Molecular genetic analysis of siRNA biogenesis and function in *Arabidopsis thaliana*.

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Todd Lucas Blevins

aus Oregon in den Vereinigten Staaten

Friedrich-Miescher-Institut

Basel, September 2009

creative commons

Attribution-Noncommercial-No Derivative Works 2.5 Switzerland

---

You are free:

to Share — to copy, distribute and transmit the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.

Your fair dealing and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license) available in German:
http://creativecommons.org/licenses/by-nc-nd/2.5/ch/legalcode.de

Disclaimer:
The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license. Creative Commons is not a law firm and does not provide legal services. Distributing of, displaying of, or linking to this Commons Deed does not create an attorney-client relationship.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/deed.en Datum: 3.4.2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von Prof. Dr. Frederick Meins, Jr. und Dr. Michael Florian Mette.

Basel, den 18. September 2007

Prof. Dr. Hans-Peter Hauri

Dekan

Table of contents

Blevins, T., Pontes, O., Pikaard, C.S., Meins, F., Jr. (2009). Heterochromatic siRNAs and DDM1 Independently Silence Aberrant 5S rDNA Transcripts in Arabidopsis. PLoS ONE **4,** e5932.

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.

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.

1. Summary

In diverse eukaryotes, small RNA products of Dicer-like (DCL) proteins regulate mRNA stability or translation, and direct chromatin modifications to genomic regions, phenomena collectively known as RNA silencing. In plants, different types of small RNAs generated from double-stranded RNA, called short interfering RNAs (siRNAs), mediate RNA-directed DNA methylation (RdDM) to endogenous repeats and defense against viruses. Here, I studied the biogenesis and function of siRNAs from two types of tandem repeat: the 180 bp and 5S rDNA arrays of *Arabidopsis thaliana*. Furthermore, I analyzed siRNAs derived from two DNA viruses -- *Cabbage Leaf Curl Virus* (CaLCuV) and *Cauliflower Mosaic Virus* (CaMV) -- in *Arabidopsis*. Using a reverse genetics approach, I found that accumulation of specific size classes of ~20-24 nt siRNAs depends on particular *Arabidopsis* DCLs and also HEN1, which methylates siRNA 3'-ends. Upstream of these steps, biogenesis of tandem repeat-derived siRNAs required the RNA polymerase IV (Pol IV) pathway, known to include NRPD1a and RDR2. 5S rDNA methylation was reduced in mutants deficient for NRPD1a, RDR2 or the chromatin remodeling factor DDM1, but not in strains deficient for four *Arabidopsis* DCLs. Genetic crosses showed that DDM1 and the Pol IV pathway play distinct but intertwined roles in 5S rDNA methylation: analysis of a strain deficient for both DDM1 and DCL3 suggests that siRNA function in cytosine methylation is most crucial when CpG-methylation maintenance is impaired. Integrating my data with reports on 5S rRNA gene regulation, I propose a dynamic model to describe RdDM affecting the 5S rDNA arrays. A similar mechanism might regulate genes that impact growth in *Arabidopsis*, since double mutants deficient for DDM1 and the Pol IV pathway showed reduced fresh weight. Unlike endogenous siRNAs, viral siRNA accumulation did not require upstream components of known RNA silencing pathways, such as RDR2 or RDR6. However, RDR-independent mechanisms leading to viral siRNA biogenesis may include: (i) overlapping sense / antisense transcription about the circular viral DNA, and (ii) folding of the CaMV 35S RNA transcript leader. Supporting the latter hypothesis, CaMV-infected plants accumulate siRNA from three "hotspots", all matching the structured leader. Furthermore, DCL1 -- thought to excise microRNAs from stem-loop hairpin precursors -- appears to mediate production of ~21 nt siRNAs from the leader or related dsRNA. In general, my results are consistent with a "branched" pathway model for siRNA biogenesis, in which substrates from various sources can be funneled through the same DCL and HEN1 steps, before the siRNA products are incorporated into specialized effectors.

2. List of abbreviations

A. thaliana	*Arabidopsis thaliana*, or the model organism *Arabidopsis*
AGO	Argonaute-like protein
bp	Base pair
BSA	Bovine serum albumin
CaLCuV	*Cabbage Leaf Curl Virus*
CaMV	*Cauliflower Mosaic Virus*
cDNA	Complementary DNA
Col-0	Columbia-0 ecotype of *A. thaliana*
DCL	Dicer-like protein
d23	*dcl2 dcl3*
d24	*dcl2 dcl4*
d34	*dcl3 dcl4*
d234	*dcl2 dcl3 dcl4*
$s1^{qm}$	*dcl1-8 dcl2 dcl3 dcl4* (*sin1-2* based quadruple *dcl*-mutant)
$c1^{qm}$	*dcl1-9 dcl2 dcl3 dcl4* (*caf* based quadruple *dcl*-mutant)
DEPC	Diethylpyrocarbonate
dpi	Days post-inoculation
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
EDTA	Ethylenediamine tetraacetic acid
EMS	Ethyl methanesulfonate-induced mutation
GABI	“Genomanalyse im biologischen System Pflanze” T-DNA insertion
HMW	High molecular weight
LMW	Low molecular weight
miRNA	microRNA
nt	Nucleotide
OD_{260}	Optical density measured at the 260 nm wavelength
PCR	Polymerase Chain Reaction
PTGS	Post-transcriptional Gene Silencing
qPCR	Quantitative RT-PCR
RdDM	RNA-directed DNA Methylation
RDR	Putative RNA-dependent RNA polymerase in *Arabidopsis*
RdRP	RNA-dependent RNA polymerase activity
RISC	RNA-induced silencing complex

RITS	RNA-induced Transcriptional Silencing (complex)
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
SAIL	Syngenta *Arabidopsis* Insertion Library T-DNA insertion
SALK	Salk Institute T-DNA insertion
siRNA	Short interfering RNA
smRNA	Small RNA (includes siRNAs and miRNAs)
ssRNA	Single-stranded RNA
T-DNA	*Agrobacterium tumefaciens* mediated T-DNA insertion mutant
ta-siRNA	*trans*-acting siRNA
TAIR ID#	The *Arabidopsis* Information Resource identification number
TGS	Transcriptional Gene Silencing
Tris	Tris(hydroxymethyl)-amino-methane
Tris-HCl	Tris(hydroxymethyl)-amino-methane hydrochloric acid
WT	Wild type (Col-0, unless otherwise stated)
VIGS	Virus-induced Gene Silencing

3. Introduction

Small RNAs (smRNAs) specify targets of a wide range of RNA silencing processes in eukaryotes (Parker and Barford, 2006). RNA silencing refers to phenomena in which double-stranded RNAs (dsRNAs) or stem-loop hairpin RNAs are processed by a family of RNase III proteins into ~20-25 nt smRNAs, which guide different effector complexes to cleave specific RNA transcripts, block productive mRNA translation or direct chromatin modifications to specific DNA regions (Matzke et al., 2001; Baulcombe, 2004; Meins et al., 2005). For example, smRNA-directed heterochromatin formation around repetitive DNA is thought to maintain genome integrity in fission yeast, plants and animals (O'Donnell K and Boeke, 2007; Zaratiegui et al., 2007). In plants and mammals, this can involve cytosine methylation of DNA repeats (Bender, 2004; Matzke et al., 2005; Aravin et al., 2007). RNA silencing also mediates a general virus defense in plants and invertebrates (Lecellier and Voinnet, 2004; Wang et al., 2006), and regulates expression of developmentally important genes in plants and animals (Bartel, 2004; Meins et al., 2005; Poethig et al., 2006). Because it targets specific nucleic acid sequences, RNA silencing has important agricultural, clinical and research applications (Grunweller and Hartmann, 2005; Watson et al., 2005).

Our current understanding of RNA silencing emerged from molecular genetic and biochemical experiments in different eukaryotic systems that pointed to a evolutionarily conserved, core mechanism (Tijsterman et al., 2002; Voinnet, 2003). In contrast to most animal systems, plants possess a particularly wide array of silencing modules, which can be grouped into functionally distinct smRNA biogenesis pathways (Herr and Baulcombe, 2004; Meins et al., 2005). This introduction briefly charts discoveries that laid a foundation for the RNA silencing paradigm, and then describes protein families whose members are required for steps in the core mechanism. Three endogenous RNA silencing pathways of *Arabidopsis* are then described. Particular attention is given to the role of repeat-associated smRNAs in genomic DNA methylation, which is the first topic of my dissertation. Finally, the relevance of RNA silencing pathways to virus-derived smRNA biogenesis is covered, which is the second focus of my work.

3.1. Early RNA silencing discoveries

RNA silencing was discovered by experiments with transgenes in plants. Introducing multiple transgenes driven by the same promoter into a plant line frequently triggered silencing of the transgenes, which was correlated with methylation of their promoters (Matzke et al., 1989; Park et al., 1996). This effect, which was linked to reduced transcription of the transgenes , was later called transcriptional gene silencing (TGS). Although DNA-DNA interactions would suffice to explain homology-dependent DNA methylation, experiments with viroids in tobacco indicated that RNA could target cytosine methylation to homologous DNA sequences in the plant nucleus (Wassenegger et al., 1994).

In independent experiments, Napoli *et al.* (1990) and van der Krol *et al.* (1990) introduced a chalcone synthase (*CHS*) transgene into petunia plants by *Agrobacterium*-mediated transformation. CHS is required for the biosynthesis of anthocyanin pigments, so they expected that its overexpression would result in deeply pigmented flowers. Instead the plants developed purple-white variegated and white flowers. This phenomenon, called co-suppression, was due to inhibition *in trans* of the accumulation of transgene and endogenous gene encoded *CHS* mRNAs. Co-suppression of *CHS* in petunia and related phenomena in plants often resulted from increased degradation of mRNA, or post-transcriptional gene silencing (PTGS, de Carvalho et al., 1992; Ingelbrecht et al., 1994; van Blokland et al., 1994). PTGS was found to reset post-meiotically, meaning that plants showing a silent state produced embryos that regained expression of the transgene (Kunz et al., 2001). In contrast, TGS and the associated DNA methylation was heritable through multiple self-fertilized generations (Kilby et al., 1992; Park et al., 1996). In this dissertation, I use the term RNA silencing to refer to both PTGS and TGS processes for which dsRNA is essential (Meister and Tuschl, 2004).

Further silencing phenomena were uncovered by inserting specific sequences into vectors that transiently express RNA in plants. For example, RNA silencing was shown to limit replication of recombinant RNA viruses with homology to a plant transgene by enhanced degradation of the viral RNA (Lindbo and Dougherty, 1992; Goodwin et al., 1996). Then, the similarity between RNA silencing and natural defense against viruses was noted (Covey et al., 1997; Ratcliff et al., 1997). Importantly, plant viruses were found to encode proteins that suppress PTGS of transgenes (Anandalakshmi et al., 1998). These silencing suppressors were found to be determinates of viral pathogenicity, which suggested that viruses had evolved a counter-defense to evade RNA silencing (Brigneti et al., 1998; Kasschau and Carrington, 1998).

Conversely, viruses carrying portions of an expressed reporter transgene could induce silencing of that gene, a technology known as virus-induced gene silencing (VIGS, Ruiz et al., 1998). VIGS, and another transient system based on *Agrobacterium* infiltration, allowed rapid assessment of the temporal and spatial dynamics of silencing (Schöb et al., 1997; Voinnet et al., 1998; Jones et al., 1999). Local induction of silencing and grafting experiments showed that silent cells elaborate a mobile silencing signal that can move locally from cell-to-cell and systemically through the phloem (Palauqui et al., 1997; Voinnet and Baulcombe, 1997; Crete et al., 2001).

Using four independent plant RNA silencing systems, Hamilton and Baulcombe (1999) showed that both PTGS and VIGS correlated with the accumulation of ~25 nt smRNAs complementary (antisense) to the targeted mRNAs. This discovery was important on multiple levels: (i) the smRNAs were large enough to convey sequence specificity; (ii) they were, in theory, small enough to facilitate their systemic spread in plants; and (iii) they were reminiscent of 22 nt small temporal RNAs expressed in *Caenorhabditis elegans* that were hypothesized to block translation of an endogenous mRNA by base-pairing to its 3'-UTR (Lee et al., 1993; Wightman et al., 1993). Thereafter, RNA isolated from uninfected, wild-type tobacco and *Arabidopsis* was shown to contain a similar ~21-23 nt smRNA pool; this suggested the existence of smRNAs derived from endogenous genes, which were down regulated by PTGS (Hutvagner et al., 2000).

In the late 1990's, evidence came from both plants and animals for the role of long dsRNAs in silencing. Fire *et al.* (1998) showed that introducing pure dsRNA molecules into *C.elegans* induced sequence-specific silencing of genes homologous to that dsRNA – a process the authors called RNA interference (RNAi). Furthermore, plant transgenes containing a transcribed inverted repeat homologous to a promoter region induced cytosine methylation and TGS of the corresponding antibiotic resistance gene (Mette et al., 1999). This induction depended on the generation of an inverted repeat via recombination *in planta*, which correlated with the biogenesis of smRNAs derived from the inverted repeat (Mette et al., 2000). Equally, transcribed inverted repeats directed against an mRNA or direct bombardment of plants with dsRNA homologous to an mRNA, induced efficient PTGS of the corresponding target (Wesley et al., 2001; Klahre et al., 2002). These and other experiments suggested that PTGS and TGS could be mechanistically linked in some systems (Sijen et al., 2001; Fojtova et al., 2003).

To elucidate the mechanism of RNA silencing, forward genetic screens for mutants deficient in transgene silencing were carried out in *Neurospora crassa*, *A. thaliana* and *C. elegans* (Cogoni and Macino, 1999; Tabara et al., 1999; Dalmay et al., 2000; Mourrain et al., 2000). Concurrent biochemical and genetic work from animal systems found enzymatic

activities and gene products that mediate key steps of RNAi (Zamore et al., 2000; Ketting et al., 2001; Williams and Rubin, 2002). Among the genes were those encoding proteins conserved in animals, plants, fungi and protists, showing that RNA silencing is an evolutionary ancient mechanism with functions in many organisms (Fagard et al., 2000; Cerutti and Casas-Mollano, 2006).

Throughout this dissertation, homozygous recessive mutants for an *Arabidopsis* gene (e.g., *DCL1*) are designated by a lowercase italic gene symbol (e.g., *dcl1*). However, the notation (-/-) is used to simplify diagrams in which homozygous mutant plants are compared to heterozygous and wild-type plants -- e.g., *DCL1*(-/-), *DCL1*(+/-) and *DCL1*(+/+).

