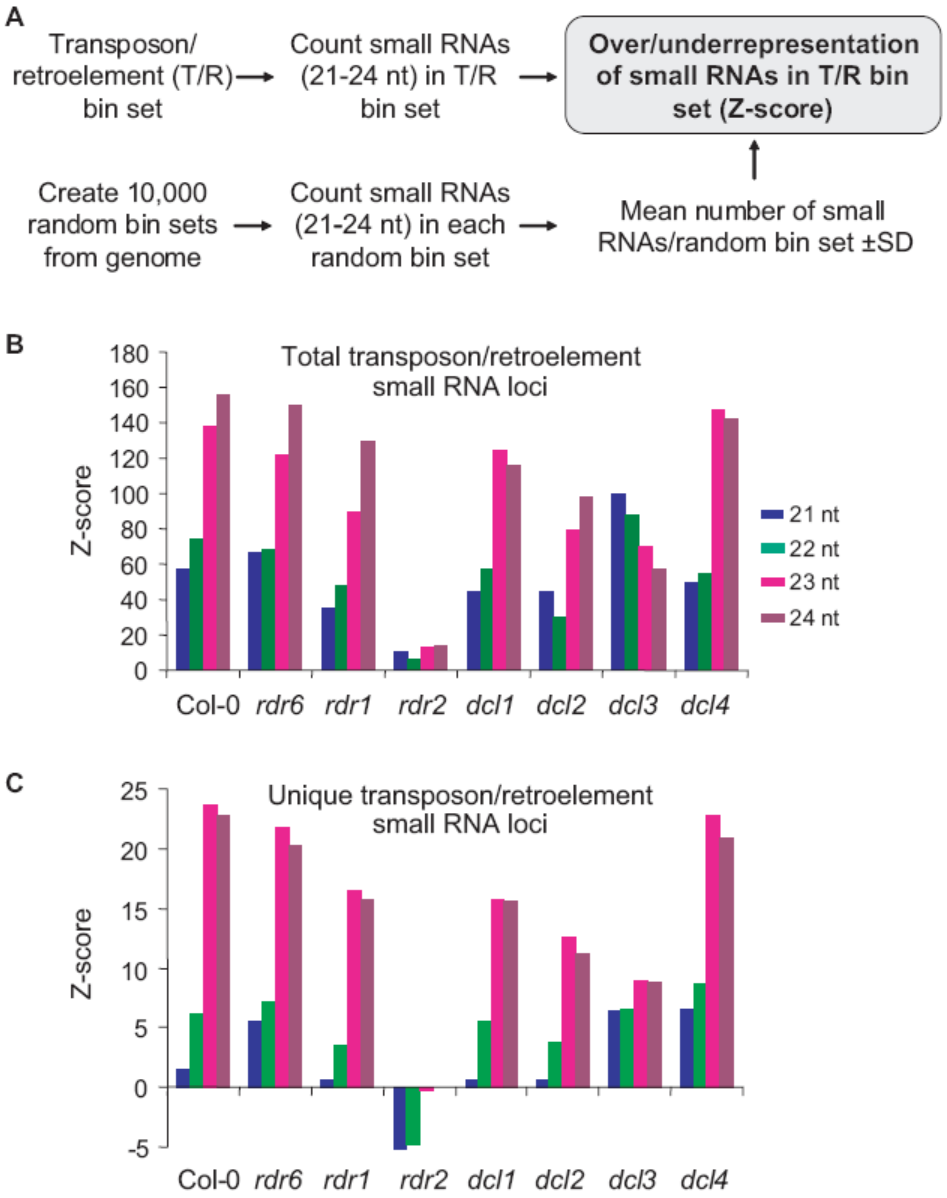

Fig. 4.12. Analysis of small RNAs proximal to transposons and retroelements (T/R). (A) Method to analyze over- or under-representation of small RNAs in T/R bins. (B) and (C) Z-score plots showing over-representation (positive) or under-representation (negative) of small RNA loci in T/R bins from wt Col-0 and mutant plants. Independent analyses were done for each size class from total (B) and unique (C) small RNA loci.

Fig. 4.13. Expression profiling of miRNA families using high-throughput pyrosequencing. (A) Comparison of most-abundant miRNA families between biological replicates of Col-0 inflorescence (inf.) tissue. Normalized reads for each miRNA family member were consolidated. Note that *MIR159*- and *MIR319*- derived members were counted separately, even though they are frequently assigned to the same family [31,57]. (B) Fold-change of miRNAs in dcl1-7 inflorescence versus Col-0 inflorescence (left axis, bars). Total number of reads for each family is indicated (right axis, green line). As the dcl1-7 mutant contained no reads for many miRNA families, fold-change was calculated using normalized reads+1. This had the effect of dampening fold-change values for low-abundance families. (C) Fold-change of miRNA family reads in leaves at 1 hr and 3 hr post-inoculation with *P. syringae* (DC3000*hrcC*) (left axis, bars). Fold-change relative to uninoculated leaves was calculated based on normalized reads as described in panel (B). Total number of reads in the control and inoculated samples is shown (right axis, green line). Grey dashed lines indicate the $p = 0.05$ upper and lower thresholds.