3.2 Classes of proteins required for RNA silencing

3.2.1. RNA-dependent RNA polymerases

RNA-dependent RNA polymerases (RdRPs) in plants, *N. crassa*, *C. elegans*, *Schizosaccharomyces pombe* (fission yeast) and other eukaryotes are thought to produce dsRNA silencing intermediates from single-stranded RNAs (ssRNAs) or amplify RNA silencing signals (Smardon et al., 2000; Xie et al., 2004; Wassenegger and Krczal, 2006). The first endogenous RdRP activity of a eukaryote was discovered in tomato (Schiebel et al., 1993; Schiebel et al., 1998), which helped explain why a homologous protein, QDE-1, is required for RNA silencing in *N. crassa* (Cogoni and Macino, 1999). Six genes identified in *Arabidopsis* (*RDR1 to RDR6*) encode proteins similar to the tomato RdRP (Wassenegger and Krczal, 2006). At least two of these genes, *RDR2* and *RDR6*, are required for endogenous RNA silencing pathways (Dalmay et al., 2000; Mourrain et al., 2000; Vazquez et al., 2004b; Xie et al., 2004). Similar findings were made in other Embryophyta -- i.e., in land plants (Alleman et al., 2006; Talmor-Neiman et al., 2006). Biochemical work established that *N. crassa* QDE-1 and *S. pombe* Rdp1 possess RdRP activities *in vitro* (Makeyev and Bamford, 2002; Sugiyama et al., 2005), but similar confirmation is lacking for the *Arabidopsis* proteins. *C. elegans* is the only animal known to have a putative RdRP gene (Smardon et al., 2000; Grishok et al., 2005; Lee et al., 2006).

3.2.2. Dicer RNase III endoribonucleases

Dicers, which are RNase III type endoribonucleases, cleave dsRNA substrates into smRNA duplexes in animals (Bernstein et al., 2001). Dicer products were studied in great

detail using *Drosophila* cell extracts (Elbashir et al., 2001a). They are ~22 nt long and have the following chemical and structural features: single phosphate groups on their 5' ends, two to three nt 3' overhangs, and free hydroxyl groups on their 3' ends (**Figure 1**). Several protein domains were identified by analyzing animal Dicers, including an RNA helicase domain, a DUF283 (domain of unknown function), a PAZ domain, two RNase III catalytic domains, and a dsRNA binding domain (Bernstein et al., 2001; Meister and Tuschl, 2004). All organisms that are competent for RNA silencing have at least one Dicer-like (DCL) protein (Susi et al., 2004; Sontheimer, 2005).

Using the evolutionarily conserved domains as criteria, four *Arabidopsis* genes were found to encode DCLs (Golden et al., 2002; Vazquez, 2006). DCL products processed from perfect dsRNA are called short interfering RNAs (siRNAs), which are the sequence-specifying molecules of PTGS / RNAi in eukaryotes (Fire et al., 1998; Hamilton and Baulcombe, 1999; Zamore et al., 2000; Elbashir et al., 2001b; Tang and Zamore, 2004). This distinguishes them from DCL products called microRNAs (miRNAs), which derive from endogenous hairpin RNA precursors with stem sections containing mismatches (Jones-Rhoades et al., 2006). Small temporal RNAs of *C. elegans* were the first example of miRNAs, which have since been identified in various plants and animals (Ambros et al., 2003; Jones-Rhoades et al., 2006). Different *Arabidopsis* DCL proteins have been genetically linked to the biogenesis of endogenous siRNAs and miRNAs (Reinhart et al., 2002; Schauer et al., 2002; Xie et al., 2004; Gasciolli et al., 2005; Bouche et al., 2006; Henderson et al., 2006).

3.2.3. Argonaute proteins

Argonaute (AGO) proteins were named after the *ago1* mutants of *Arabidopsis* (Bohmert et al., 1998). They have characteristic N-terminal PAZ domains and C-terminal PIWI domains (Carmell et al., 2002; Hunter et al., 2003). AGO proteins are key components of RNA-induced Silencing Complexes (RISCs), which directly mediate RNA silencing effects (**Figure 1**). During siRNA-programmed RNA degradation, RISC contains an AGO protein bound to an siRNA strand, which guides cleavage of complementary mRNA targets (Martinez et al., 2002; Baumberger and Baulcombe, 2005). Additionally, RISC-like complexes are thought to direct histone modifications and/or DNA methylation to specific DNA repeats (Motamedi et al., 2004; Pikaard, 2006). In fission yeast this is called the RNA-induced Transcriptional Silencing (RITS) complex (Verdel et al., 2004).

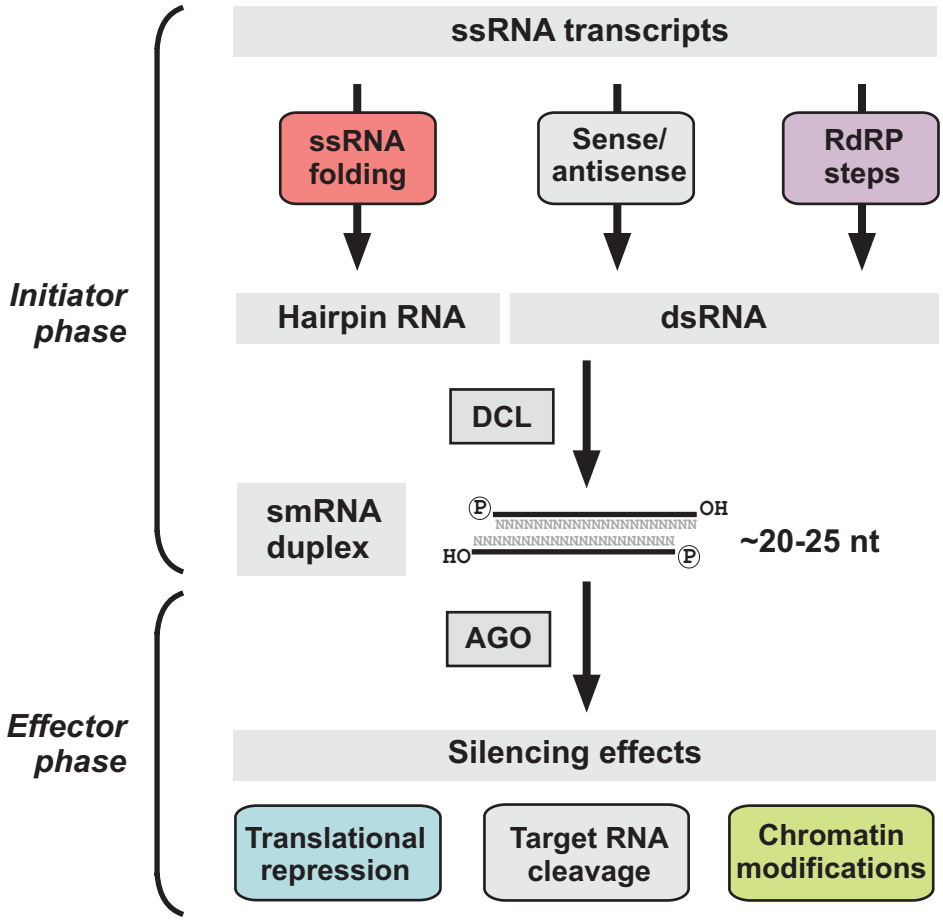


Figure 1. The core RNA silencing mechanism. The core mechanism of RNA silencing can be divided into two phases: During the ***initiator phase***, partially or perfectly dsRNA forms by one of three processes: a ssRNA folds into a hairpin, sense and antisense transcripts anneal to form dsRNA, or a ssRNA is converted into dsRNA by an RNA-dependent RNA polymerase (RdRP). Dicer-like (DCL) proteins process these molecules into ~20-25 nt duplex products. In the ***effector phase*** specific strands of these duplex smRNAs are integrated into silencing effector complexes, including an Argonaute protein (AGO), to guide translational repression of mRNAs, cleavage of RNA transcripts, or covalent histone modifications (and/or DNA methylation) to specific genomic regions. The most evolutionarily conserved factors and steps are depicted in gray.

Based on sequence similarity analysis, the *Arabidopsis* genome contains 10 genes that encode AGO proteins (Fagard et al., 2000; Vazquez, 2006). Five of these genes, *AGO1*, *AGO4*, *AGO6*, *AGO7* and *AGO10*, have been linked to specific RNA silencing processes (Morel et al., 2002; Hunter et al., 2003; Zilberman et al., 2003; Baumberger and Baulcombe, 2005; Kidner and Martienssen, 2005; Zheng et al., 2007).

3.3. The core RNA silencing mechanism

The biochemical steps of RNA silencing can be divided into two phases: the *initiator phase* resulting in smRNA biogenesis and the *effector phase*, which achieves sequence-specific silencing effects (**Figure 1**). During the initiator phase, partially or perfectly dsRNA forms by one of three processes: (i) a ssRNA precursor folds into a hairpin, (ii) complementary sense/antisense transcripts anneal or (iii) an RNA-dependent RNA polymerase (RdRP) uses a ssRNA template to synthesize a second, comlementary RNA strand. The resulting hairpin RNA or dsRNA molecules are then processed by a DCL protein into duplexes of ~20-25 nt smRNAs. These short duplexes are subsequently bound by an AGO protein, which mediates effector steps via RISCs (Collins and Cheng, 2005; Parker and Barford, 2006).

Based on their detailed phylogenetic analysis, Cerutti and Casas-Mollano (2006) propose that a basic set of RNA silencing enzymes (RdRPs, DCLs and AGOs) arose after the emergence of single-celled eukaryotes and before the advent of multicellular organisms. According to their model, the vertebrate lineage kept RNA silencing (RNAi) but lost the archetypical RdRP that is conserved in plants, fungi, *C. elegans* and some protists (Cerutti and Casas-Mollano, 2006). Embryophyta are extraordinary because they show *all* major silencing effects, including RNA-directed DNA methylation (RdDM) and virus defense (Matzke et al., 2005; Brodersen and Voinnet, 2006). This makes the genetic model plant *Arabidopsis* an excellent organism in which to analyze different RNA silencing systems and their interactions.

3.4. Endogenous RNA silencing pathways in *Arabidopsis*

Studies of silencing, its suppression by viruses and the pattern of smRNA accumulation in deficiency mutants have identified several different routes for endogenous smRNA biogenesis in *Arabidopsis* (**Figure 2**) (Vance and Vaucheret, 2001; Herr and Baulcombe, 2004; Meins et al., 2005; Vazquez, 2006). This work showed that subpopulations of the ~20-25 nt smRNAs cloned from *Arabidopsis* and other plants associate with functionally

distinct, conserved RNA silencing pathways (Xie et al., 2004; Sunkar et al., 2005; Talmor-Neiman et al., 2006).

3.4.1. The miRNA pathway

Arabidopsis miRNAs are excised from imperfectly paired regions in the stem of hairpin RNA precursors and are predominantly 20-22 nt long (Park et al., 2002; Reinhart et al., 2002; Ambros et al., 2003). They are thought to guide AGO1 to longer RNA transcripts, which contain a reverse complement binding site (cognate RNAs), and direct AGO1-mediated cleavage to the center of the pairing region (**Figure 2A**, Llave et al., 2002; Vaucheret et al., 2004; Baumberger and Baulcombe, 2005). DCL1 and the dsRNA binding protein Hyponastic Leaves 1 (HYL1) mediate miRNA biogenesis (Park et al., 2002; Reinhart et al., 2002; Han et al., 2004; Kurihara and Watanabe, 2004; Vazquez et al., 2004a; Kurihara et al., 2006). DCL1 is thought to process the miRNA precursor by two endonucleolytic cuts to generate a miRNA/miRNA* duplex, where miRNA* is the duplex strand not used to guide cognate RNA cleavage. Similarly to what was shown in animals, the duplex is loaded on to AGO1, which stabilizes the miRNA, while the miRNA* strand is thought to be degraded (Vaucheret et al., 2004; Baumberger and Baulcombe, 2005; Matranga et al., 2005). The 3'-ends of miRNAs appear to be stabilized by methylation of their 2'-hydroxyl groups, which is mediated by the protein HEN1 (Park et al., 2002; Boutet et al., 2003; Yu et al., 2005).

miRNA precursor genes, called *MIRNA* genes, can be annotated using the predicted folding structure of their transcripts as one criterion (Jones-Rhoades and Bartel, 2004). Diverse *MIRNA* genes are known in *Arabidopsis*, with experimental evidence for at least 64 gene families, whose miRNA products target dozens of developmentally important mRNAs (Reinhart et al., 2002; Allen et al., 2004; Axtell and Bartel, 2005; Mallory and Vaucheret, 2006). Thus, *dcl1* null mutants appear to be lethal, and even hypomorphic mutants have strong, pleiotropic phenotypes (Schauer et al., 2002). Although deep evolutionary conservation was originally thought to be a characteristic of all *MIRNA*s (Ambros et al., 2003; Axtell and Bartel, 2005), recent work has demonstrated that plants from different taxa may express specific pools of miRNA that specify developmental programs limited to those species (Rajagopalan et al., 2006; Kutter et al., 2007).

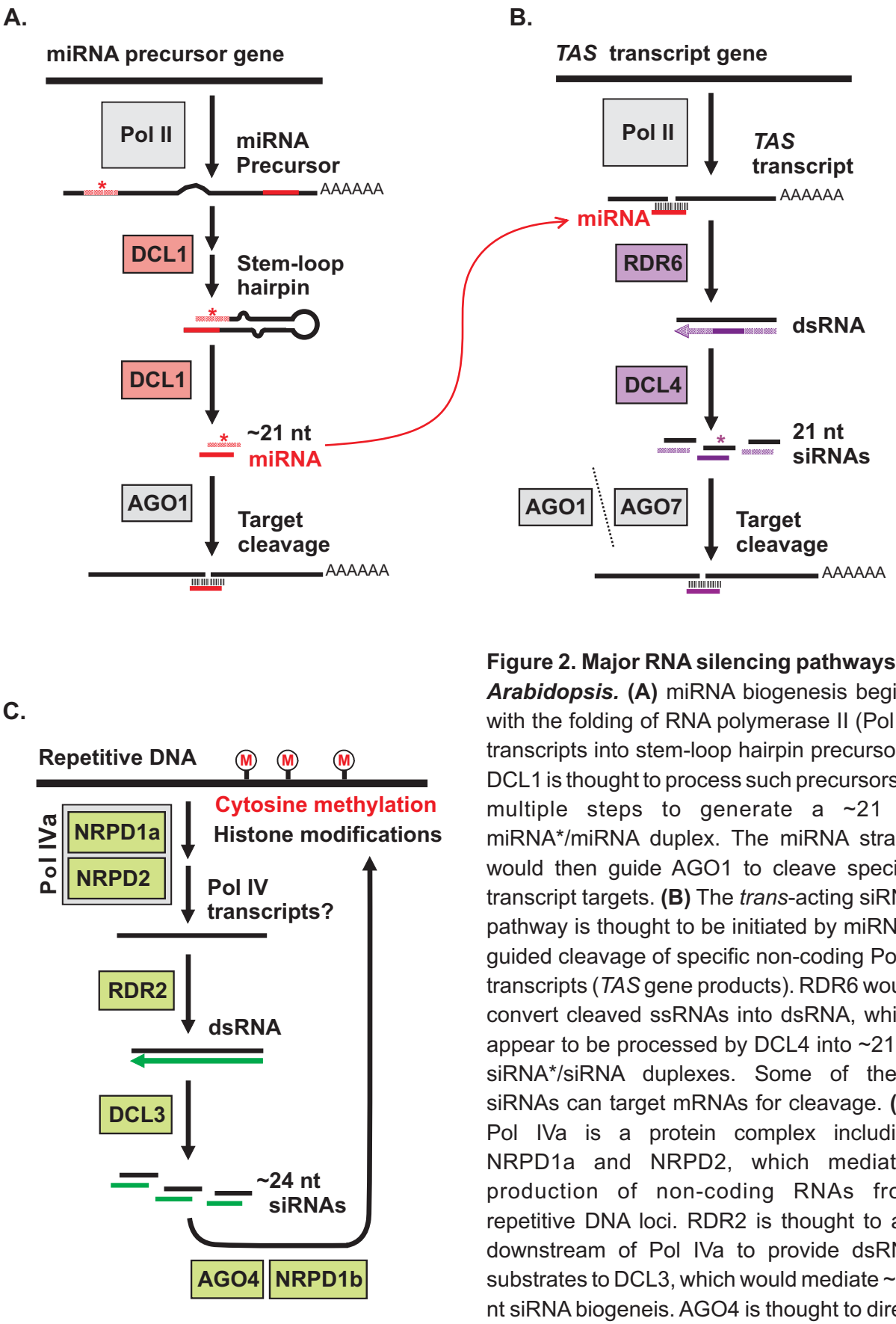


Figure 2. Major RNA silencing pathways in *Arabidopsis.* (A) miRNA biogenesis begins with the folding of RNA polymerase II (Pol II) transcripts into stem-loop hairpin precursors. DCL1 is thought to process such precursors in multiple steps to generate a ~21 nt miRNA*/miRNA duplex. The miRNA strand would then guide AGO1 to cleave specific transcript targets. **(B)** The *trans*-acting siRNA pathway is thought to be initiated by miRNA-guided cleavage of specific non-coding Pol II transcripts (*TAS* gene products). RDR6 would convert cleaved ssRNAs into dsRNA, which appear to be processed by DCL4 into ~21 nt siRNA*/siRNA duplexes. Some of these siRNAs can target mRNAs for cleavage. **(C)** Pol IVa is a protein complex including NRPD1a and NRPD2, which mediates production of non-coding RNAs from repetitive DNA loci. RDR2 is thought to act downstream of Pol IVa to provide dsRNA substrates to DCL3, which would mediate ~24 nt siRNA biogeneis. AGO4 is thought to direct DNA methylation and repressive chromatin modifications to repeats using siRNAs as guides in an effector complex with NRPD1b.