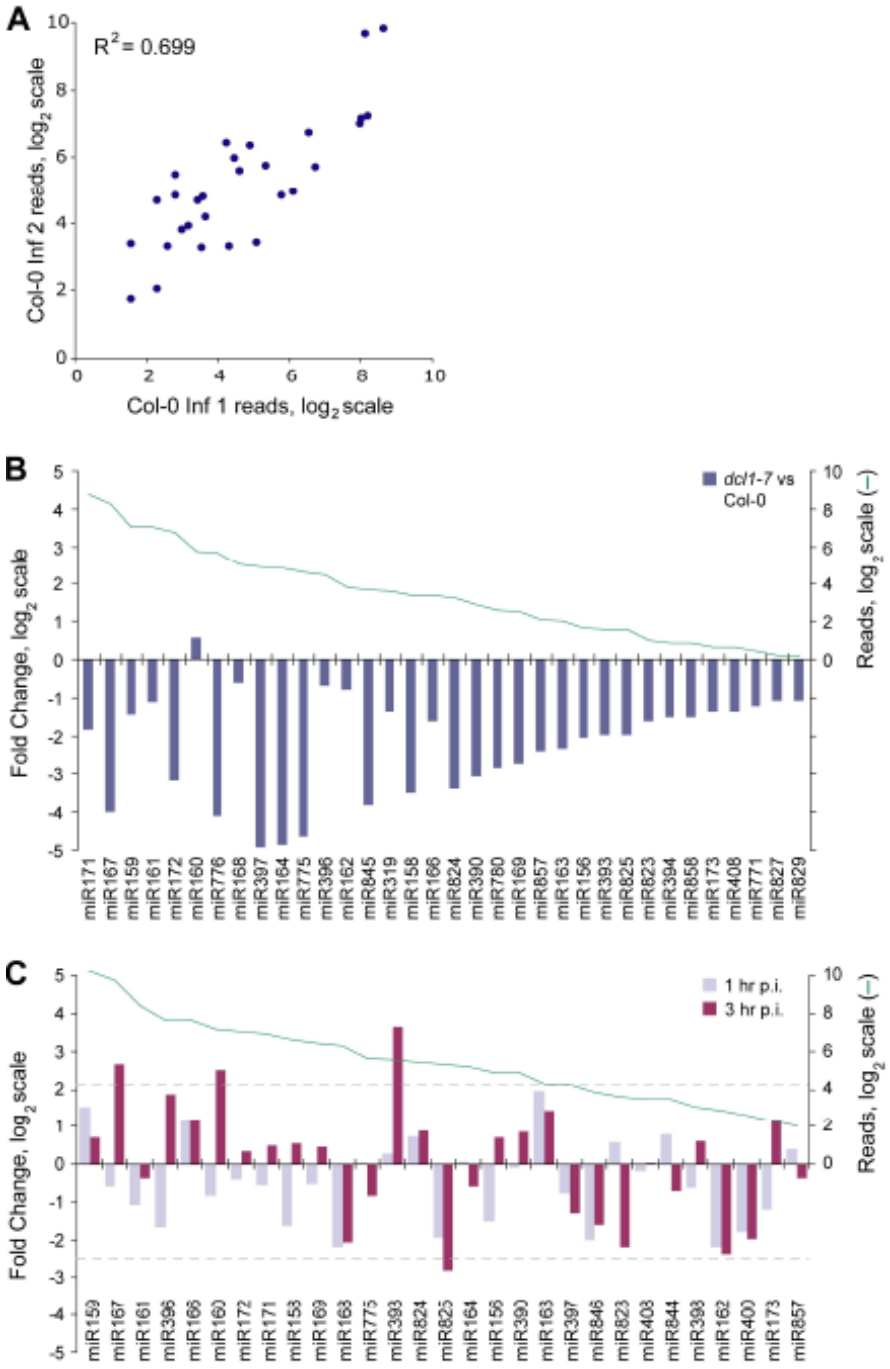

Fig. 4.13. Expression profiling of miRNA families using high-throughput pyrosequencing.

CHAPTER 5
GENERAL CONCLUSION

A Revised Model for Suppression of Antiviral Silencing

Initial evidence of a biological function for RNA silencing came from studies of plant viruses and viral RNA silencing suppressors [190, 192, 193]. The host antiviral RNA silencing response is mediated by virus-derived 21-24nt siRNAs that function to attenuate virus replication and trigger systemic silencing [47, 80, 151]. Systemic virus infection by most, if not all, plant viruses requires expression of one or more RNA silencing suppressor proteins [234]. Although the potyvirus suppressor P1/HC-Pro was among the first suppressors identified [190, 192, 193] and was the subject of many additional studies [91, 92, 200, 202] the suppression mechanism remained unclear. Our work definitively demonstrates small RNA duplex binding by P1/HC-Pro and the unrelated suppressors p21 and p19. Thus, these three suppressors deprive RISC of the sequence-specificity determinant required for targeting and cleavage of viral RNA targets. Our proposed mechanisms provide an explanation consistent with previous observations derived from different systems. Together with recent results presented by others [58, 128, 234], we now have an understanding of the suppression mechanisms of several viruses (Figure 5.1). Three of the suppressors we studied here act through small RNA duplex sequestration; 2b suppresses RISC-mediated target cleavage through a direct, inhibitory association with AGO1, and TCV CP likely inhibits DCL4 activities.

Suppressor Effects on Endogenous RNAi Pathways

The findings that P1/HC-Pro, p21 and p19 capture virus-derived small RNA duplexes were expanded to endogenous microRNA pathways in plants. Each of these suppressors interacted *in vivo* with miRNA/miRNA* duplexes and inhibited HEN1-dependent microRNA methylation (Chapter 3, Figures 3.5 and 3.6). Loss of microRNA methylation is associated with diminished microRNA stability and decreased and heterogenous electrophoretic mobility, both of which may be consequences of microRNA uridylation [26]. Interestingly, decreases in stability and electrophoretic mobility were not detected in suppressor-expressing plants. This suggests that the interactions between suppressors and miRNA/miRNA* duplexes prevent access to uridylation factors as well as to HEN1 methylase.

Suppressor sequestration of miRNA/miRNA* duplexes was associated with ectopic expression of microRNA targets genome-wide (Chapter 2, Figures 2.5 and 2.7) and triggered a series of morphological and developmental abnormalities in plants (Chapter 2, Figures 2.3 and 2.4, and Table 2.2). These abnormalities overlap the panel of symptoms observed in plants expressing these suppressors in virus infections [185, 187]. It seems likely that suppression of

RNA silencing is responsible, at least in part, for the symptomology associated with virus infection. Our work therefore provides answers to long-standing questions about virus disease in plants.

Although TCV CP did not associate with microRNA, moderate effects on microRNA targets were detected (Chapter 3, Figures 2.5 and 2.7). CP inhibited siRNA accumulation from dsGFP (an inverted repeat predicted to form a long double-stranded RNA hairpin) (Chapter 2, Figure 2.1) and thus may interact with DCL4, which processes siRNA from long double-stranded RNA (discussed in Chapter 1). Therefore, CP may interact with other DCL proteins, and even weak inhibition of DCL1 by CP, which is abundant in the *35S:CPHA* transgenic lines, may be sufficient for the observed moderate effect on microRNA targets. An alternative possibility is that CP may indirectly inhibit microRNA pathways through inhibition of DCL4. TCV inhibits accumulation of DCL4/DRB4-dependent tasiRNA and is proposed to inhibit the DCL4/DRB4 complex [57]. Distinct DCL/DRB complexes contribute to biogenesis of tasiRNA and microRNA [32, 78]. In the absence of one or more DCL activities, (proposed) antagonistic interactions and shuffling among the remaining DCL and DRB proteins affect accumulation of different small RNA classes [57]. Thus, expression of CP in transgenic plants may indirectly affect microRNA pathways by altering the pool of DCL proteins available to participate in DCL/DRB interactions.

Genome-Wide Endogenous Small RNA Patterns in *Arabidopsis*

The small RNA profiles generated for *Arabidopsis* provide a genome-wide map of the distribution of small RNA loci as well as the relative contribution of each locus to the small RNA population. Small RNA loci are particularly rich in centromeric/pericentromeric regions, and they are underrepresented in *dcl3* and *rdr2* mutants. The DCL3/RDR2 pathway is associated with the maintenance of heterochromatin and DNA methylation (discussed in Chapter 1) and thus, the abundance of small RNA loci in centromeric/pericentromeric regions may indicate the deployment of small RNA to modulate or repress expression from these regions. Conversely, clusters of small RNA loci are underrepresented in core promoter elements of protein-coding genes [59]. This may indicate that on an evolutionary scale, maintenance of gene expression from protein-coding genes requires prevention of invasion of core promoter elements by small RNA-generating loci. Results of these and other analyses of small RNA profiles [60, 142, 147, 149, 156, 169] demonstrate that small RNA profiling is a powerful tool for the

development of models of gene expression on a genome-wide level, genome organization, and genome evolution.

Silencing Suppressors as a Small RNA Profiling Tool

Duplex intermediates are common to multiple RNA silencing pathways and conserved in all eucaryotes. Suppressors from plant viruses are functional in insect and animal model systems (Fig. 3.1; [291]). Thus, the use of small RNA duplex-binding suppressors to capture small RNA populations for profiling has broad appeal. Expression of the suppressor in a cell type-specific manner would narrow the focus of such a study and reduce the risk of secondary effects due to interference with microRNA-mediated developmental pathways. In a striking example of this, the p19 suppressor expressed in mouse embryonic stem cells selectively coimmunoprecipitates with rRNA-derived small RNA and interacts *in vitro* with small RNA complexed to a longer RNA [292]. These results raise the intriguing possibility that p19 interacts with small RNA-target pairs in certain cell types. This study also revealed a novel small RNA pathway proposed to regulate rRNA expression, although the details of this pathway are unclear.

Identification of microRNA* species is a useful parameter for identification of new microRNA species, and variations in microRNA* sequences from different *MIRNA* family members may indicate preferential expression from one or more loci in a family. The accumulation of miRNA* species in the presence of suppressors suggests that small RNA profiles from suppressor immunoprecipitates could be useful for identification and expression profiling of new microRNA species and loci. Similarly, several microRNA families accumulated to higher levels in suppressor-expressing plants (Fig. 2.6; [185, 201]) and thus suppressors may be useful as an enrichment tool for recovery of rare microRNA species.