3.4.2. The *trans*-acting siRNA pathway

Most endogenous siRNAs in *Arabidopsis* are cleaved from long perfect dsRNAs. The biogenesis of these dsRNAs requires the activity of either of two putative RdRPs: RDR6 and RDR2 (**Figure 2B/C**). Multiple lines of evidence suggest that RDR6 and RDR2 act upstream of siRNA biogenesis in their respective pathways (Vazquez et al., 2004b; Allen et al., 2005; Yoshikawa et al., 2005; Pontes et al., 2006)

An important example is the evolutionarily conserved, non-coding transcript encoded by *TAS3* in *Arabidopsis* (Allen et al., 2005; Adenot et al., 2006; Fahlgren et al., 2006). Remarkably, homologous *TAS* transcripts influence development in various Embryophyta via the same RNA silencing mechanism (Axtell et al., 2006; Hunter et al., 2006; Talmor-Neiman et al., 2006; Nogueira et al., 2007). Biogenesis of siRNAs from *TAS3* and similar genes apparently depends on miRNA-directed cleavage of primary *TAS* transcripts (**Figure 2B**, Allen et al., 2005; Axtell et al., 2006). In *Arabidopsis*, the cleavage event is thought to generate substrates for RDR6, whose dsRNA products are processed by DCL4 (Gasciolli et al., 2005; Xie et al., 2005; Yoshikawa et al., 2005). Certain ~21 nt siRNA products of DCL4, defined by their "phase" within *TAS* transcripts relative to the miRNA cleavage site, are then thought to associate with AGO1 or AGO7 and target specific mRNAs for cleavage *in trans* (**Figure 2B**, Peragine et al., 2004; Vazquez et al., 2004b; Allen et al., 2005; Baumberger and Baulcombe, 2005; Adenot et al., 2006).

In RDR6-deficient mutants, these *trans*-acting siRNAs (ta-siRNAs) fail to accumulate, while the *TAS* transcripts themselves overaccumulate (Vazquez et al., 2004b). These data are consistent with RDR6 acting as an RdRP on *TAS* ssRNA templates. Supporting this model, mRNAs targeted by specific ta-siRNAs in the wild type, overaccumulate in deficiency mutants *dcl1*, *rdr6*, *dcl4* and *ago1* (Peragine et al., 2004; Vazquez et al., 2004b; Xie et al., 2005).

3.4.3. RNA polymerase IV (Pol IV) and repeat-associated siRNAs

A recently discovered fourth class of *Arabidopsis* RNA polymerase, with catalytic subunits similar to those of RNA polymerase II (*Arabidopsis*_Genome_Initiative, 2000), has been shown to have functions in RNA silencing (Herr et al., 2005; Onodera et al., 2005). Two related RNA polymerase IV (Pol IV) complexes have been identified: Pol IVa containing the subunits NRPD1a and NRPD2, and Pol IVb containing the subunits NRPD1b and NRPD2 (Kanno et al., 2005b; Pontier et al., 2005; Pontes et al., 2006). Together, these Pol IV complexes are required for the biogenesis of most endogenous siRNAs in

Arabidopsis and are conserved in Embryophyta (Alleman et al., 2006; Kasschau et al., 2007; Luo and Hall, 2007; Zhang et al., 2007).

Pol IV-dependent siRNAs frequently derive from DNA repeats. For example, 24 nt siRNAs derived from a short-interspersed element (SINE), *AtSN1*, accumulate in wild-type plants but not in the *NRPD1a* deficient mutant *sde4* (Hamilton et al., 2002; Herr et al., 2005). The same effect was observed in *RDR2* deficient plants (Xie et al., 2004). Importantly, the authors showed that the *AtSN1* genomic DNA region is hypomethylated in *sde4* and *rdr2* mutants (Hamilton et al., 2002; Xie et al., 2004). In another example, SINE-like repeats in the promoter of the *FWA* gene are a source of 24 nt siRNAs (Lippman et al., 2004). Deficiencies in NRPD1a or RDR2 lead to a late flowering phenotype that is linked to hypomethylation and ectopic expression of *FWA* (Chan et al., 2004; Chan et al., 2006a; Kinoshita et al., 2007). Similar siRNAs also arise from genes encoding the 5S rRNA of *Arabidopsis* (Xie et al., 2004). In eukaryotes, these 5S rRNA genes are arranged in large tandem repeat arrays (5S rDNA) that are transcribed by RNA polymerase III (Douet and Tourmente, 2007). Distinct 5S rDNA arrays are situated in pericentromeric regions of *Arabidopsis* Chromosomes 3, 4 and 5, for a total of about 1000 repeat copies (Cloix et al., 2000; Vaillant et al., 2007b).

3.4.4. DNA methylation and histone modifications mediated by the Pol IV pathway

Cytosine methylation of DNA is required for normal development in both plants and animals, where it controls genomic imprinting and influences other important biological processes (Bird, 2002; Chan et al., 2005; Goll and Bestor, 2005). DNA methylation is associated with inhibition of promoter activity and may function to protect the genome from transposable elements (Yoder et al., 1997; Hirochika et al., 2000; Weber and Schubeler, 2007). In plants, the methylation status of certain genes regulates their epigenetic state (Takeda and Paszkowski, 2006; Henderson and Jacobsen, 2007): that is, a heritable but potentially reversible state of gene activity not explained by changes in DNA sequence (Meins, 1996; Pikaard, 2000). When “normal” epigenetic states are perturbed by a global reduction in DNA methylation, epialleles can form that are linked to altered DNA methylation and expression of particular genes (Finnegan, 2002; Yi et al., 2004).

Eukaryotic DNA is packaged into chromatin, which serves as a platform for regulating access to the genome and modulating transcription (Wolffe, 1998; Berger, 2007). Covalent histone modifications, such as histone 3 lysine 9 dimethylation (H3K9), are associated with condensed chromatin known as heterochromatin (Tariq and Paszkowski, 2004; Peters and Schubeler, 2005). In *Arabidopsis*, the histone

methyltransferase KRYPTONITE (KYP1), which mediates H3K9 dimethylation to *AtSN1*, has a domain that is postulated to bind methylated DNA (Jackson et al., 2002; Johnson et al., 2007). This suggests that DNA methylation is reinforced by H3K9 dimethylation via the recruitment of KYP1 (Johnson et al., 2007) and is consistent with immunostaining and chromatin-immunoprecipitation observations in mutants deficient for DNA methylation (Soppe et al., 2002; Tariq et al., 2003). Conversely, cytosine methylation in the CpNpG sequence context requires KYP1 and the H3K9 dimethylation mediated by that protein (Jackson et al., 2002).

The Pol IV pathway directs DNA methylation and H3K9 methylation to genomic repeats, such as *AtSN1*, the *FWA* promoter and 5S rRNA genes (Xie et al., 2004; Herr et al., 2005; Onodera et al., 2005; Huettel et al., 2006). For the purposes of my dissertation, however, I focused on DNA methylation mediated by the Pol IV pathway. Methylation of cytosines in non-CpG sequences contexts is at least partly dependent on DRM2, a *de novo* cytosine methyltransferase (Cao et al., 2003; Chan et al., 2006b). A function for DRM2 in Pol IV-mediated methylation is suggested by the reduced accumulation of repeat-associated siRNAs in the *drm1 drm2* double mutant (Onodera et al., 2005). Forward genetic screens have indicated, however, that multiple DNA methyltransferases cooperate to achieve RdDM in non-CpG and CpG contexts (Malagnac et al., 2002; Aufsatz et al., 2004; Matzke et al., 2005). RdDM is also facilitated by the SNF2-like protein DRD1 (Kanno et al., 2004; Kanno et al., 2005a).

In a simplified model, Pol IVa produces or amplifies non-coding transcripts, which are converted to dsRNA by RDR2 and cleaved into 24 nt siRNAs by DCL3 (**Figure 2C**) (Xie et al., 2004; Onodera et al., 2005). These siRNAs would then guide silencing effectors, involving Pol IVb and AGO4 proteins, to homologous sequences in the genome to mediate RdDM (Zilberman et al., 2003; Kanno et al., 2005b; Pontes et al., 2006). Studies of the immunological localization of components of the Pol IV pathway (Pontes et al., 2006) and the cloning of smRNAs associated with AGO1 and AGO4 (Qi et al., 2006; Zhang et al., 2007) support the hypothesis that 24 nt siRNAs arising from 5S rRNA genes and other repeats are loaded on to an AGO4 effector complex.

Nonetheless, no experiment has directly shown that siRNAs guide an effector complex to sequences within genomic DNA. One report suggests that nascent RNA, rather than genomic DNA, is the proximate target of the Pol IVa protein in the nucleus (Pontes et al., 2006). Furthermore, studies of centromeric heterochromatin formation in fission yeast found that RITS targets nascent RNA transcripts from repeats, rather than genomic DNA (Moazed et al., 2006). With the plant viroid system described earlier, 30 nt of complementarity between viroid RNA and nuclear DNA were sufficient to mediate

RdDM limited to the targeted region (Pelissier and Wassenegger, 2000). This ~30 nt threshold is slightly longer than repeat-associated siRNAs but much shorter than DNA looped around a single nucleosome, ~170 bp (Wolffe, 1998). This fact strengthens the proposition, based on genetic experiments, that siRNAs can induce precise methylation of genomic DNA targets (Pikaard, 2006), but would not exclude an intermediary role for non-coding RNA transcripts.

Complicating the above described model, production of repeat-derived siRNAs appears to depend on multiple, partially redundant DCLs in plants (Hamilton et al., 2002; Xie et al., 2004; Gasciolli et al., 2005; Henderson et al., 2006; Margis et al., 2006). This has functional significance for RdDM in *Arabidopsis*: Henderson *et al.* (2006) showed that methylation of *AtSN1* was more strongly reduced in the *dcl2 dcl3 dcl4* triple mutant than in *dcl3* alone. Other DNA repeat classes showed only moderate hypomethylation in multiple *dcl*-mutants, however. Methylation of 5S rDNA repeats was not yet analyzed in multiple *dcl*-mutant backgrounds, although it is reduced in mutants for upstream steps of the Pol IV pathway, such as *nrpd1a* or *rdr2* (Onodera et al., 2005).

3.5. Maintenance of 5S rDNA repeat methylation in *Arabidopsis*

Once DNA methylation is established at a DNA repeat, its maintenance over cell divisions and successive plant generations is not necessarily RNA silencing-dependent (Matzke et al., 2005). Methylated cytosines in the CpG sequence context are maintained during DNA replication by the methyltransferase MET1 (Vongs et al., 1993; Kankel et al., 2003; Saze et al., 2003) acting in concert with the SWI2/SNF2-like chromatin remodeling factor DDM1 (Jeddeloh et al., 1999; Steimer et al., 2000). Maintenance of CpNpG site methylation depends on the methyltransferase CMT3 (Bartee et al., 2001; Lindroth et al., 2001). Mutations in MET1, DDM1 and CMT3 have each been shown to release silencing of specific genes and transposable elements (Jeddeloh et al., 1998; Hirochika et al., 2000; Bartee et al., 2001; Saze et al., 2003).

5S rDNA methylation in CpG, and to some extent in non-CpG, contexts is sensitive to DDM1-deficiency (Vongs et al., 1993; Kakutani et al., 1995; Mathieu et al., 2003). Contradictory models have been proposed for the relationship between DDM1-maintained methylation and the Pol IV pathway. Lippman *et al.* (2003 and 2004) showed that the accumulation of *AtSN1* and tandem repeat-derived siRNAs is reduced in the DDM1-deficient mutant, *ddm1-2* (Vongs et al., 1993). This suggested that DDM1 and/or the heterochromatic, methylated state of tandem DNA repeats is a precondition for siRNA biogenesis. In contrast, Onodera *et al.* (2005) found that 5S rDNA repeat-derived siRNAs

are over-expressed in *ddm1-2*. The later result favors a model in which DDM1 and/or DNA methylation represses production of dsRNA precursors for repeat-associated siRNAs. In both sets of experiments, *met1* mutants had the same effect as *ddm1-2*, indicating that DNA hypomethylation rather than DDM1-deficiency *per se* was at issue (Lippman et al., 2003; Onodera et al., 2005). This example illustrates the puzzling relationship between siRNA accumulation and DNA methylation of genomic tandem repeats in plants.

The function of 5S rDNA methylation is uncertain, since the steady state level of 5S rRNA is unaffected by hypomethylation (Mathieu et al., 2002). However, mutants deficient for DDM1, MET1, CMT3, or AGO4, show release of silencing affecting “minor” 5S rRNA genes (Vaillant et al., 2006; Vaillant et al., 2007a). These reports suggest that methylation at CpG and CpNpG sites, which is mediated by the abovementioned proteins, regulates the transcription of certain 5S rDNA repeat units.

3.6. Plant virus infection

Viruses are nucleoproteins that multiply in living cells and have the ability to cause disease (Agrios, 1997). There are approximately 450 species of plant-pathogenic viruses, many of which reduce yields, increase sensitivity to frost and drought, or cause all out crop failure (Hull, 2004; Soosaar et al., 2005). Symptoms of virus infection include leaf chlorosis, necrosis, mosaic patterns and plant stunting (Hammond-Kosack and Jones, 2000). Severe crop losses are caused by the RNA virus, *Sugarcane mosaic virus* (*Potyviridae*), and the DNA viruses, *Tomato yellow leaf curl virus*, *Africa cassava mosaic virus* (both *Geminiviridae*), and *Banana streak virus* (Caulimoviridae), to name just a few examples (Hull et al., 2000; Moriones and Navas-Castillo, 2000; Legg and Fauquet, 2004). Plant virus infection and defense against viruses are therefore of great economic and social import.

3.7. Plant defense responses to viruses

The defense response of plants to virus infection involves both specific host resistance (R) genes and generalized host defense systems (Voinnet, 2001; Hull, 2004; Kang et al., 2005). Most R genes encode nucleotide binding-leucine rich repeat (NB-LRR) proteins that interact with pathogen-encoded effectors, known as avirulence proteins (Jones and Dangl, 2006). Upon recognition of the avirulence protein, a signal transduction cascade induces various basal defenses, including the accumulation of salicylic acid (SA) and expression of pathogenesis-related proteins; this sometimes culminates in a type of programmed cell death, known as the hypersensitive response (HR, Lam et al., 2001). For example, an NB-LRR protein called Rx confers extreme resistance to Potato virus X (PVX) by recognizing the viral coat protein (Bendahmane et al., 1995; Bendahmane et al., 1999). However, PVX infection does not induce HR. On the other hand, the *N* gene of tobacco encodes a protein that mediates recognition of the helicase domain of the tobacco mosaic virus (TMV) replicase (Whitham et al., 1994; Erickson et al., 1999). The resulting HR limits TMV infection to localized necrotic lesions (Hull, 2004). That said, the most general defense response of plants to virus infection is the processing of viral dsRNA into siRNAs, and targeted cleavage of viral genomic RNA or transcripts by silencing effectors (Lecellier and Voinnet, 2004).

.

3.8. RNA silencing as a defense against viruses in plants

Early in the 1990s, PTGS was found to limit the replication of recombinant RNA viruses containing sequences homologous to transgenes in the host genome (Lindbo et al., 1993; Goodwin et al., 1996). However, viral defense mediated by RNA silencing does not require genomic host DNA homologous to the virus: a protected state can also be conveyed by prior infection with a related RNA virus (Ratcliff et al., 1999). This is the likely mechanism of cross-protection, whereby pre-inoculating a plant with a mild virus strain protects it from subsequent inoculation by a related, severe strain (Valle et al., 1988; Voinnet, 2001; Hohn et al., 2007). Such pre-established silencing states, while advantageous for preventing severe disease, are not absolutely required for a successful defense response. RNA silencing contributes to lessen the impact of viral diseases at early stages of infection, which can forestall systemic infection and lead to asymptomatic new growth, sometimes called ‘recovery’ (Voinnet, 2001; Baulcombe, 2004).

3.8.1 Suppressor proteins

Despite RNA silencing-based defense, plant viruses still establish robust infection in susceptible hosts, in part by suppressing RNA silencing. Infection with different RNA and DNA viruses can suppress PTGS affecting transgenes (Voinnet et al., 1999). This is accomplished by proteins expressed by the virus, which are also determinates of viral pathogenicity (Brigneti et al., 1998; Kasschau and Carrington, 1998). Suppressor proteins are diverse in sequence and structure, but many of them bind short or long dsRNA (Silhavy and Burgyan, 2004; Lakatos et al., 2006; Merai et al., 2006). Furthermore, suppressor proteins inactivate multiple types and steps of RNA silencing (Anandalakshmi et al., 1998; Beclin et al., 1998; Brigneti et al., 1998; Kasschau and Carrington, 1998; Di Serio et al., 2001). An elegant case is the p19 protein of *Carnation Italian ringspot virus* and other tombusviruses, which is a molecular caliper that selectively binds siRNAs based on their duplex length (Vargason et al., 2003). In another example, the 2b protein of *Cucumber Mosaic Virus* (CMV) was shown to block SA-mediated virus resistance, impair the spread of intercellular silencing, and reduce accumulation of multiple size classes of CMV-derived siRNAs (Ji and Ding, 2001; Diaz-Pendon et al., 2007).

3.8.2 Virus infection in RNA silencing deficient plants

Another indication that RNA silencing mediates plant defense against viruses came from the study of mutants deficient for siRNA biogenesis pathways. For example, *rdr6* mutants infected with CMV showed overaccumulation of viral genomic RNA compared to the infected wild type (Mourrain et al., 2000). Interestingly, the tobacco ortholog of RDR6 is required for systemic transgene silencing across tissue grafts. This systemic signal is possibly a component of plant virus defense that protects the shoot apical meristem from infection (Schwach et al., 2005). Work with single mutants for *Arabidopsis DCL* genes showed that *dcl2* was hypersusceptible to infection with *Turnip Crinkle Virus* (TCV, Xie et al., 2004). This apparently results from suppression of DCL4 function by the TCV coat protein, p38, leaving mainly DCL2 to compensate for DCL4 deficiency in infected wild-type plants (Deleris et al., 2006). Similar studies showed that triple and quadruple *dcl*-mutants are hypersusceptible to CMV infection (Bouche et al., 2006; Fusaro et al., 2006; Diaz-Pendon et al., 2007). Hypersusceptibility was also observed in *hen1* and *ago1* mutants infected with CMV (Morel et al., 2002; Boutet et al., 2003). Together, these studies indicate that RDR6, DCL2, DCL4, HEN1 and AGO1 all contribute to the RNA silencing-mediated defense of *Arabidopsis* against RNA virus infection.

3.8.3 Engineered virus resistance

The introgression of natural R genes into crop plants is one strategy to generate virus-resistant plants (Soosaar et al., 2005). However, this approach is limited to viruses for which R genes are known. Previous studies of recombinant viruses inspired the application of RNA silencing to engineer crops resistant to various wild-type viruses: for example, transgenic plants expressing dsRNA homologous to the virus. This research yielded, amongst others, rice resistant to *Rice yellow mottle virus, Vigna mungo* plants resistant to *Mungbean yellow mosaic virus*, and cassava plants resistant *to Africa cassava mosaic virus* (Pinto et al., 1999; Pooggin et al., 2003; Zhang et al., 2005a). The natural integration and selection of viral DNA sequences in plant genomes might also be a form of RNA silencing-based virus resistance (Hull et al., 2000; Mette et al., 2002). Understanding the multiple mechanisms of viral dsRNA and siRNA biogenesis will be important for improving virus resistance in such plants (Sudarshana et al., 2007; Vanderschuren et al., 2007).

3.9 Viruses as tools to study RNA silencing in *Arabidopsis*

3.9.1. Virus induced gene silencing (VIGS)

VIGS constructs are important experimental tools for studying RNA silencing (Lu et al., 2003; Dalmay, 2005). They are typically recombinant RNA viruses that contain sequences homologous to a host-encoded mRNA; infection with these constructs induces cleavage and degradation of the targeted messenger (Pantaleo et al., 2007). TGS can also be initiated by VIGS directed against a transgene promoter, although this approach is rarely taken (Al-Kaff et al., 1998; Jones et al., 2001). Many studies have used VIGS to silence GFP transgenes, or host genes whose deficiency causes bleaching (Kjemtrup et al., 1998; Ruiz et al., 1998; Jones et al., 1999; Dalmay et al., 2000; Deleris et al., 2006; Dunoyer et al., 2007; Gammelgard et al., 2007). This allows a visual evaluation of the extent of VIGS, and facilitates a genetic analysis of the process (Dalmay, 2005).

Similar to transgene-induced silencing, VIGS can spread cell-to-cell and systemically (Voinnet, 2005). The exact nature of the underlying silencing signal is unknown, but local cell-to-cell spread of transgene silencing correlates with DCL4-dependent, 21 nt siRNA expression (Dunoyer et al., 2005). Long-range, cell-to-cell movement of transgene silencing requires RDR6, which is otherwise coupled to DCL4 in the biogenesis of endogenous 21 nt siRNAs (Himber et al., 2003; Xie et al., 2005; Yoshikawa et al., 2005). It is speculated that RDR6 mediates a relay amplification of the silencing signal (Himber et al., 2003). Forward genetic screens recently showed that NRPD1a and RDR2 are required for the intercellular spread of RNA silencing induced by inverted repeat transgenes, but not for production of siRNAs from those silencing inducers (Dunoyer et al., 2007; Smith et al., 2007). However, the function that NRPD1a and RDR2 have in this form of RNA silencing does not extend to VIGS mediated by an RNA virus (Dunoyer et al., 2007).

3.9.2. Production of virus-derived siRNAs

As was observed for endogenous siRNAs, multiple size classes of viral siRNAs accumulate in virus infected plants (Hamilton et al., 2002). Viral siRNAs cloned from plants infected with the RNA viruses *Cymbidium ringspot tombusvirus* (CymRSV) or *Tobacco mosaic virus* (TMV-Cg) were predominantly ~20-22 nt in length (Molnar et al., 2005; Kurihara et al., 2007). This range of size classes was also observed by blot hybridization of low molecular weight RNA from plants infected with various RNA viruses (Bouche et al., 2006; Deleris et al., 2006; Fusaro et al., 2006; Diaz-Pendon et al., 2007). Recently, ~21-22

nt siRNA classes whose biogenesis required DCL2 and DCL4 were implicated in VIGS and in plant defense against RNA virus infection, as mentioned above (Xie et al., 2004; Deleris et al., 2006; Diaz-Pendon et al., 2007).

It is not clear whether VIGS and virus-derived siRNA biogenesis follow pathways similar to those of endogenous RNA silencing. The *trans*-acting siRNA and Pol IV pathways require an RDR step before the DCL-dependent production of siRNAs (**Figure 2B/C**). Dalmay *et al.* (2000) noted that viral RdRPs may complement RDR-deficiency, meaning that viral siRNA biogenesis would be autonomous of host RDRs. In spite of this, RDR6 appears to facilitate virus defense and VIGS in plants, based on the phenotype of infected *rdr6* mutants (Mourrain et al., 2000; Muangsan et al., 2004). Furthermore, *Arabidopsis* RDR1 and its tobacco ortholog, which are induced by virus infection or treatment with SA, reduce the susceptibility of plants to TMV and PVX (Xie et al., 2001b; Yu et al., 2003). Recently, *Arabidopsis* RDR1 was also implicated in viral siRNA biogenesis during CMV infection; however, the effect was only apparent when the 2b suppressor was deleted from CMV, and did not correlate with hypersusceptibility (Diaz-Pendon et al., 2007).

3.10. Experimental systems based on DNA viruses

DNA viruses, in contrast to RNA viruses, pass through both nuclear and cytoplasmic phases during replication and do not encode RdRPs (Hull, 2004). Until the work presented here, little was known about the genetic requirements for DNA virus-derived siRNA biogenesis. Therefore, a comprehensive model was not available, but the basic components of siRNA biogenesis induced by DNA virus infection would likely incorporate RDR and DCL-dependent steps (**Figure 3**), similar to those described for endogenous RNA silencing pathways (Meins et al., 2005). Alternatively, dsRNA formation could be RdRP-independent, if one posits the annealing of overlapping sense / antisense transcripts produced from the circular viral genome (Voinnet, 2001). To test different elements of this speculative model, I used the two experimental systems described below.

3.10.1 The geminivirus *Cabbage Leaf Curl Virus* (CaLCuV)

The geminivirus *Cabbage Leaf Curl Virus* (CaLCuV), which infects plants of the Brassicaceae family (Hill et al., 1998), has a bipartite ssDNA genome and is thought to replicate in the nucleus by a rolling-circle mechanism (Abouzid et al., 1992; Hanley-Bowdoin et al., 2000). This mechanism involves both sense and antisense viral DNA,

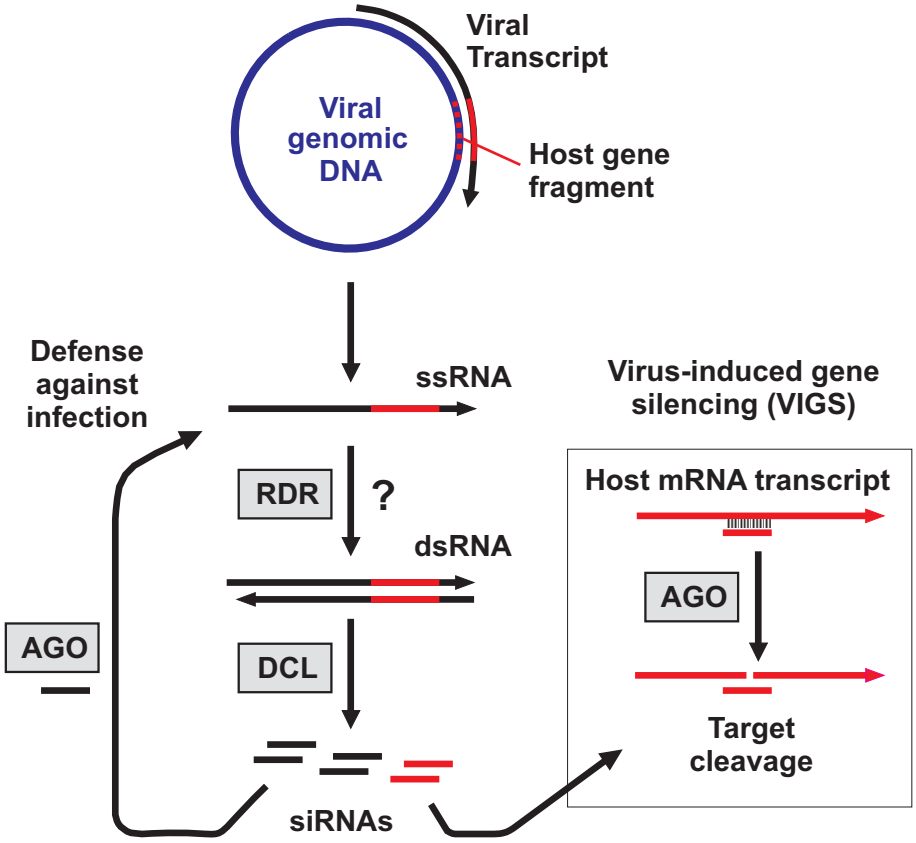


Figure 3. Putative mechanisms of DNA virus-derived siRNA biogenesis and function. The mechanism of dsRNA and siRNA biogenesis from transcripts of DNA viruses is unknown. Potentially, an RDR protein could act on viral ssRNA transcripts to synthesize a complementary RNA strand. The resulting dsRNA could then be processed by one or multiple *Arabidopsis* DCLs. Depending on their sequence homology, these siRNAs would target either viral RNA species (Defense against infection) or host mRNA transcripts (VIGS).

but no dsRNA intermediates (Gutierrez, 2000). CaLCuV is also an effective VIGS vector in *Arabidopsis* (Turnage et al., 2002; Muangsan et al., 2004). RDR6 appears to be required for VIGS triggered by a particular CaLCuV construct designed to silence the *Arabidopsis* gene *CHLI* (Muangsan et al., 2004). In addition, slightly higher levels of viral DNA were detected in an infected *rdr6* mutant than in the wild type (Muangsan et al., 2004). However, a full molecular genetic analysis CaLCuV infection, *CHLI* silencing and related viral siRNA biogenesis have yet to be undertaken.