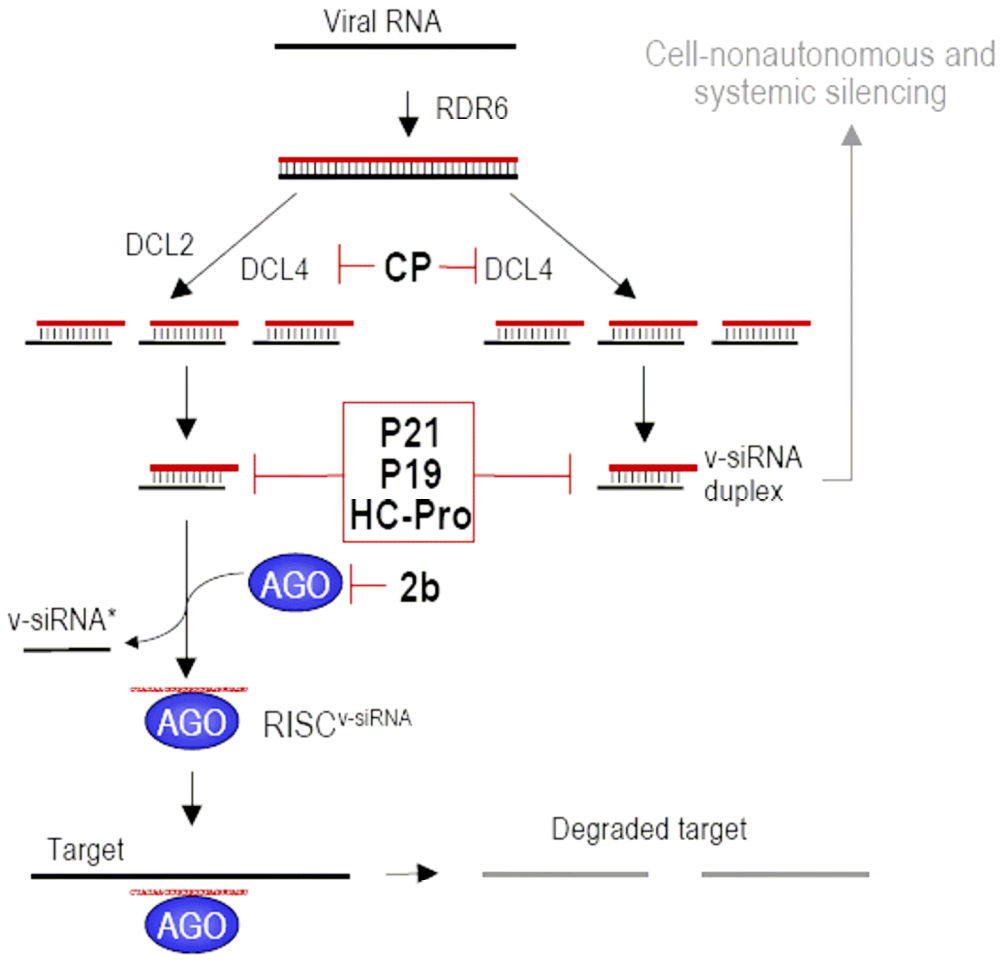

Figure 5.1. Model for suppression of antiviral RNA silencing. *Turnip crinkle virus* suppresses DCL4 activity and the TCV suppressor, CP, may inhibit DCL4 directly. P21, P19 and HC-Pro, suppressors from unrelated viruses *Beet yellows virus,* viruses in the Tombusvirus group, and *Turnip mosaic virus,* are small RNA duplex-binding proteins and sequester a key intermediate in small RNA pathways. The 2b suppressor from *Cucumber mosaic virus* interacts with and inhibits AGO1, a component of the RNA-induced silencing complex degrades target RNA.

BIBLIOGRAPHY

1. Fire, A., Xu, S., Montgomery, M.K., Kostos, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.
2. Kennerdell, J.R., and Carthew, R.W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell *95*, 1017-1026.
3. Waterhouse, P.M., Graham, M.W., and Wang, M.B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. USA *95*, 13959-13964.
4. Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proc Natl Acad Sci USA *95*, 14687-14692.
5. Cogoni, C., Irelan, J.T., Schumacher, M., Schmidhauser, T.J., Selker, E.U., and Macino, G. (1996). Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. EMBO J *15*, 3153-3163.
6. Hamilton, A.J., and Baulcombe, D.C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science *286*, 950-952.
7. Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.
8. Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.
9. Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 428-429.
10. Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. *15*, 188-200.
11. Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.
12. Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. Cell *123*, 1279-1291.
13. Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Jr., Zhu, J.K., Staskawicz, B.J., and Jin, H. (2006). A pathogen-inducible endogenous siRNA in plant immunity. Proc Natl Acad Sci USA *103*, 18002-18007.
14. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.
15. Bartel, D. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.
16. Chan, S.W., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in Arabidopsis thaliana. Nat Rev Genet *6*, 351-360.
17. Du, T., and Zamore, P.D. (2005). microPrimer: the biogenesis and function of microRNA. Development *132*, 4645-4652.

18. Chalfie, M., Horvitz, H.R., and Sulston, J.E. (1981). Mutations that lead to reiterations in the cell lineages of C. elegans. Cell *24*, 59-69.
19. Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.
20. Chen, X. (2005). microRNA biogenesis and function in plants. FEBS Lett *579*, 5923-5931.
21. Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol *57*, 19-53.
22. Wienholds, E., and Plasterk, R.H. (2005). MicroRNA function in animal development. FEBS Lett *579*, 5911-5922.
23. Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.
24. Nykanen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.
25. Ebhardt, H.A., Thi, E.P., Wang, M.B., and Unrau, P.J. (2005). Extensive 3' modification of plant small RNAs is modulated by helper component-proteinase expression. Proc Natl Acad Sci U S A *102*, 13398-13403.
26. Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. Curr Biol *15*, 1501-1507.
27. Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. Science *307*, 932-935.
28. Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.
29. Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. Science *306*, 1377-1380.
30. Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.
31. Vazquez, F., Gasciolli, V., Crete, P., and Vaucheret, H. (2004). The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. Curr Biol *14*, 346-351.
32. Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., Koiwa, H., Seki, M., Shinozaki, K., and Fukuhara, T. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. Plant Mol Biol *57*, 173-188.
33. Han, M.H., Goud, S., Song, L., and Fedoroff, N. (2004). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. Proc Natl Acad Sci U S A *101*, 1093-1098.
34. Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouche, N., Gasciolli, V., and Vaucheret, H. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Curr Biol *16*, 927-932.
35. Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.
36. Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev *20*, 2214-2222.

37. Grivna, S.T., Pyhtila, B., and Lin, H. (2006). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci USA *103*, 13415-13420.
38. Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. Cell.
39. Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev *20*, 1709-1714.
40. Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.
41. Margis, R., Fusaro, A.F., Smith, N.A., Curtin, S.J., Watson, J.M., Finnegan, E.J., and Waterhouse, P.M. (2006). The evolution and diversification of Dicers in plants. FEBS Lett *580*, 2442-2450.
42. Schauer, S.E., Jacobsen, S.E., Meinke, D.W., and Ray, A. (2002). *DICER-LIKE1*: blind men and elephants in *Arabidopsis* development. Trends Plant Sci. *7*, 487-491.
43. Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195-198.
44. Cerutti, H., and Casas-Mollano, J.A. (2006). On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet *50*, 81-99.
45. Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. Science *297*, 2232-2237.
46. Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.
47. Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. EMBO J *21*, 4671-4679.
48. Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716-719.
49. Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.
50. Pal-Bhadra, M., Bhadra, U., and Birchler, J.A. (2002). RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila. Mol Cell *9*, 315-327.
51. Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G., Yates, J.R., 3rd, and Mello, C.C. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.
52. Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654-2659.
53. Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr Biol *13*, 807-818.
54. Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna *12*, 589-597.