3.10.2 The caulimovirus *Cauliflower Mosaic Virus* (CaMV)

Cauliflower Mosaic Virus (CaMV), an important model virus of the *Caulimoviridae* family (Hohn, 1999), has a circular dsDNA genome that replicates in two phases: (i) in the plant nucleus, the viral genome forms a minichromosome from which a more-than-genome length, 35S transcript is synthesized by the host RNA polymerase II. (ii) in the cytoplasm, the 35S transcript is reverse transcribed into dsDNA by a viral protein (Pfeiffer and Hohn, 1983; Hull, 2004). CaMV contains seven protein coding regions, whose products are translated from the 35S transcript and its spliced variants in the cytoplasm. The 35S transcript leader (~600 nt) contains numerous stem-loop hairpin structures and regulatory elements involved in splicing, polyadenylation, translation, reverse transcription, and packaging (Hemmings-Mieszczak et al., 1997; Pooggin et al., 1998).

Pfeffer *et al.* (2005) found that infection of human cells with the DNA virus, Epstein-Barr virus (EBV), involves expression of virus-encoded miRNAs. Another mammalian DNA virus, Simian virus 40, expresses a miRNA from a late viral gene that down-regulates expression of early viral genes, a process that may enhance the probability of successful infection (Sullivan et al., 2005). Most recently, a human cytomegalovirus-derived miRNA was shown to down-regulate a host immune system gene, suggesting a viral miRNA-based immunoevasion mechanism (Stern-Ginossar et al., 2007). With its highly folded structure, the 35S RNA leader of CaMV would be a good substrate from which to expect miRNA-like species.

3.11. Scope of this dissertation

When the work described here was started in 2002, it was known that RNA silencing functions in developmental regulation, targeting methylation to genomic DNA repeats, and protection against infection of RNA and DNA viruses in plants. Early studies of silencing mutants and the cloning of smRNAs had established that different size classes of smRNAs occur naturally and that key components of RNA silencing pathways, particularly in plants, are often encoded by gene families. At that time it was unclear, however, whether different members of the same protein family and size classes of siRNAs have specialized functions in RNA silencing; whether endogenous and foreign RNA sequences are silenced by the same or different mechanisms; and what role siRNAs play in methylation of endogenous genomic repeats. My major objective was to identify and functionally characterize siRNA biogenesis pathways associated with methylation of tandem genomic repeats. A second, related objective was to elucidate the pathways required for production of viral siRNAs in plants infected by DNA viruses differing in their mode of replication. I adopted a reverse genetics approach that involved establishing a panel of single and multiple mutant strains deficient in RNA silencing. These mutants were used to study patterns of siRNA expression, DNA methylation, and responses to virus infection.

This dissertation addresses several important questions concerning the biogenesis and function of siRNAs in *Arabidopsis*:

1. What are the genetic requirements for the biogenesis of siRNAs derived from tandem genomic repeats? (**Chapter 5**)
2. Do siRNAs representing the 5S rDNA tandem repeats contribute to cytosine methylation in these repeats? (**Chapters 5 and 6**)
3. How are factors required for global maintenance of CpG-methylation, such as DDM1, related to the siRNA pathways implicated in methylation of 5S rDNA repeats? Are these processes redundant and/or interacting? (**Chapter 6**)
4. Do proteins encoded by *DCL* and *RDR* gene families have specific functions in the biogenesis and function of the different size classes of viral siRNAs, which accumulate in plants infected with DNA viruses? Are these functions redundant? (**Chapter 7**)
5. Does viral siRNA biogenesis depend on the sequence and/or structure of viral RNA precursors? (**Chapter 8**)

4. Materials and Methods

4.1. Materials

4.1.1 *Arabidopsis thaliana* strains

In this study, wild-type *Arabidopsis thaliana* plants of the ecotypes Columbia-0 (Col-0) and Landsberg *erecta* (La-*er*) were used. A panel of *Arabidopsis* mutant strains (**Table 1**) deficient for RNA silencing-related genes was assembled. Col-0 was the reference wild-type ecotype for comparison to mutant strains unless otherwise stated., The *Arabidopsis* Information Resource identification numbers (TAIR ID#) in **Table 1** refer to gene models curated in the database http://www.arabidopsis.org/. Most mutant lines used were *Agrobacterium tumefaciens* T-DNA insertion lines from the SALK, SAIL, and GABI collections (Sessions et al., 2002; Alonso et al., 2003; Rosso et al., 2003), where the T-DNA interrupts an exon thought to be critical for protein activity. In additional, well-characterized ethyl methanesulfonate-induced (EMS) mutation lines were used, which produce an mRNA containing a premature stop codon (prem. stop), a misspliced exon (msp. Ex), or that prematurely terminates (prem. term.). The mutant *dcl1-8*, which is a point mutation in DCL1 helicase domain, and the mutant *dcl1-9*, which is a truncation of the DCL1 C-terminal double-stranded RNA binding domain, are known to be hypomorphic (Schauer et al., 2002). Homozygous, recessive mutants for an *Arabidopsis* gene (e.g., *DCL1*) were designated by a lowercase italic gene symbol (e.g., *dcl1*). However, the notation (-/-) was used to simplify diagrams, where homozygous mutants are compared to heterozygous and wild-type individuals -- e.g., *DCL1*(-/-), *DCL1*(+/-) and *DCL1*(+/+).

4.1.2 Plant virus strains

The *Cabbage Leaf Curl Virus* (CaLCuV) strain used in this study was originally cloned, identified and sequenced by (Abouzid et al., 1992): NCBI accessions NC_003866 (DNA-A) and NC_003887 (DNA-B). I used CaLCuV constructs generated by (Turnage et al., 2002), which allow efficient inoculation of *Arabidopsis*: NCBI accessions AY279346 (DNA-A in pMTCbLCVA.008) and AY279344 (DNA-B in pCPCbLCVB.002). The *Cauliflower Mosaic Virus* (CaMV) strain was CM1841 (Gardner et al., 1981); a redundant plasmid clone of CM1841 (pCa122, Kobayashi et al., 2002) was used for viral inoculation.

Locus	Allele	TAIR ID#	Type of mutation	Ecotype	Literature reference
AGO4	*ago4-1*	At2g27040	EMS, prem.term.@595aa	La-*er*	(Zilberman et al., 2003)
DCL1	*dcl1-8* (*sin1-2*)	At1g01040	EMS, mutation I431K	Col-*gl1*	(Schauer et al., 2002)
DCL1	*dcl1-9* (*caf*)	At1g01040	T-DNA, Del. after T1835	La-*er*	(Schauer et al., 2002)
DCL1	*dcl1-9* (BC5)	At1g01040	*same as above*	Col-0	*none*
DCL2	*dcl2-5*	At3g03300	T-DNA, SALK_123586	Col-0	(Akbergenov et al., 2006)
DCL3	*dcl3-1*	At3g43920	T-DNA, SALK_005512	Col-0	(Xie et al., 2004)
DCL3	*dcl3-2*	At3g43920	T-DNA, GABI_327D02	Col-0	*none*
DCL4	*dcl4-2*	At5g20320	T-DNA, GABI_160G05	Col-0	(Xie et al., 2005)
DDM1	*ddm1-2*	At5g66750	EMS, G→A msp. Ex11	Col-0	(Vongs et al., 1993)
HEN1	*hen1-5*	At4g20910	T-DNA, SALK_049197	Col-0	(Vazquez et al., 2004a)
HYL1	*hyl1-2*	At1g09700	T-DNA, SALK_064863	Col-0	(Vazquez et al., 2004a)
MET1	*met1-3*	At5g49160	T-DNA, private collection	Col/C24	(Saze et al., 2003)
NRPD1a	*nrpd1a-3*	At1g63020	T-DNA, SALK_128428	Col-0	(Herr et al., 2005)
NRPD1b	*drd3-7*	At2g40030	EMS, prem.stop: Q605X	Col-0	(Kanno et al., 2005b)
NRPD2a	*nrpd2a-1*	At3g23780	T-DNA, SALK_095689	Col-0	(Herr et al., 2005)
NRPD2a	*drd2-4*	At3g23780	EMS, prem.stop: R494X	Col-0	(Kanno et al., 2005b)
RDR2	*rdr2-1*	At4g11130	T-DNA, SAIL_1277	Col-0	(Xie et al., 2004)
RDR2	*rdr2-2*	At4g11130	T-DNA, SALK_059661	Col-0	(Vazquez et al., 2004b)
RDR6	*rdr6-15*	At3g49500	T-DNA, SAIL_617	Col-0	(Allen et al., 2005)

Table 1. Mutant strains of *Arabidopsis thaliana.*

T-DNA ***Agrobacterium tumefaciens*** **mediated T-DNA insertion mutants:**

GABI	"Genomanalyse im biologischen System Pflanze" T-DNA insertion	(Alonso et al., 2003)
SALK	Salk institute T-DNA insertion	(Sessions et al., 2002)
SAIL	Syngenta Arabidopsis Insertion Library T-DNA insertion	(Rosso et al., 2003)

EMS **Ethyl methanesulfonate-induced mutation lines:**

AαB	DNA point mutation that changes amino acid A to B at position α in the protein
G→A, Ex11	DNA point mutation from guanosine to adenosine causing misspliced Exon 11
Prem. term.	Premature termination of transcription caused by point mutation
Prem. stop	Premature stop codon introduced by point mutation

4.1.3. Bacterial strains and growth medium

The standard *Escherichia coli* (*E. coli*) strain DH5α was used to propagate TOPO pCRII clones (Invitrogen, Carlsbad, USA) containing cloned smRNA inserts and plasmids containing probes for DNA blot hybridization. *E. coli* were grown in liquid LB medium (Luria-Bertani medium: 1% Bacto-tryptone, 0.5% (w/v) Bacto yeast extract, 0.5% (w/v) NaCl) or on LB plates (supplemented with 1.5% Bacto-agar). The antibiotics Ampicillin (100 µg/ml) or Kanamycin (50 µg/ml) were added for selection of plasmids.

4.1.4. Plasmid vectors

DNA fragments for hybridization probes were cloned in the plasmid vectors pBluescript SK- (Stratagene, La Jolla, USA) or pCRII (Invitrogen). Sequence inserts for conventional sequencing of smRNAs were ligated into the pCRII vector.

4.1.5. Enzymes and reagents

The enzymes used in this study were purchased from Ambion (Austin, USA), Amersham (GE Healthcare, Chalfont St. Giles, UK), New England Biolabs (Ipswich, USA), Promega (Madison, USA) and Roche Diagnostics (Basel, Switzerland). Chemicals were obtained from Fermentas (Burlington, Canada), Fluka (Buchs, Switzerland), Merck (Darmstadt, Germany), and Sigma-Aldrich (St. Louis, USA), and were of analytical grade. Radioactively labeled ^{32}P was obtained from Amersham.

4.1.6. Oligonucleotides

Oligonucleotides were designed manually and then synthesized by Microsynth (Balgach, Switzerland). Sequences of oligonucleotides used as probes for RNA blot hybridization experiments are listed in Appendix, **Table A1**. Sequences of oligonucleotides used as primers for Polymerase Chain Reaction are listed in Appendix, **Table A2**.

4.2. Methods

4.2.1 Plant growth conditions

Arabidopsis wild-type and mutant plants used for viral infection experiments were grown from seeds in soil in a free-standing phytochamber (Sanyo Co., Tokyo, Japan) with 12 h day and 12 h night at 20°C. Plants used for analysis of endogenous DNA methylation and siRNA biogenesis were grown from seeds in soil in a walk-in phytotron with 16 h day and 8 h night (long day), at 21°C with 65% humidity. The light intensity on shelves in the phytotron was adjusted to ~10'000 lux (cool white 36W; Philips N.V., Amsterdam, Netherlands).

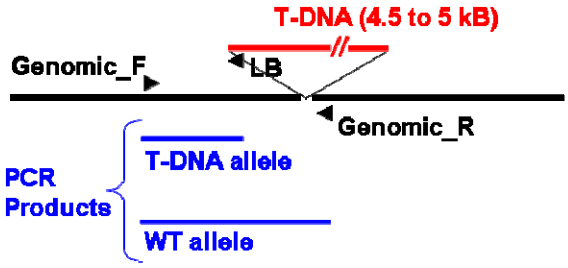
4.2.2 Genomic DNA extraction for genotyping mutant lines

One leaf from each plant was placed in a 1.5 mL tube (Eppendorf, Hamburg, Germany). To each tube was added one 3 mm tungsten carbide bead (Qiagen N.V., Venlo, The Netherlands) and 300 µL crude genomic DNA extraction buffer (0.2 M Tris-HCl, 0.25 M NaCl, 0.025 M EDTA, 0.5% SDS, pH 7.2). Up to 48 tubes were simultaneously processed in a Mixer Mill 300 (Retsch GmbH, Haan, Germany) using the 2x 24 Mixer Mill Adaptor Set (Qiagen). The leaves were ground 2 min. with the 30 Hz setting. The extracts were spun in a microcentrifuge for 5 min. (10'000 g), and then the supernatants were pipetted into new 1.5 mL tubes through Miracloth squares (2 cm^2) to remove leaf debris. These crude DNA solutions were mixed with 300 µL isopropanol, incubated for 15 min. at RT and spun for 15 min. (10'000 g). The pellets were washed with 75% ethanol, and after air-drying 15 min., they were resuspended in 40 µL of 10 mM Tris-HCl buffer (pH 8.5).

4.2.3 Polymerase Chain Reaction

Preparative Polymerase Chain Reactions (PCR) were performed using a standard 50 uL reaction mixture consisting of 41 µL H_2O, 5 µL 10x PCR buffer (Eppendorf), 1 µL dNTPs (2 mM each, Fermentas), 1 µL forward primer (10 µM), 1 µL reverse primer (10 µM) and 0.25 µL Taq polymerase enzyme (5 U/µL, Eppendorf). Approximately 20-200 ng of DNA template (in 1-2 µL H_2O) was added to each reaction. PCR was performed in a T3 Thermocycler (Biometra Analytik GmbH, Göttingen, Germany). For genotyping purposes, 0.5x reactions (25 µL) were used to test each DNA sample.

4.2.4 Genotyping T-DNA insertions



PCR was performed to identify homozygous plants from populations segregating a T-DNA insertion. Primer combinations required to genotype specific lines are shown in **Table 2** and the primer sequences are provided in Appendix, **Table A2**. Genotypes of individuals were determined by two parallel PCR reactions: (i) with primers designed for wild-type genomic DNA sequences flanking the T-DNA insertion (**Genomic_F** / **Genomic_R**); and (ii) with a genomic primer and a T-DNA border-specific primer (**Genomic_F / LB**). If only the first reaction produced a product, then the plant was wild type. If only the second reaction produced a product, then the plant was a homozygous mutant. If both reactions produced products, then the plant was hemizygous for the T-DNA insertion.

Mutant	Genomic_F	Genomic_R	T-DNA_LB
dcl1*-9*	dcl1_geno_F1	dcl2_geno_R1	LB-pGV_CAF1_2
dcl2*-5*	dcl2_geno_F2	dcl2_geno_R2	LB-new_SALK
dcl3*-1*	dcl3-1_geno_F1	dcl3-1_geno_R1	LB-new_SALK
dcl3*-2*	dcl3-2_geno_F1	dcl3-2_geno_R1	GABI_8409A
dcl4*-2*	dcl4_geno_F1	dcl4_geno_R1	GABI_8409A
nrpd1*-3*	nrpd1_geno_F1	nrpd1_geno_R1	LB-new_SALK
rdr2*-1*	rdr2-1_geno_F1	rdr2-1_geno_R1	LBb1_SAIL
rdr2*-2*	rdr2-2_geno_F1	rdr2-2_geno_R1	LB-new_SALK
ddm1*-2*	ddm1-2_F1	ddm1-2_R1	n / a

Table 2. PCR primer combinations for genotyping mutant strains of *Arabidopsis*.

4.2.5 Genotyping the *ddm1-2* point mutation

The *ddm1-2* point mutation is a G→A mutation that causes missplicing of Exon 11 in the DDM1 protein (Jeddeloh et al., 1999). A 400 bp region including this mutated base pair position was PCR amplified (primers, **Table 2**) from genomic DNA extracted from individual plants. The PCR products were purified using QIAquick columns (Qiagen). Sequencing these purified products generated a chromatogram from which to call the relevant base pair: i.e., in the forward-oriented sequence, 'G' for a homozygous *DDM1-2* (+/+) or 'A' for homozygous *DDM1-2* (-/-) plants. If the chromatogram was ambiguous

(both G and A peaks) then the sampled plant was heterozygous *DDM1-2* (+/-). Genotypes were confirmed using the reverse-oriented sequencing reaction.

4.2.6 DNA blot hybridization

This method, adapted from the original "Southern blot" (Southern, 1975), was used to analyze cytosine methylation of genomic DNA in the manner described by (Cedar et al., 1979) and numerous others subsequently (Oakeley, 1999). Molecular techniques that follow were performed using (Sambrook and Russell, 2001) for general guidance.

Genomic DNA isolation. Genomic DNA was isolated from *Arabidopsis* leaves by grinding them in liquid nitrogen and extraction using the *Nucleon PhytoPure, plant and fungal DNA extraction kit* (Amersham) with two modifications: (i) RNase A was added after suspension of the plant tissue in Reagent #1 with an incubation for 30 min. at 37°C, and (ii) after chloroform/resin extraction the aqueous phase was incubated on ice for 30 min. and then clarified by centrifugation for 15 min. (>5000 g). The DNA was sedimented by adding an equal volume of isopropanol, incubating for 10 min. at RT and then 10-20 min. on ice, followed by centrifugation for 30 min. (>5000 g). After washing with 75% ethanol, the DNA pellet was dissolved in sterile H_2O for 16-24 h at 4°C. DNA concentrations were calculated from optical densities measured at the 260 nm wavelength (OD_{260}).

Restriction enzyme digestion. Genomic DNA (3 ug) was dissolved in 32 uL sterile water and mixed with 4 uL of 10X Restriction Enzyme Buffer (New England Biolabs) in an Eppendorf tube. After incubation for 1-2 h at 4°C, 1.5 µL of *Hpa* II restriction enzyme (10 U/µL, New England Biolabs) was added and the tube was incubated for an initial 3-4 h at 37°C. A second 1.5 µL aliquot of *Hpa* II was added and the tubes were incubated in a 37°C-oven overnight (12-14 h). The next day, a third aliquot of 1.0 µL of *Hpa* II was added and the tubes incubated for 3-4 hours at 37°C. The total *Hpa* II used for each sample digestion was thus 40 U.

***Agarose gel electrophoresis*.** A 15 cm-long 0.8% agarose gel (1x TBE: 0.09 M Tris, 0.09 M Borate, 0.002 M EDTA) containing Ethidium Bromide was prepared. After removing the large comb (1.5 cm wide x 1mm thick slots) the slots were washed out with 1x TBE using a syringe. To each genomic DNA digest were added 10 µL of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in H_2O). The digests were loaded along side a 1 kb Plus DNA Size Marker (Invitrogen) lane. The gel was run at 70-

80 V for about 6 h using a standard power supply (Bio-Rad), until the bromophenol blue color marker had run ~3/4 the length of the gel. The position of size markers and digestion of DNA were verified using a UV transilluminator (Bio imaging System, Syngene/Synoptics Ltd., Cambridge, UK).

Capillary gel transfer. The gel was bathed in *Depurination Solution* (0.2 N HCl) for 30 min. (until the colors of the size markers changed). Afterward, it was washed with bidest H2O three times for 5 min, then placed in *Denaturation Solution* (0.5 M NaOH; 1.5 M NaCl) for 30 min. Finally the gel was equilibrated in *Neutralization Buffer* (1.5 M NaCl; 1 mM EDTA; 0.5 M Tris-HCl) twice for 15 min. Capillary gel transfer to a Hybond-N nylon membrane (Amersham) was carried out overnight with 10x SSC (75 mM NaCl, 7.5 mM Na_3Citrate, pH 7.0). The membrane was then UV cross-linked (energy = 140 mJ with a Stratalinker apparatus; Stratagene).

***Hybridization*.** 30 mL of hybridization buffer (0.25 M Sodium Phosphate buffer, pH 7.2; 1 mM EDTA; 6.6 % SDS; 1% BSA) -- i.e., Church Buffer (Church and Gilbert, 1984) -- was added to a large hybridization tube containing the cross-linked nylon membrane, which was then pre-hybridized at 50°C for ~2 h (Hybridiser HB-1D, Techne Inc., Burlington). DNA probes were labeled internally with $\alpha[^{32}P]$-dCTP with the RadPrime DNA Labeling System (Invitrogen) using DNA fragments representing the 5S rDNA or 180-bp repeats (for sequences, **Figure A3**); the probe was added to the tube and allowed to hybridize overnight at 50°C. The membrane was then washed twice for 5 min. with 2x SSC, 0.1% SDS; and twice 20 min. with 0.1 x SSC, 0.1% SDS. Finally, the membrane was wrapped in plastic wrap and exposed to either a Phospho-imager screen at RT or a Kodak BioMax MR Film (Eastman Kodak Co., Rochester, USA) at -80°C.

4.2.7 Viral inoculation of *Arabidopsis* by particle bombardment

For virus infection experiments, rosette-stage plants were inoculated 4-5 weeks post-germination with infectious plasmids containing the genomic sequences of either CaLCuV::*ChlI* or CaMV. Inoculation was performed by delivering 150 µg of plasmid-coated, 1 µm gold particles to each individual rosette using a biolistic particle delivery system (PDS-1000/He) according to the manufacturers instructions (Bio-Rad Laboratories, Hercules, USA). For CaLCuV::*ChlI* inoculations, 0.5 µg each of the plasmids pMTCbLCVA.008 and pCPCbLCVB.002 were used as a plasmid mixture (Turnage et al.,

2002). For CaMV inoculations, 1 µg of the plasmid pCa122 was used (Kobayashi et al., 2002).

4.2.8 Determination of virus titers

One month post-inoculation, unless otherwise stated, virus-infected plants were harvested in pools and ground in liquid N_2 for total RNA and DNA preparations. Virus titers were measured as described by (Pooggin et al., 1998). In brief, total DNA was isolated from a small amount of the pooled tissue powder and semi-quantitative duplex PCR was performed using the following primers: as an internal control the 18S rDNA-specific primers 5'-TGT GAT GCC CTT AGA TGT TCT GGG-3' and 5'-ATC ATT CAA TCG GTA GGA GCG ACG-3' (yielding a 234 bp product) along with either the CaLCuV-specific primers 5'-GCA TCA CTA AGA GCG TGG ACT AC-3' and 5'-CTC TCA TCG AAG TCG TCC AGA CT-3' (393 bp) or the CaMV-specific primers 5'-GGG TTC TTA TAG GGT TTC GCT-3' and 5'-ACG TAC CTC TAT CAA ATT TCC A-3' (397 bp).

4.2.9 Extraction of total RNA

Total RNA was extracted from aliquots of *Arabidopsis* tissues using the TRIzol (Invitrogen) method. In brief, 1 g aliquots of frozen plant tissue (ground in liquid N_2 and stored at -80°C) were transferred to 15 mL centrifuge tubes. TRIzol reagent was added (10 mL), the tubes were vortexed several times and incubated 5 min. at RT. Then 2 mL of chloroform was added and the tubes vortexed for 20 s and incubated 3 min. at RT. The samples were centrifuged for 15 min. (11'500 g) at 4°C. The aqueous phase (~6 mL) was transferred to new tubes. To each tube were added 6 mL of isopropanol and tube was inverted to mix. The solution was incubated at RT for 10 min. and then transferred to ice for 30-60 min. The RNA was sedimented by centrifugation for 30 min. (11'500 g) at 4°C. After discarding the isopropanol, the pellet was washed with 75% ethanol in diethylpyrocarbonate (DEPC)-treated H_2O. The ethanol was discarded and the tube allowed to air-dry 15 min. Then 60 µL of DEPC-treated H_2O (65°C) was pipetted down the side of the tubes and carefully swirled across their entire interior surface. The tubes were then centrifuged at low speed to collect aqueous RNA. This solution was transferred to 1.5 mL Eppendorf tubes, new 60 µL aliquots of DEPC-treated H_2O were added to the large tubes and the resulting RNA combined with the first aliquots. RNA samples were stored at -20°C.

4.2.10 RNA blot hybridization

High Molecular Weight (HMW) RNA blots. TRIzol-extracted total RNA was further purified using RNA Easy Columns (Qiagen). 8 μg total RNA was run on a 1% agarose (formaldehyde) gel and transferred to a nylon membrane (N+; Amersham). DNA probes were labeled internally with $\alpha[^{32}P]$-dCTP using the RadPrime DNA Labeling System (Invitrogen). Hybridization was performed overnight at 50°C in 0.5M sodium phosphate buffer (pH 7.5) supplied with 1mM Na-EDTA, 6.6% SDS and 1% BSA. The membrane was washed twice for 5 min in 2x SSC, 0.1% SDS and twice for 20 min in 0.1x SSC, 0.1% SDS at 60°C. The *CHLI.1* gene (At4g18480) probe was synthesized using a 411 bp cDNA fragment (nt 41 to 451, TAIR gene model) located 5' of the VIGS-target (nt 470-836) as the template; and the *ACT2* gene (At3g18780) probe was synthesized using a 404 bp cDNA fragment (nt 1095-1498, TAIR gene model) as the template.

Low Molecular Weight (LMW) RNA blots. Total RNA was extracted from ~1.0 g plant tissue using TRIzol reagent (Invitrogen). In most cases, total RNA was fractionated using RNAeasy columns (Qiagen) following the "RNA cleanup" protocol. LMW RNA was recovered from the column flow-through by mixing it with an equal volume of isopropanol, incubating overnight at 4°C, and sedimentation by centrifugation for 30 min. (11'500 g) at 4°C. The isopropanol was discarded and the pellet washed with 75% ethanol in DEPC-treated H_2O. The ethanol was discarded and the pellet air-dried for 15 min. LMW RNA was resuspended in 50-100 μL DEPC-treated H_2O. A 15 μg aliquot of each sample was transferred to a new tube, dried in a Speed-Vac apparatus (Thermo Scientific, Waltham, USA), and re-suspended in 10 μL of loading buffer (95% formamide, 20 mM EDTA pH 8.0, 0.05% bromophenol blue, 0.05% xylene cyanol). These samples were heated at 95°C for 2 min. and loaded on to an 18% polyacrylamide gel (a 19:1 ratio of acrylamide to bis-acrylamide, 8M urea). Gel electrophoresis was performed with 1x TBE in a vertical chamber at 350-450 V (13-15 mW) for 2-3 h using a Constant Power Supply 3000/150 (Pharmacia, Uppsala, Sweden) until the xylene cyanol FF marker was ~3 cm from the lower buffer reservoir. Then the smRNA was transferred to a Hybond N+ membrane by electroblotting in 1x TBE buffer at 10 V overnight. Blot hybridizations were performed at 35°C for 12 to 18 h in PerfectHyb Plus buffer (Sigma-Aldrich). The probes used (**Table A1**) were end-labeled with $\gamma[^{32}P]$-ATP (Amersham) using polynucleotide kinase (Roche) and purified using a MicroSpin™ G-25 column (Amersham). The blot was washed three times with 2x SSC, 0.5% SDS for 30 min. at 35°C. The signal was detected after 1 to 3 days exposure to a phosphor screen using the Typhoon 9400 Variable Mode Imager (Molecular

Dynamics Inc., Sunnyvale, USA), or on Kodak MR film 1-2 days (at -80°C). For repeated hybridization the membrane was stripped twice with 0.1% SDS for 30 min. at 80°C.

4.2.11 Quantitative reverse-transcription PCR

Poly-dT cDNAs were made using Superscript III Reverse Transcriptase (Invitrogen) and 2 µg of RNA as the template. Real-time PCR was performed in 96-well titer plates on an ABI PRISM 7000 SDS apparatus with SYBR GREEN PCR Master Mix (Applied Biosystems, Foster City, USA) following the manufacturers' recommendations (95°C for 5 min., followed by 40 cycles: 95°C for 30 s, 60°C for 45 s). Primers were: for *CHLI.1,* 5'-AAA TGG CGT CTC TTC TTG GAA CAT C-3' and 5'-GCC TTA TTT GGA TTC CTG CAT TTA ACT TG-3'; and, for normalization with ACT2, 5'-GCA CCC TGT TCT TCT TAC CG-3' and 5'-AAC CCT CGT AGA TTG GCA CA-3'. Uncertainties were propagated from standard errors for triplicate measurements of cDNA (from column-purified RNA of 3-4 plants).

4.2.12 Sequencing of smRNAs

Ligation of *Arabidopsis* smRNAs and cloning by PCR amplification were performed essentially as described in Llave *et al.* (2002a) with modifications described in Arazi *et al.* (2005). Briefly, total RNA (400-500 µg) was isolated from two sample pools: CaMV infected Col-0 (CaMV Col-0) and CaMV infected triple *dcl*-mutant *dcl2 dcl3 dcl4* (CaMV *d234*). The total RNA from each sample was loaded into two large wells of a 15% polyacrylamide gel along with 15, 21, 24 and 30 nt RNA oligonucleotide standards at the edges. The gel was run for 3 h and RNA is the excised region corresponding to the ~16-28 nt size range was eluted by extraction with two aliquots of 3 M NaCl in DEPC-treated H_2O. The eluted RNA was stored overnight at -20°C in the presence of 70% (v/v) ethanol, 200 mM NaOAc (pH 5.2) and 200 µg glycogen, and then sedimented by centrifugation for 30 min. (10'000 g). smRNAs in samples were sequentially ligated to 5'- and 3'-oligonucleotide adapters (**Table 3**; Dharmacon, Lafayette, USA) using T4 RNA ligase (Ambion).