55. Simmer, F., Tijsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., and Plasterk, R.H. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Curr Biol *12*, 1317-1319.
56. Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature *427*, 645-649.
57. Bouche, N., Lauressergues, D., Gasciolli, V., and Vaucheret, H. (2006). An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. EMBO J *25*, 3347-3356.
58. Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K.D., Carrington, J.C., and Voinnet, O. (2006). Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science *313*, 68-71.
59. Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-Wide Profiling and Analysis of Arabidopsis siRNAs. PLoS Biol *5*, e57.
60. Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet *38*, 721-725.
61. Kadotani, N., Nakayashiki, H., Tosa, Y., and Mayama, S. (2004). One of the two Dicer-like proteins in the filamentous fungi Magnaporthe oryzae genome is responsible for hairpin RNA-triggered RNA silencing and related small interfering RNA accumulation. J Biol Chem *279*, 44467-44474.
62. Catalanotto, C., Pallotta, M., ReFalo, P., Sachs, M.S., Vayssie, L., Macino, G., and Cogoni, C. (2004). Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in Neurospora crassa. Mol Cell Biol *24*, 2536-2545.
63. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.
64. Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol *3*, e235.
65. Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol *3*, e236.
66. Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. Cell *117*, 69-81.
67. Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.
68. Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.
69. Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep *6*, 961-967.
70. Papp, I., Mette, M.F., Aufsatz, W., Daxinger, L., Schauer, S.E., Ray, A., van der Winden, J., Matzke, M., and Matzke, A.J. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiol. *132*, 1382-1390.

71. Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol *12*, 1484-1495.
72. Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev *16*, 1616-1626.
73. Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol *2*, E104.
74. Kurihara, Y., and Watanabe, Y. (2004). *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Natl Acad Sci U S A *101*, 12753-12758.
75. Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. Rna *12*, 206-212.
76. Song, L., Han, M.H., Lesicka, J., and Fedoroff, N. (2007). Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc Natl Acad Sci U S A *104*, 5437-5442.
77. Qi, Y., Denli, A.M., and Hannon, G.J. (2005). Biochemical specialization within Arabidopsis RNA silencing pathways. Mol Cell *19*, 421-428.
78. Nakazawa, Y., Hiraguri, A., Moriyama, H., and Fukuhara, T. (2007). The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 in vivo and functions in the trans-acting siRNA pathway. Plant Mol Biol *63*, 777-785.
79. Finnegan, E.J., Margis, R., and Waterhouse, P.M. (2003). Posttranscriptional gene silencing Is not compromised in the *Arabidopsis* CARPEL FACTORY (DICER-LIKE1) mutant, a homolog of Dicer-1 from Drosophila. Curr. Biol. *13*, 236-240.
80. Moissiard, G., and Voinnet, O. (2006). RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. Proc Natl Acad Sci U S A *103*, 19593-19598.
81. Gasciolli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol *15*, 1494-1500.
82. Pontes, O., Li, C.F., Nunes, P.C., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell *126*, 79-92.
83. Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E. (2004). RNA silencing genes control de novo DNA methylation. Science *303*, 1336.
84. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. Genes Dev *18*, 2368-2379.
85. Xie, Z., Allen, E., Wilken, A., and Carrington, J.C. (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A *102*, 12984-12989.
86. Dunoyer, P., Himber, C., and Voinnet, O. (2005). DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. Nat Genet *37*, 1356-1360.
87. Qu, F., Ye, X., Hou, G., Sato, S., Clemente, T.E., and Morris, T.J. (2005). RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in Nicotiana benthamiana. J Virol *79*, 15209-15217.

88. Blevins, T., Rajeswaran, R., Shivaprasad, P.V., Beknazariants, D., Si-Ammour, A., Park, H.S., Vazquez, F., Robertson, D., Meins, F., Jr., Hohn, T., and Pooggin, M.M. (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res *34*, 6233-6246.
89. Pantaleo, V., Szittya, G., and Burgyan, J. (2007). Molecular Bases of Viral RNA Targeting by Viral Small Interfering RNA-Programmed RISC. J Virol *81*, 3797-3806.
90. Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C., and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. EMBO J *22*, 4523-4533.
91. Dunoyer, P., Lecellier, C.H., Parizotto, E.A., Himber, C., and Voinnet, O. (2004). Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. Plant Cell *16*, 1235-1250.
92. Chapman, E.J., Prokhnevsky, A.I., Gopinath, K., Dolja, V.V., and Carrington, J.C. (2004). Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. Genes Dev. *18*, 1179-1186.
93. Lakatos, L., Szittya, G., Silhavy, D., and Burgyan, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. EMBO J *23*, 876-884.
94. Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E.J., Carrington, J.C., Liu, Y.P., Dolja, V.V., Calvino, L.F., Lopez-Moya, J.J., and Burgyan, J. (2006). Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. EMBO J *25*, 2768-2780.
95. Qi, Y., Zhong, X., Itaya, A., and Ding, B. (2004). Dissecting RNA silencing in protoplasts uncovers novel effects of viral suppressors on the silencing pathway at the cellular level. Nucleic Acids Res *32*, e179.
96. Lu, R., Maduro, M., Li, F., Li, H.W., Broitman-Maduro, G., Li, W.X., and Ding, S.W. (2005). Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. Nature *436*, 1040-1043.
97. Wilkins, C., Dishongh, R., Moore, S.C., Whitt, M.A., Chow, M., and Machaca, K. (2005). RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. Nature *436*, 1044-1047.
98. Li, H., Li, W.X., and Ding, S.W. (2002). Induction and suppression of RNA silencing by an animal virus. Science *296*, 1319-1321.
99. Mylonakis, E., and Aballay, A. (2005). Worms and flies as genetically tractable animal models to study host-pathogen interactions. Infect Immun *73*, 3833-3841.
100. Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. Science *297*, 2053-2056.
101. Doench, J.G., Petersen, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev *17*, 438-442.
102. Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev *20*, 515-524.
103. Bruno, I., and Wilkinson, M.F. (2006). P-bodies react to stress and nonsense. Cell *125*, 1036-1038.
104. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.

105. Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature *425*, 411-414.
106. Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev *16*, 2491-2496.
107. Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev *16*, 2497-2508.
108. Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.
109. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197.
110. Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.
111. Rivas, F.V., Tolia, N.H., Song, J.J., Aragon, J.P., Liu, J., Hannon, G.J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol *12*, 340-349.
112. Baumberger, N., and Baulcombe, D.C. (2005). *Arabidopsis* ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A *102*, 11928-11933.
113. Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. EMBO J *17*, 170-180.
114. Williams, R.W., and Rubin, G.M. (2002). ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. Proc Natl Acad Sci U S A *99*, 6889-6894.
115. Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.
116. Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., and Barton, M.K. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. Development *126*, 469-481.
117. Moussian, B., Schoof, H., Haecker, A., Jurgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. EMBO J *17*, 1799-1809.
118. Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. Trends Biochem Sci *25*, 481-482.
119. Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the Argonautes. Nat Chem Biol *3*, 36-43.
120. Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747-757.

121. Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.
122. Megosh, H.B., Cox, D.N., Campbell, C., and Lin, H. (2006). The role of PIWI and the miRNA machinery in Drosophila germline determination. Curr Biol *16*, 1884-1894.
123. Yuan, Y.R., Pei, Y., Chen, H.Y., Tuschl, T., and Patel, D.J. (2006). A potential protein-RNA recognition event along the RISC-loading pathway from the structure of A. aeolicus Argonaute with externally bound siRNA. Structure *14*, 1557-1565.
124. Parker, J.S., and Barford, D. (2006). Argonaute: A scaffold for the function of short regulatory RNAs. Trends Biochem Sci *31*, 622-630.
125. Qi, Y., He, X., Wang, X.J., Kohany, O., Jurka, J., and Hannon, G.J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature *443*, 1008-1012.
126. Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D.P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev *18*, 1187-1197.
127. Morel, J.B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. Plant Cell *14*, 629-639.
128. Zhang, X., Yuan, Y.R., Pei, Y., Lin, S.S., Tuschl, T., Patel, D.J., and Chua, N.H. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. Genes Dev *20*, 3255-3268.
129. Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-Dependent 80S Complex Cleaves Targeted mRNAs during RNAi in Drosophila. Cell *117*, 83-94.
130. Zilberman, D., Cao, X., Johansen, L.K., Xie, Z., Carrington, J.C., and Jacobsen, S.E. (2004). Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. Curr Biol *14*, 1214-1220.
131. Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.
132. Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev *12*, 3715-3727.
133. Grishok, A., Tabara, H., and Mello, C.C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science *287*, 2494-2497.
134. Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861-871.
135. Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H.A., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 1-20.
136. Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244-247.
137. Steiner, F.A., and Plasterk, R.H. (2006). Knocking out the Argonautes. Cell *127*, 667-668.

138. Lozovskaya, E.R., Hartl, D.L., and Petrov, D.A. (1995). Genomic regulation of transposable elements in Drosophila. Curr Opin Genet Dev *5*, 768-773.
139. Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev *18*, 1655-1666.
140. Rehwinkel, J., Natalin, P., Stark, A., Brennecke, J., Cohen, S.M., and Izaurralde, E. (2006). Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. Mol Cell Biol *26*, 2965-2975.
141. Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.
142. Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes Dev *20*, 3407-3425.
143. Onodera, Y., Haag, J.R., Ream, T., Nunes, P.C., Pontes, O., and Pikaard, C.S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell *120*, 613-622.
144. Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. (2005). RNA polymerase IV directs silencing of endogenous DNA. Science *308*, 118-120.
145. Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev *19*, 2030-2040.
146. Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell *126*, 93-106.
147. Lu, C., Kulkarni, K., Souret, F.F., Muthuvalliappan, R., Tej, S.S., Poethig, R.S., Henderson, I.R., Jacobsen, S.E., Wang, W., Green, P.J., and Meyers, B.C. (2006). MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. Genome Res *16*, 1276-1288.
148. Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., and Green, P.J. (2005). Elucidation of the small RNA component of the transcriptome. Science *309*, 1567-1569.
149. Zhang, X., Henderson, I.R., Lu, C., Green, P.J., and Jacobsen, S.E. (2007). Role of RNA polymerase IV in plant small RNA metabolism. Proc Natl Acad Sci U S A *104*, 4536-4541.
150. Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D.C. (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell *101*, 543-553.
151. Himber, C., Dunoyer, P., Moissard, G., Ritzenthaler, C., and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. EMBO J *22*, 4523-4533.
152. Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T.A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell *101*, 533-542.
153. Chan, S.W., Zhang, X., Bernatavichute, Y.V., and Jacobsen, S.E. (2006). Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol *4*, e363.

154. Alleman, M., Sidorenko, L., McGinnis, K., Seshadri, V., Dorweiler, J.E., White, J., Sikkink, K., and Chandler, V.L. (2006). An RNA-dependent RNA polymerase is required for paramutation in maize. Nature *442*, 295-298.
155. Axtell, M.J., Jan, C., Rajagopalan, R., and Bartel, D.P. (2006). A Two-Hit Trigger for siRNA Biogenesis in Plants. Cell *127*, 565-577.
156. Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D., and Carrington, J.C. (2007). Genome-Wide Analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 Pathway in Arabidopsis Reveals Dependency on miRNA- and tasiRNA-Directed Targeting. Plant Cell.
157. Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P. (2004). Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. Mol Cell *16*, 69-79.
158. Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell *121*, 207-221.
159. Yoshikawa, M., Peragine, A., Park, M.Y., and Poethig, R.S. (2005). A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. Genes Dev *19*, 2164-2175.
160. Aravin, A.A., Klenov, M.S., Vagin, V.V., Bantignies, F., Cavalli, G., and Gvozdev, V.A. (2004). Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line. Mol Cell Biol *24*, 6742-6750.
161. Gvozdev, V.A., Aravin, A.A., Abramov, Y.A., Klenov, M.S., Kogan, G.L., Lavrov, S.A., Naumova, N.M., Olenkina, O.M., Tulin, A.V., and Vagin, V.V. (2003). Stellate repeats: targets of silencing and modules causing cis-inactivation and trans-activation. Genetica *117*, 239-245.
162. Slotkin, R.K., Freeling, M., and Lisch, D. (2003). Mu killer causes the heritable inactivation of the Mutator family of transposable elements in Zea mays. Genetics *165*, 781-797.
163. Slotkin, R.K., Freeling, M., and Lisch, D. (2005). Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. Nat Genet *37*, 641-644.
164. Tuteja, J.H., Clough, S.J., Chan, W.C., and Vodkin, L.O. (2004). Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in Glycine max. Plant Cell *16*, 819-835.
165. Senda, M., Masuta, C., Ohnishi, S., Goto, K., Kasai, A., Sano, T., Hong, J.S., and MacFarlane, S. (2004). Patterning of virus-infected Glycine max seed coat is associated with suppression of endogenous silencing of chalcone synthase genes. Plant Cell *16*, 807-818.
166. Fobis-Loisy, I., Miege, C., and Gaude, T. (2004). Molecular evolution of the s locus controlling mating in the brassicaceae. Plant Biol (Stuttg) *6*, 109-118.
167. Kusaba, M., Dwyer, K., Hendershot, J., Vrebalov, J., Nasrallah, J.B., and Nasrallah, M.E. (2001). Self-incompatibility in the genus Arabidopsis: characterization of the S locus in the outcrossing A. lyrata and its autogamous relative A. thaliana. Plant Cell *13*, 627-643.
168. Allen, E., Xie, Z., Gustafson, A.M., Sung, G.H., Spatafora, J.W., and Carrington, J.C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. Nat Genet *36*, 1282-1290.
169. Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L., and Carrington, J.C. (2007). High-Throughput Sequencing of Arabidopsis microRNAs: Evidence for Frequent Birth and Death of MIRNA Genes. PLoS ONE *2*, e219.

170. Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet *5*, 396-400.
171. Nair, V., and Zavolan, M. (2006). Virus-encoded microRNAs: novel regulators of gene expression. Trends Microbiol *14*, 169-175.
172. Zhao, Y., and Srivastava, D. (2007). A developmental view of microRNA function. Trends Biochem Sci.
173. Hull, R. (2001). Matthews' Plant Virology, Fourth Edition (Academic Press).
174. Voinnet, O., Pinto, V.M., and Baulcombe, D.C. (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proc. Natl. Acad. Sci. USA *96*, 14147-14152.
175. Waterhouse, P.M., Wang, M.B., and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. Nature *411*, 834-842.
176. Vaucheret, H., Beclin, C., and Fagard, M. (2001). Post-transcriptional gene silencing in plants. J Cell Sci. *114*, 3083-3091.
177. Tijsterman, M., Ketting, R.F., and Plasterk, R.H. (2002). The genetics of RNA silencing. Annu Rev Genet. *36*, 489-519.
178. Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses, and RNA silencing. Science *296*, 1270-1273.
179. Finnegan, E.J., and Matzke, M.A. (2003). The small RNA world. J Cell Sci *116*, 4689-4693.
180. Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.
181. Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell *113*, 673-676.
182. Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., and Bartel, D.P. (2002). Prediction of plant microRNA targets. Cell *110*, 513-520.
183. Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature *425*, 257-263.
184. Xie, Z., Kasschau, K.D., and Carrington, J.C. (2003). Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. Curr Biol *13*, 784-789.
185. Kasschau, K.D., Xie, Z., Allen, E., Llave, C., Chapman, E.J., Krizan, K.A., and Carrington, J.C. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. Dev Cell *4*, 205-217.
186. Reed, J.C., Kasschau, K.D., Prokhnevsky, A.I., Gopinath, K., Pogue, G.P., Carrington, J.C., and Dolja, V.V. (2002). Suppressor of RNA silencing encoded by Beet yellows virus INCORRECT-use2003. Virology *306*, 203-209.
187. Silhavy, D., Molnar, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., and Burgyan, J. (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. EMBO J *21*, 3070-3080.
188. Qu, F., Ren, T., and Morris, T.J. (2003). The coat protein of *Turnip crinkle virus* suppresses posttranscriptional gene silencing at an early step. J Virol. *77*, 511-522.
189. Thomas, C.L., V., L., Lederer, C., and Maule, A.J. (2003). Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. Virology *306*, 33-41.
190. Brigneti, G., Voinnet, O., Wan-Xiang, L., Ding, S.W., and Baulcombe, D.C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing. EMBO J *17*, 6739-6746.

191. Guo, H.S., and Ding, S.W. (2002). A viral protein inhibits the long range signaling activity of the gene silencing signal. EMBO J *21*, 398-407.
192. Anandalakshmi, R., Pruss, G.J., Ge, X., Marathe, R., Smith, T.H., and Vance, V.B. (1998). A viral suppressor of gene silencing in plants. Proc. Natl. Acad. Sci. USA *95*, 13079-13084.
193. Kasschau, K.D., and Carrington, J.C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. Cell *95*, 461-470.
194. Dolja, V.V., and Koonin, E.V. (1991). Phylogeny of capsid proteins of small icosahedral RNA plant viruses. J Gen. Virol. *72*, 1481-1486.
195. Koonin, E.V., Choi, G.H., Nuss, D.L., Shapira, R., and Carrington, J.C. (1991). Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive-strand RNA plant viruses. Proc. Natl. Acad. Sci. USA *88*, 10647-10651.
196. Vargason, J.M., Szittya, G., Burgyan, J., and Tanaka Hall, T.M. (2003). Size selective recognition of siRNA by an RNA silencing suppressor. Cell *115*, 799-811.
197. Kasschau, K.D., Cronin, S., and Carrington, J.C. (1997). Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. Virology *228*, 251-262.
198. Chiba, M., Reed, J.C., Prokhnevsky, A.I., Chapman, E.J., Mawassi, M., Koonin, E.V., Carrington, J.C., and Dolja, V.V. (2006). Diverse suppressors of RNA silencing enhance agroinfection by a viral replicon. Virology *346*, 7-14.
199. Johansen, L.K., and Carrington, J.C. (2001). Silencing on the spot: induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. Plant Physiol. *126*, 930-938.
200. Mallory, A.C., Ely, L., Smith, T.H., Marathe, R., Anandalakshmi, R., Fagard, M., Vaucheret, H., Pruss, G., Bowman, L., and Vance, V.B. (2001). HC-Pro Suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal. Plant Cell *13*, 571-583.
201. Mallory, A.C., Reinhart, B.J., Bartel, D., Vance, V.B., and Bowman, L.H. (2002). A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. Proc. Natl. Acad. Sci. USA *99*, 15228-15233.
202. Llave, C., Kasschau, K.D., and Carrington, J.C. (2000). Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. Proc Natl Acad Sci U S A *97*, 13401-13406.
203. Vaistij, F.E., Jones, L., and Baulcombe, D.C. (2002). Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. Plant Cell *14*, 857-867.
204. Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Gorlach, J. (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. Plant Cell *13*, 1499-1510.
205. Ye, K., Malinina, L., and Patel, D.J. (2003). Recognition of small interfering RNA by a viral suppressor of RNA silencing. Nature *426*, 874-878.
206. Peng, C.W., and Dolja, V.V. (2000). Leader proteinase of the beet yellows closterovirus: mutation analysis of the function in genome amplification. J Virol. *74*, 9766-9770.
207. Carrington, J.C., and Freed, D.D. (1990). Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. J Virol. *64*, 1590-1597.

208. Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *257*, 8569-8572.
209. Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J *16*, 735-743.
210. Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19*, 185-193.
211. Jones-Rhoades, M.W., and Bartel, D.P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell *14*, 787-799.
212. Hochberg, Y., and Benjamini, Y. (1990). More powerful procedures for multiple significance testing. Stat Med *9*, 811-818.
213. Slade, D.E., Johnston, R.E., and Dougherty, W.G. (1989). Generation and characterization of monoclonal antibodies reactive with the 49-kDa proteinase of tobacco etch virus. Virology *173*, 499-508.
214. Mello, C.C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. Nature *431*, 338-342.
215. Zamore, P.D., and Haley, B. (2005). Ribo-gnome: the big world of small RNAs. Science *309*, 1519-1524.
216. Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356-363.
217. Yu, L., Yu, X., Shen, R., and He, Y. (2005). HYL1 gene maintains venation and polarity of leaves. Planta *221*, 231-242.
218. Leuschner, P.J., Obernosterer, G., and Martinez, J. (2005). MicroRNAs: Loquacious speaks out. Curr Biol *15*, R603-605.
219. Hutvagner, G. (2005). Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Lett *579*, 5850-5857.
220. Hammond, S.M. (2005). Dicing and slicing The core machinery of the RNA interference pathway. FEBS Lett *579*, 5822-5829.
221. Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell *123*, 607-620.
222. Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.
223. Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. EMBO J *23*, 4727-4737.
224. Hall, T.M. (2005). Structure and function of argonaute proteins. Structure (Camb) *13*, 1403-1408.
225. Herr, A.J. (2005). Pathways through the small RNA world of plants. FEBS Lett *579*, 5879-5888.
226. Qi, Y., and Hannon, G.J. (2005). Uncovering RNAi mechanisms in plants: biochemistry enters the foray. FEBS Lett *579*, 5899-5903.
227. Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J. (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet *37*, 761-765.
228. Matzke, M.A., and Birchler, J.A. (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet *6*, 24-35.

229. Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res *34*, 667-675.
230. Boutet, S., Vazquez, F., Liu, J., Beclin, C., Fagard, M., Gratias, A., Morel, J.B., Crete, P., Chen, X., and Vaucheret, H. (2003). *Arabidopsis* HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Curr Biol *13*, 843-848.
231. Tang, G., Reinhart, B.J., Bartel, D.P., and Zamore, P.D. (2003). A biochemical framework for RNA silencing in plants. Genes & Dev. *17*, 49-63.
232. Williams, L., Carles, C.C., Osmont, K.S., and Fletcher, J.C. (2005). A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the *Arabidopsis ARF2*, *ARF3*, and *ARF4* genes. Proc Natl Acad Sci U S A *102*, 9703-9708.
233. Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.S., Givan, S.A., Kasschau, K.D., and Carrington, J.C. (2007). Phasing Patterns and Genome-Wide Profiles of RDR6/DCL4-Dependent Small RNAs in *Arabidopsis*. The Plant Cell *19*, in press.
234. Li, F., and Ding, S.W. (2006). Virus counterdefense: diverse strategies for evading the RNA-silencing immunity. Annu Rev Microbiol *60*, 503-531.
235. Fusaro, A.F., Matthew, L., Smith, N.A., Curtin, S.J., Dedic-Hagan, J., Ellacott, G.A., Watson, J.M., Wang, M.B., Brosnan, C., Carroll, B.J., and Waterhouse, P.M. (2006). RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. EMBO Rep *7*, 1168-1175.
236. Schwach, F., Vaistij, F.E., Jones, L., and Baulcombe, D.C. (2005). An RNA-dependent RNA polymerase prevents meristem invasion by potato virus X and is required for the activity but not the production of a systemic silencing signal. Plant Physiol *138*, 1842-1852.
237. Waterhouse, P.M., and Fusaro, A.F. (2006). Plant science. Viruses face a double defense by plant small RNAs. Science *313*, 54-55.
238. Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev *20*, 759-771.
239. Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Jr., Hohn, T., and Pooggin, M.M. (2006). Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res *34*, 462-471.
240. Molnar, A., Csorba, T., Lakatos, L., Varallyay, E., Lacomme, C., and Burgyan, J. (2005). Plant virus-derived small interfering RNAs originate predominantly from highly structured single-stranded viral RNAs. J Virol *79*, 7812-7818.
241. Ho, T., Pallett, D., Rusholme, R., Dalmay, T., and Wang, H. (2006). A simplified method for cloning of short interfering RNAs from Brassica juncea infected with Turnip mosaic potyvirus and Turnip crinkle carmovirus. J Virol Methods *136*, 217-223.
242. Voinnet, O. (2005). Induction and suppression of RNA silencing: insights from viral infections. Nat Rev Genet *6*, 206-220.
243. Silhavy, D., and Burgyan, J. (2004). Effects and side-effects of viral RNA silencing suppressors on short RNAs. Trends Plant Sci *9*, 76-83.
244. Merai, Z., Kerenyi, Z., Molnar, A., Barta, E., Valoczi, A., Bisztray, G., Havelda, Z., Burgyan, J., and Silhavy, D. (2005). Aureusvirus P14 is an efficient RNA silencing suppressor that binds double-stranded RNAs without size specificity. J Virol *79*, 7217-7226.