Adapter	Sequence		Sample
5.1	TGG GAA TTC CTC ACT **aaa**	5'-adapter:	CaMV Col-0
5.2	TGG GAA TTC CTC ATG **aaa**	5'-adapter:	CaMV *d234*
3.454C	Ⓟ **uuu** TGA TCC ATG GAC TGT idT	3'-adapter:	CaMV Col-0 & CaMV *d234*

Table 3. Adapters used for smRNA ligation and sequencing. Bold, lowercase letters represents ribonucleotides, while the remainder of the adapter is composed of DNA. Blue letters in 5.1 and 5.2 are "bar codes" to distinguish the two samples.

The 5'-Adapter products were gel purified after each ligation reaction, and the final product used as the template for a reverse transcription reaction (Superscript II, Invitrogen) with primers designed to bind the 3'-adapter sequences. The PCR products obtained with forward and reverse primers that bind the 5'- and 3'-adapter sequences were purified by extraction from a 10% non-denaturing polyacrylamide gel electrophoresis. TOPO pCRII TA-based cloning was used to prepare 5'-adapter::smRNA-insert::3'-adapter products for conventional sequencing. cDNA containing ligated smRNAs was sequenced by Fasteris Life Sciences (Plan-les-Ouates, Switzerland) using Solexa high-throughput technology.

4.2.13 Bioinformatic analysis of smRNAs

A custom Perl script was used to parse insert sequences from the Solexa-determined full-length sequences containing the 5' and 3'-adapters. The two samples had been sequenced in parallel using barcodes in the 5'-adapters (**Table 3**) to differentiate them. CaMV Col-0 and CaMV *d234* sequences were sorted into separate files at this step. This script yielded 127'499 unique inserts with intact adapter ends from the Col-0 CaMV sequences and 126'952 from the *d234* CaMV sequences. Thus, the scale of the two sequence datasets is roughly equivalent.

The *Basic Local Alignment Search Tool* (BLAST, NCBI version 2.2.16) was used to compare cloned smRNA sequences to the CaMV and *Arabidopsis* genomes (Altschul et al., 1990). For the complete CaMV genome (Franck et al., 1980), the accession NC_001497 was obtained from http://www.ncbi.nlm.nih.gov. For *Arabidopsis*, the assembled *Arabidopsis* pseudo-chromosomes were obtained from ftp://ftp.arabidopsis.org -- ATH1_chr1.1con, ATH1_chr2.1con, ATH1_chr3.1con, ATH1_chr4.1con,

ATH1_chr5.1con, 01/22/2004 versions (*Arabidopsis*_Genome_Initiative, 2000). The viral and plant genomes were formatted as separate BLAST databases (formatdb) (Korf et al., 2003). Searches were carried out using the executable "blastall –blastn" with the parameters :

Genome	**-blastn parameters**
CaMV	`-W 8 -F F -G -3 -E -3 -m 8 -e 0.00001`
Arabidopsis	`-W 8 -F F -G -3 -E -3 -m 8 -e 0.005`

The smaller expect value (-e 0.00001) used for the CaMV Blast search reflects the much smaller viral genome size (8 kb) compared to the *Arabidopsis* genome (120 000 kb). Using these -blastn parameters, the following number of 19-26 nt perfect genomic matches were found: 31839 (25.1%) endogenous and 35641 (28.1%) viral from the CaMV Col-0 reads; 37978 (29.7%) endogenous and 24540 (19.2%) viral from the CaMV *d234* reads.

4.2.14 Images

Digital photographs were taken using a Canon EOS-D30 camera with either a 28mm lens (1:2.8) or a 100mm Macro EF lens (1:2.8) (Canon Inc., Tokyo, Japan). Minimal adjustments were made to the color and contrast settings of these images using Adobe Photoshop (version 9.0.2; Adobe, San Jose, USA).

5. *Arabidopsis* DCLs and the methylation of genomic tandem repeats

Endogenous repeat sequences in many eukaryotic genomes are subject to RNA silencing, DNA methylation and/or repressive histone modifications (Zaratiegui et al., 2007). In *Arabidopsis*, the RNA polymerase IV (Pol IV) pathway is clearly linked to RNA-directed DNA methylation (RdDM) of endogenous repeats. However, the role of *Arabidopsis* DCLs in RdDM of high copy number, genomic tandem repeats remains puzzling. Moreover, maintenance of methylation in these repeats is primarily dependent on MET1 and DDM1, which are not necessarily guided by siRNAs. In this chapter I analyze the genetic requirements for biogenesis of tandem repeat-derived siRNAs and examine how the accumulation of these siRNAs correlates with cytosine methylation in the repeats.

5.1. Tandem repeat-derived siRNAs

Numerous tandem repeat-derived siRNAs were identified by analyzing sequences from a wild-type *Arabidopsis* cDNA library of 20-25 nt smRNAs (in part described in Kutter *et al.*, 2007). The complete dataset contained approximately 250 unique clones matching the *Arabidopsis* genome. At least 35% of these clones matched repetitive sequences, whereas 5% were assigned to known miRNAs (**Figure 4A**). Two important classes of tandem genomic repeat were represented by multiple, partially overlapping clones: these sequences (prefix “At”) were compared to a publicly available dataset (prefix “siRNA”) (Gustafson et al., 2005), to consolidate clusters of siRNAs derived from the *Arabidopsis* 180 bp centromeric satellite and 5S rDNA repeats (**Figure 4B/C**). The latter cluster included siRNA1003, which has been used to analyze 5S rDNA-derived siRNA biogenesis (Xie et al., 2004; Onodera et al., 2005). DNA sequences from two representative BAC clones of the *Arabidopsis* genome are shown for comparison to the siRNA sequences (**Figure 4B/C**).

5S rRNA genes are arranged in large tandem repeat arrays, which contain hundreds of ~500 bp repeat units (Cloix et al., 2000). The siRNA sequences matched 5’ and 3’ of the 120 bp transcribed part of the 5S rRNA gene, in a region known as the “spacer” (**Figure 4D**). Since At157 and the other clones of this cluster do not overlap with the transcript region, it can be excluded that they arise directly from mature 5S rRNAs. This is an important distinction because recently published high-throughput sequencing

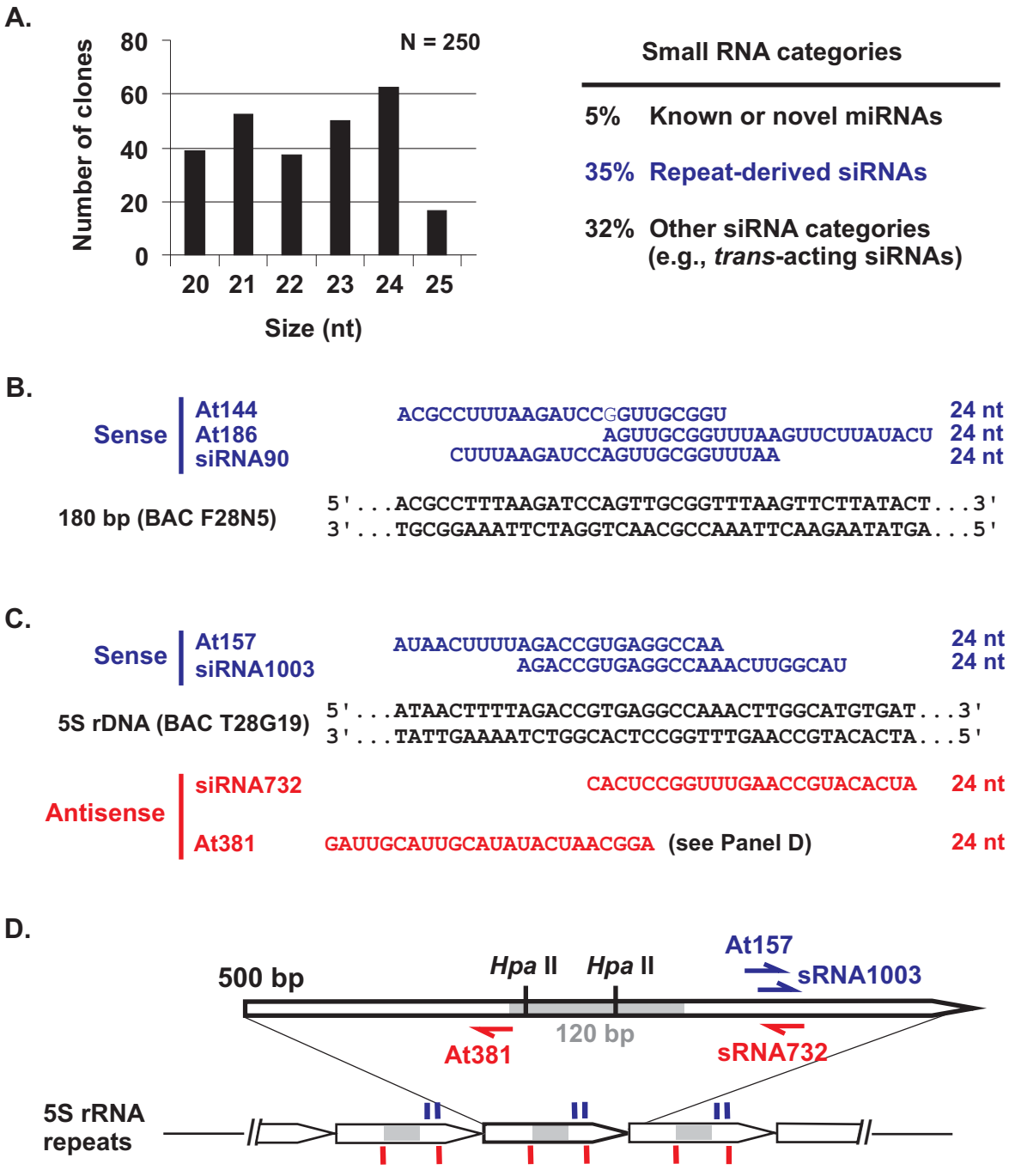


Figure 4. Cloned siRNAs corresponding to *Arabidopsis* genomic tandem repeats. (A) Size distribution of smRNA clones from RNA extracted from wild-type, flowering *Arabidopsis* plants (250 unique clones, 20-25 nt). Bioinformatic analysis and publicly available databases were used to annotate the smRNA categories (i.e., siRNA or miRNA). **(B) 180 bp satellite repeat sequences:** Three 24 nt siRNAs and one 21 nt siRNA (not shown) clustered to the consensus sequence of 180 bp satellite repeats. **(C) 5S rDNA repeat sequences:** Four 24 nt siRNAs clustered to 5S rRNA gene repeats. **(D) Arrangment of siRNAs derived from the 500 bp, 5S rDNA repeat unit.** These siRNAs matched the spacer of the 5S rRNA genes (white), but not the mature transcript region (gray).

data suggests that 19-26 nt RNAs also arise from non-specific degradation of 5S rRNAs (Rajagopalan et al., 2006). For the purposes of this dissertation, I focused on At157, which is representative of the 3' siRNA cluster.

Previous studies established that the biogenesis of 5S rDNA-derived siRNAs requires proteins of the Pol IV pathway (Xie et al., 2004; Onodera et al., 2005). The 180 bp satellite repeats are similar to 5S rRNA genes, in that they are arranged in large tandem arrays, but differ from them in showing only a very low level of transcription (May et al., 2005). Additionally, 180 bp repeats are of much higher copy number (>15'000 copies) than the 5S rDNA repeats (~1000 copies) (Campell et al., 1992; Heslop-Harrison et al., 2003; Jiang et al., 2003). Cytosines in both types of tandem repeat are known to be highly methylated via MET1/DDM1-depenent processes (Vongs et al., 1993). The discovery of 24 nt siRNA clusters derived from these repeats suggested that siRNAs could play a role in methylation of high-copy number, endogenous tandem repeat arrays.

5.2. Redundancy of DCLs for repeat-associated siRNA biogenesis

The expression of siRNAs cloned from 5S rDNA repeats (represented by At157) and 180 bp satellite repeats (represented by At144) was analyzed by blot hybridization of low molecular weight RNA. Both At157 and At144-specific probes detected 23-24 nt bands in RNA from wild-type plants. While absent in *nrpd1a* and *rdr2* deficiency mutants, these bands were present in the *rdr6* deficiency mutant (**Figure 5A**). *nrpd1a* is a null allele for the second largest subunit of Pol IVa (Herr et al., 2005; Onodera et al., 2005). This confirmed that NRPD1a and RDR2 are required for the accumulation of these repeat-associated siRNAs. The *trans*-acting siRNA pathway is dispensable for their production, since RDR6 was not required (Vazquez et al., 2004b). As controls, the membranes were hybridized with a probe for siR255 (21 nt), whose biogenesis requires RDR6 (Allen et al., 2005), and a probe for miR173 (22 nt), which is produced by the miRNA pathway (Reinhart et al., 2002).

DCL3 mediates production of ~24 nt siRNAs downstream of RDR2 in the Pol IV pathway (Xie et al., 2004; Pontes et al., 2006). In early experiments I found that the At144 and At157 expression patterns (23-24 nt) in various tissues required DCL3 (data not shown). However, I also observed an aberrant set of siRNA bands in the deficiency mutant *dcl3* (**Figure 5B**); this suggested that alternative DCLs function with the Pol IV pathway in the absence of DCL3. To test this idea I made genetic crosses between the three *dcl*-mutants known to impair siRNA biogenesis (*dcl2*, *dcl3*, and *dcl4*). I obtained multiple