245. Li, W.X., Li, H., Lu, R., Li, F., Dus, M., Atkinson, P., Brydon, E.W., Johnson, K.L., Garcia-Sastre, A., Ball, L.A., Palese, P., and Ding, S.W. (2004). Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. Proc Natl Acad Sci U S A *101*, 1350-1355.
246. Pfeffer, S., Dunoyer, P., Heim, F., Richards, K.E., Jonard, G., and Zeigler-Graff, V. (2002). P0 of Beet Western Yellows Virus is a suppressor of posttrancriptional gene silencing. J Virol. *76*, 6815-6824.
247. Pazhouhandeh, M., Dieterle, M., Marrocco, K., Lechner, E., Berry, B., Brault, V., Hemmer, O., Kretsch, T., Richards, K.E., Genschik, P., and Ziegler-Graff, V. (2006). F-box-like domain in the polerovirus protein P0 is required for silencing suppressor function. Proc Natl Acad Sci U S A *103*, 1994-1999.
248. Anandalakshmi, R., Marathe, R., Ge, X., Herr, J.M., Mau, C., Mallory, A., Pruss, G., Bowman, L., and Vance, V.B. (2000). A calmodulin-related protein that suppresses posttranscriptional gene silencing in plants. Science *290*, 142-144.
249. Ruiz-Ferrer, V., Boskovic, J., Alfonso, C., Rivas, G., Llorca, O., Lopez-Abella, D., and Lopez-Moya, J.J. (2005). Structural analysis of tobacco etch potyvirus HC-pro oligomers involved in aphid transmission. J Virol *79*, 3758-3765.
250. Alefelder, S., Patel, B.K., and Eckstein, F. (1998). Incorporation of terminal phosphorothioates into oligonucleotides. Nucleic Acids Res *26*, 4983-4988.
251. Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.
252. Yu, B., Chapman, E.J., Yang, Z., Carrington, J.C., and Chen, X. (2006). Transgenically expressed viral RNA silencing suppressors interfere with microRNA methylation in Arabidopsis. FEBS Lett *580*, 3117-3120.
253. Ye, K., and Patel, D.J. (2005). RNA silencing suppressor p21 of Beet yellows virus forms an RNA binding octameric ring structure. Structure *13*, 1375-1384.
254. Uhrig, J.F., Canto, T., Marshall, D., and MacFarlane, S.A. (2004). Relocalization of nuclear ALY proteins to the cytoplasm by the tomato bushy stunt virus P19 pathogenicity protein. Plant Physiol *135*, 2411-2423.
255. Reed, J.C., Kasschau, K.D., Prokhnevsky, A.I., Gopinath, K., Pogue, G.P., Carrington, J.C., and Dolja, V.V. (2003). Suppressor of RNA silencing encoded by Beet yellows virus. Virology *306*, 203-209.
256. Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. (2005). Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci U S A *102*, 3691-3696.
257. Zamore, P.D. (2004). Plant RNAi: How a viral silencing suppressor inactivates siRNA. Curr Biol *14*, R198-200.
258. Dolja, V.V., Kreuze, J.F., and Valkonen, J.P. (2006). Comparative and functional genomics of closteroviruses. Virus Res *117*, 38-51.
259. Koonin, E.V., Boyko, V.P., and Dolja, V.V. (1991). Small cysteine-rich proteins of different groups of plant RNA viruses are related to different families of nucleic acid-binding proteins. Virology *181*, 395-398.
260. Chellappan, P., Vanitharani, R., and Fauquet, C.M. (2005). MicroRNA-binding viral protein interferes with Arabidopsis development. Proc Natl Acad Sci U S A *102*, 10381-10386.
261. Haley, B., Tang, G., and Zamore, P.D. (2003). In vitro analysis of RNA interference in Drosophila melanogaster. Methods *30*, 330-336.

262. Almeida, R., and Allshire, R.C. (2005). RNA silencing and genome regulation. Trends Cell Biol *15*, 251-258.
263. Tomari, Y., and Zamore, P.D. (2005). Perspective: machines for RNAi. Genes Dev *19*, 517-529.
264. Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.
265. Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. Trends Genet *22*, 268-280.
266. Bernstein, E., and Allis, C.D. (2005). RNA meets chromatin. Genes Dev *19*, 1635-1655.
267. Yu, D., Fan, B., MacFarlane, S.A., and Chen, Z. (2003). Analysis of the involvement of an inducible *Arabidopsis* RNA-dependent RNA polymerase in antiviral defense. Mol Plant Microbe Interact *16*, 206-216.
268. Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature *431*, 364-370.
269. Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol *1*, 420-428.
270. Fahlgren, N., Montgomery, T.A., Howell, M.D., Allen, E., Dvorak, S.K., Alexander, A.L., and Carrington, J.C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. Curr Biol *16*, 939-944.
271. Garcia, D., Collier, S.A., Byrne, M.E., and Martienssen, R.A. (2006). Specification of leaf polarity in Arabidopsis via the trans-acting siRNA pathway. Curr Biol *16*, 933-938.
272. Hunter, C., Willmann, M.R., Wu, G., Yoshikawa, M., de la Luz Gutierrez-Nava, M., and Poethig, S.R. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. Development *133*, 2973-2981.
273. Dunoyer, P., and Voinnet, O. (2005). The complex interplay between plant viruses and host RNA-silencing pathways. Curr Opin Plant Biol *8*, 415-423.
274. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., and Rothberg, J.M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376-380.
275. Lau, N.C., Lim, E.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.
276. Sambrook J, F.E., Maniatis T (1989). Molecular Cloning: a laboratory manual, 2nd edn. *Cold Spring Harbor, NY*, Cold Spring Harbor Laboratory Press.
277. Ho, C.K., Wang, L.K., Lima, C.D., and Shuman, S. (2004). Structure and mechanism of RNA ligase. Structure *12*, 327-339.
278. Gustafson, A.M., Allen, E., Givan, S., Smith, D., Carrington, J.C., and Kasschau, K.D. (2005). ASRP: the Arabidopsis Small RNA Project Database. Nucleic Acids Res *33*, D637-640.

279. Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., and Carrington, J.C. (2005). Expression of Arabidopsis MIRNA genes. Plant Physiol *138*, 2145-2154.
280. Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. Trends Genet *16*, 418-420.
281. Jones, J.D., and Dangl, J.L. (2006). The plant immune system. Nature *444*, 323-329.
282. Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O., and Jones, J.D. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science *312*, 436-439.
283. Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V., and Martienssen, R. (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.
284. May, B.P., Lippman, Z.B., Fang, Y., Spector, D.L., and Martienssen, R.A. (2005). Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. PLoS Genet *1*, e79.
285. Lisch, D., Carey, C.C., Dorweiler, J.E., and Chandler, V.L. (2002). A mutation that prevents paramutation in maize also reverses Mutator transposon methylation and silencing. Proc Natl Acad Sci U S A *99*, 6130-6135.
286. McGinnis, K.M., Springer, C., Lin, Y., Carey, C.C., and Chandler, V. (2006). Transcriptionally silenced transgenes in maize are activated by three mutations defective in paramutation. Genetics *173*, 1637-1647.
287. Woodhouse, M.R., Freeling, M., and Lisch, D. (2006). Initiation, Establishment, and Maintenance of Heritable MuDR Transposon Silencing in Maize Are Mediated by Distinct Factors. PLoS Biol *4*, e339.
288. Woodhouse, M.R., Freeling, M., and Lisch, D. (2006). The mop1 (mediator of paramutation1) mutant progressively reactivates one of the two genes encoded by the MuDR transposon in maize. Genetics *172*, 579-592.
289. Llave, C., Kasschau, K.D., Rector, M.A., and Carrington, J.C. (2002). Endogenous and silencing-associated small RNAs in plants. Plant Cell *14*, 1605-1619.
290. Team, R.D.C. (2006). R: A language and environment for statistical computing, 2.3.0 ed, R Foundation for Statistical Computing: Vienna, Austria.
291. Lecellier, C.H., Dunoyer, P., Arar, K., Lehmann-Che, J., Eyquem, S., Himber, C., Saib, A., and Voinnet, O. (2005). A cellular microRNA mediates antiviral defense in human cells. Science *308*, 557-560.
292. Calabrese, J.M., and Sharp, P.A. (2006). Characterization of the short RNAs bound by the P19 suppressor of RNA silencing in mouse embryonic stem cells. Rna *12*, 2092-2102.

APPENDIX

MIRNA GENE AND TARGET FAMILIES IN *ARABIDOPSIS*

MIRNA family[a]	Loci	Conserved[a]	Target family	Target validation
miR156/miR157	12	Y	Squamosa-promoter binding protein-like (SPL)	Y
miR158	2	N	Pentatricopeptide repeat (PPR)	N
miR159/miR319	6	Y	MYB transcription factor	Y
			TCP transcription factor	Y
miR160	3	Y	Auxin response factor (ARF)	Y
miR161	1	N	Pentatricopeptide repeat (PPR)	Y
miR162	2	Y	Dicer-like (DCL)	Y
miR163	1	N	S-adenosylmethionine-dependent methyltransferase (SAMT)	Y
miR164	3	Y	NAC domain transcription factor	Y
miR166/miR165	9	Y	HD-ZIPIII transcription factor	Y
miR167	4	Y	Auxin response factor (ARF)	Y
miR168	2	Y	Argonaute (AGO)	Y
miR169	14	Y	HAP2 transcription factor	Y
miR171/miR170	4	Y	Scarecrow-like (SCL)	Y
miR172	5	Y	Apetala2-like transcription factor (AP2)	Y
miR173	1	N	TAS1, TAS2	Y
miR390/miR391	3	Y	TAS3	Y
miR393	2	Y	Transport inhibitor response 1 (TIR1)/Auxin F-box (AFB)	Y
			bHLH transcription factor	Y
miR394	2	Y	F-box	Y
miR395	6	Y	ATP-sulfurylase (APS)	Y
			Sulfate transporter (AST)	Y
miR396	2	Y	Growth regulating factor (GRF)	Y
miR397	2	Y	Laccase (LAC)	Y
miR398	3	Y	Copper superoxide dismutase (CSD)	Y
			Cytochrome-c oxidase	Y
miR399	6	Y	E2 ubiquiting-conjugating protein (E2-UBC)	Y
miR400	1	N	Pentatricopeptide repeat (PPR)	N
miR402	1	N	HhH-GPD base excision DNA repair	N
miR403	1	Y	Argonaute (AGO)	Y
miR408	1	Y	Laccase (LAC)	Y
			Plantacyanin-like (PCL)	N
miR447	3	N	2-phosphoglycerate kinase-related (2-PGK)	Y
miR472	1	Y	CC-NBS-LRR	Y
miR773	1	N	DNA (cytosine-5-)-methyltransferase	Y
miR774	1	N	F-box	Y
miR775	1	N	Galactosyltransferase	Y
miR778	1	N	SET-domain	Y
miR780.1	1	N	Cation/hydrogen exchanger	Y
miR780.2	1			
miR824	1	N	MADS-box transcription factor	Y
miR827	1	N	SPX (SYG1/Pho81/XPR1) domain/Zinc finger (C3HC4-type)	Y
miR842	1	N	Jacalin lectin	Y
miR844	1	N	Kinase	Y
miR846	1	N	Jacalin lectin	Y
miR856	1	N	Cation/hydrogen exchanger	Y
miR857	1	N	Laccase	Y
miR858	1	N	MYB transcription factor	Y
miR859	1	N	F-box	Y

***MIRNA*s with only predicted targets, or no predicted targets**

***MIRNA* family**[a]	Loci	Conserved[a]	Target family	Predicted targets
miR771	1	N	-	-
miR776	1	N	Serine/threonine kinase	At5g62310
miR777	1	N	-	-
miR779	1	N	-	-
miR781	1	N	CD2-binding, MCM	At5g23480, At1g44900
miR823	1	N	Chromomethylase	At1g69770
miR825	1	N	Remorin, zinc finger homeobox family, frataxin-related	At2g41870, At5g65410, At4g03240
miR829.2	1	N	-	-
miR830-5p	1	N	RanBP1 domain, kinesin motor-related	At1g52380, At3g45850
miR830-3p	1		-	-
miR833-3p	1	N	-	-
miR833-5p	1		F-box	At1g77650
miR840	1	N	WHIRLY transcription factor	At2g02740
miR843	1	N	F-box, 1-aminocyclopropane-1-carboxylate synthase	At3g13830, At1g11810, At2g22810
miR845a	2	N	-	-
miR845b				
miR851-5p	1	N	-	-
miR852	1	N	ATPase	At5g62670
miR853	1	N	-	-
miR860	1	N	Histone deacetylase, ferrochelatase, RNA recognition motif	At5g26040, At5g26030, At3g12640
miR861-3p	1	N	-	-
miR861-5p	1		-	-
miR862-5p	1	N	-	-
miR862-3p	1		-	-
miR863-3p	1	N	-	-
miR863-5p	1		Kinase, Legumain (C13 protease)	At2g26700, At1g62710
miR864-5p	1	N	Triacylglycerol lipase	At1g06250
miR864-3p	1		Expressed protein	At4g25210
miR865-5p	1	N	Serine carboxypeptidase, sulfate transporter	At5g42240, At3g51895
miR865-3p	1		DEAD box RNA helicase, DNA-binding bromodomain-containing protein	At2g07750, At2g34900, At1g03770
miR866-3p	1	N	Kinase, electron transport SCO1/SenC, NAD-dependent G-3-P dehydrogenase	At4g21400, At4g39740, At2g41540
miR866-5p	1		Expressed protein, C2 domain-containing protein	At4g21700, At1g66360
miR867	1	N	PHD finger-related/SET domain, kinase, phospholipase/carboxylesterase	At4g27910, At3g17750, At3g15650
miR868	1	N	-	-
miR869.1	1	N	-	-
miR869.2	1	N	-	-
miR870	1	N	-	-
miR871	1	N	-	-

[a]Conserved between *Arabidopsis* and *Populus*.

[b]Dashes indicate no targets validated or high-quality targets predicted.