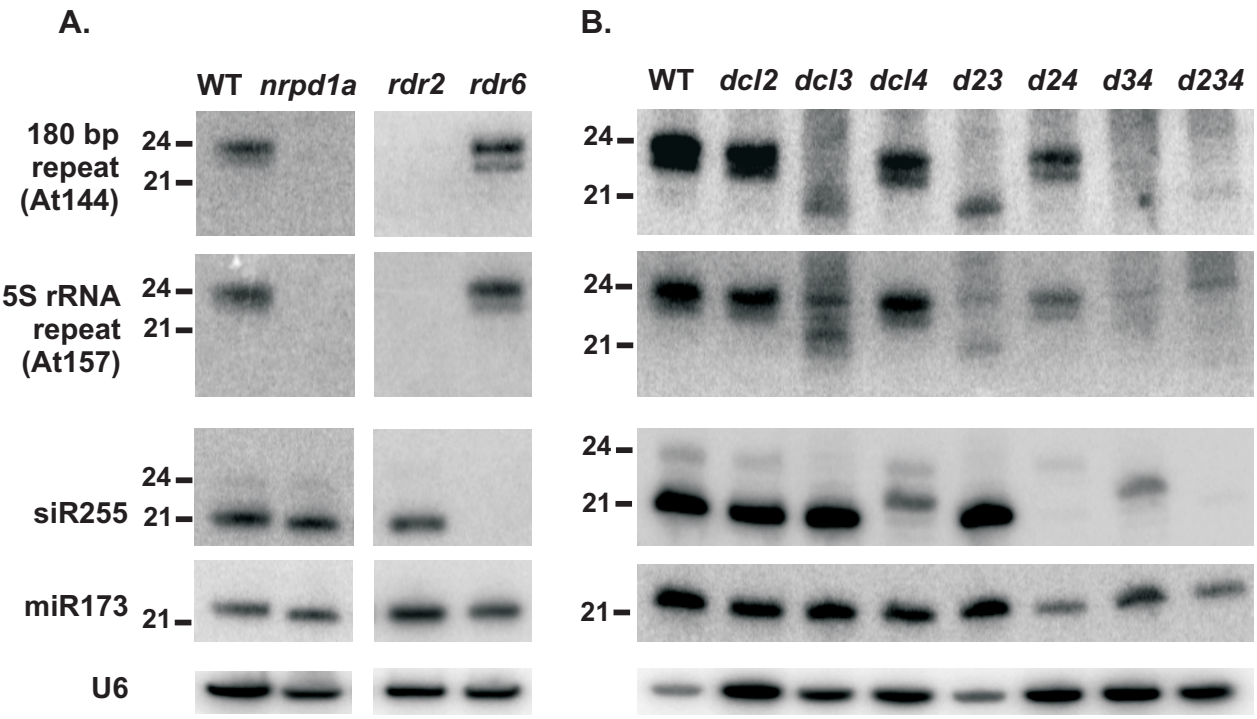


Figure 5. Tandem repeat-derived siRNA biogenesis in RNA silencing mutants. Low molecular weight RNA blot hybridization using RNA from leaves of WT plants and mutants deficient for RNA silencing. Synthetic 21 and 24 nt RNA oligos were used as size markers. Membranes were successively hybridized with DNA oligonucleotide probes for endogenous smRNA species (Appendix, **Table A1**). U6 snRNA signal is an RNA loading control. **(A)** *nrpd1a*, a null mutant for the gene encoding the largest subunit of Pol IVa; *rdr2* and *rdr6,* mutants deficient for the putative RdRPs required in distinct RNA silencing pathways. **(B)** *DCL*-deficent mutants: *dcl2*, *dcl3*, *dcl4*; double mutants *dcl2 dcl3* (*d23*), *dcl2 dcl4* (*d24)*, *dcl3 dcl4* (*d34*); and triple mutant *dcl2 dcl3 dcl4* (*d234*).

dcl-mutant combinations by crossing *dcl2* with *dcl3*, and then crossing the resultinghomozygous *dcl2 dcl3* (*d23*) with *dcl4*. The double mutants *dcl2 dcl4* (*d24*) and *dcl3 dcl4* (*d34*) and triple mutant *dcl2 dcl3 dcl4* (*d234*) were obtained from the F2, triple mutant segregating population (Appendix, **Figure A1**).

RNA blot hybridization was used to investigate how At144 expression depends on DCL proteins (**Figure 5B**). An aberrant 21 nt band observed in *dcl3* was absent in the double mutant *d34*. Furthermore, the faint 22 nt signal in *dcl3* was absent in the double mutant *d23*. The triple mutant *d234* showed no accumulation of siRNAs derived from 180 bp satellite repeats (**Figure 5B**). This suggests that each major size-class of At144-like siRNAs is the product of a particular DCL: specifically, DCL4 produces 21 nt, DCL2 produces 22 nt and DCL3 produces ~23-24 nt siRNAs. Similar conclusions can be drawn for siR255, which was cloned as a 21 nt siRNA in wild type, but accumulates as larger siRNA species in *dcl4*-mutant backgrounds (**Figure 5B**). Although the 5S rDNA-derived siRNA, At157, showed expression patterns similar to those of satellite repeat-derived siRNAs, a weak At157 signal was still detectable in *d234*. The genetic dependencies described above also applied to the antisense strands of At144 and At157 and to other repeat-derived siRNAs, such as siRNA732 and siRNA1003 (data not shown).

5.3. 5S rDNA methylation was not reduced in *dcl*-mutants

The accumulation of repeat-derived siRNAs was dramatically reduced in Pol IV pathway and *DCL*-deficient mutants. If tandem repeat methylation depends on siRNA products generated by the Pol IV pathway, then cytosine methylation in these repeats should correlate with variation in siRNA expression. I analyzed DNA methylation in 5S rDNA and 180 bp tandem repeats using the methylation sensitive restriction enzyme *Hpa* II. This enzyme normally recognizes CCGG sites within DNA, but fails to cut them when either cytosine residue is methylated – i.e., methylation in either CpG or CpNpG sequence contexts. 5S rDNA repeat units contain two closely spaced *Hpa* II sites (**Figure 4D**), whereas 180 bp repeat units contain one such site (Martinez-Zapater et al., 1986; Vongs et al., 1993).

Genomic DNA was extracted from the mutant panel, digested with *Hpa* II and analyzed by DNA blot hybridization using internally-labeled DNA probes (Appendix, **Figure A3**). The 5S rDNA arrays detected in wild-type samples showed minimal digestion, indicating that the *Hpa* II sites were highly methylated (**Figure 6A**). In contrast, *met1* and *ddm1*, which are both globally deficient for CpG-site methylation (Vongs et al., 1993; Saze et al., 2003), showed a ladder of bands at 500 bp intervals. Twin bands observed for these

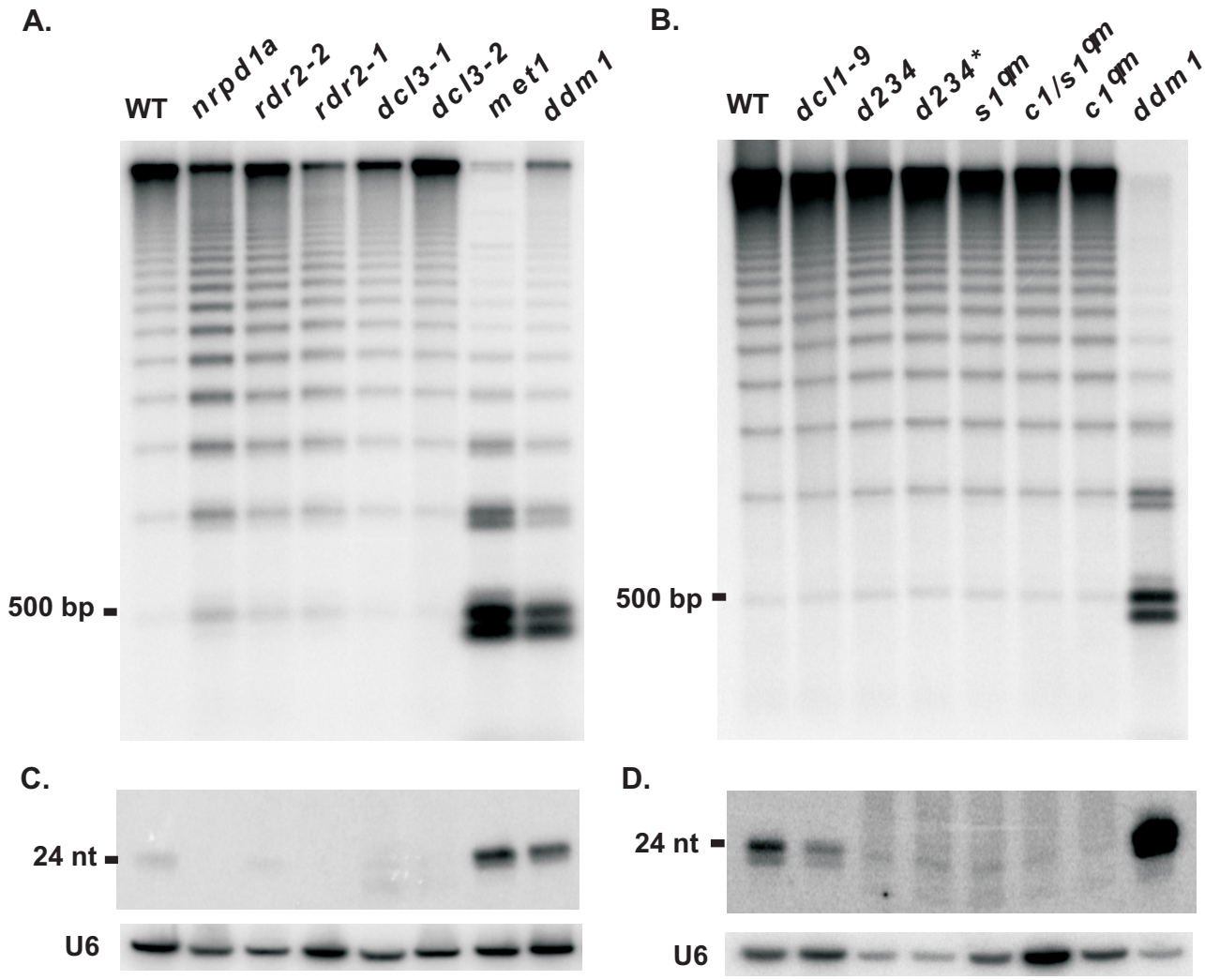


Figure 6. 5S rDNA methylation and siRNA accumulation in RNA silencing mutants. DNA blot hybridization analysis of cytosine methylation using *Hpa* II, a methylation-sensitive restriction enzyme. DNA isolated from the leaves of ca. 4-week-old plants was used. Membranes were hybridized with a probe representing the 5S rDNA repeats (Appendix, **Figure A3**). **(A)** wild-type (WT), *nrpd1a*, two alleles each of *rdr2* and *dcl3*, and mutants deficient in CpG-methylation, *met1* and *ddm1*. **(B)** WT, *dcl1-9*, triple *dcl*-mutant *d234*, another *d234** from a segregating population, quadruple *dcl*-mutants -- *dcl1-8 dcl2 dcl3 dcl4* ($s1^{qm}$), *dcl1-8/dcl1-9 dcl2 dcl3 dcl4* ($s1/c1^{qm}$), *dcl1-9 dcl2 dcl3 dcl4* ($c1^{qm}$) -- and *ddm1*. **(C and D)** smRNA blot hybridization from materials in **Panels A** and **B**, respectively, probed for 5S rDNA repeat-derived siRNAs (At157).

hypomethylated samples are likely due to polymorphism for the dual *Hpa* II sites within repeat units. Of mutants deficient for the Pol IV pathway, *nrpd1a* and two well-characterized *RDR2*-deficient mutants (*rdr2-1* and *rdr2-2*) showed some hypomethylation (**Figure 6A**). Surprisingly, 5S rDNA methylation in two *DCL3*-deficient mutants (*dcl3-1* and *dcl3-2*) was indistinguishable from wild type using the *Hpa* II assay. I propagated two *dcl3-1* lineages for three generations by self-fertilization and did not observe transgenerational variation in 5S rDNA methylation (data not shown). Methylation of 180 bp satellite repeats, which was also analyzed, was not affected in any of the Pol IV pathway mutants, including *nrpd1a* (data not shown). This latter result has also been shown for other subunits of Pol IV complexes (Kanno et al., 2005b). Taken together, these data suggest that Pol IV-dependent (e.g., 5S rDNA) and Pol IV-independent (e.g., 180 bp) pathways exist for methylating DNA in genomic tandem repeats. I focused on the less well understood, Pol IV dependent pathway, using 5S rDNA repeats as a model.

Arabidopsis DCLs are partially redundant for the biogenesis of 5S rDNA-derived siRNAs (**Figure 5B**). Thus, one explanation for the wild-type levels of 5S rDNA methylation observed in *dcl3* mutants would be that alternate DCLs can compensate for DCL3 deficiency. Unlike the T-DNA insertion mutants I obtained for *DCL2*, *DCL3* and *DCL4*, the available DCL1-deficient strains, *dcl1-8* and *dcl1-9*, are hypomorphic (Schauer et al., 2002; Xie et al., 2005; Bouche et al., 2006). To test the compensation hypothesis, I compared the methylation of DNA from the most severe, viable mutant of DCL1, *dcl1-9*, and from genetic combinations of mutants for all four *DCL* genes (**Figure 6B**). Included in this experiment were two different lines of *d234* triple mutant and three different lines of quadruple *dcl*-mutant. Two quadruple mutants were generated by crossing either *dcl1-8* or *dcl1-9* mutants to *d234*. Finally, the *dcl1-8* based quadruple mutant ($s1^{qm}$) and the *dcl1-9* based quadruple mutant ($c1^{qm}$) were also crossed with one another (Appendix, **Figure A1**). No combination of these *dcl*-mutants had a detectable effect on 5S rDNA methylation in the *Hpa* II sites tested (**Figure 6B**).

The pattern of accumulation of 5S rDNA-derived siRNAs in the *dcl*-mutants was similar to that found earlier, except that *dcl1-9* showed only ~50% of the wild-type level (**Figure 6D**). Although MET and DDM1 have distinct functions, they are both required for maintenance of CpG methylation in *Arabidopsis* (Brzeski and Jerzmanowski, 2003; Goll and Bestor, 2005). Interestingly, *met1* and *ddm1* mutants showed a striking over-accumulation of 5S rDNA-derived siRNAs (**Figure 6C**/**D**). This would argue that CpG-methylation (or an associated chromatin state) controls the titer of 5S rDNA-derived siRNAs.

5.4. Summary

Sequence data from siRNA cloning experiments identified siRNAs corresponding to *Arabidopsis* 180 bp and 5S rDNA tandem repeats. Accumulation of these siRNAs depended on *NRPD1a*, *RDR2*, and *DCL3*, which all encode proteins of the Pol IV pathway. In the absence of DCL3, however, alternate DCLs appear to produce aberrant repeat-derived siRNAs of different size classes. 5S rDNA methylation was somewhat reduced in *nrpd1a* and *rdr2* mutants, but unaffected in *dcl3* or quadruple mutants deficient for all four DCLs. In contrast, methylation of 180 bp satellite repeats did not require any key components of the Pol IV pathway. 5S rDNA-derived siRNAs overaccumulated in the mutants *ddm1* and *met1*, suggesting that CpG-methylation limits the biogenesis of these siRNAs.

6. Genetic interactions between DDM1 and Pol IV pathway mutants

Chapter 5 showed that deficiencies for DDM1, MET1 or upstream steps in the Pol IV pathway result in hypomethylation of 5S rDNA repeats. Studies with other DNA repeats have shown that deficiencies in DNA methylation can be persistent in successive sexual generations (Kankel et al., 2003; Henderson and Jacobsen, 2007). For example, Lippman *et al.* (2003) showed that effects of DDM1-deficiency on transposable element methylation persisted when homozygous *ddm1* plants were outcrossed to wild type. I performed a similar test on mutants deficient for the Pol IV pathway to evaluate the recovery of 5S rDNA methylation in F1-heterozygotes. Although Pol IV and DDM1 are both clearly implicated in 5S rDNA methylation, it is not known whether they act in the same pathway. To address this question, I generated double mutants deficient for both DDM1 and different steps in the Pol IV pathway, and tested for epistasy in terms of 5S rDNA hypomethylation. The crosses made and pedigrees of lines tested are summarized in **Figure 8**.

6.1. Stability and epistasis analysis of 5S rDNA hypomethylation

To evaluate the relative stability of hypomethylation observed in mutants from Chapter 5, I crossed homozygous *nrpd1a*, *rdr2-1* and *ddm1* to the Col-0 wild type. The resulting heterozygous *NRPD1a* (+/-) and *RDR2* (+/-) plants showed wild-type levels of 5S rDNA methylation, indicating that hypomethylation due to NRPD1a or RDR2 deficiencies did not persist (**Figure 7A**). In contrast, *DDM1* (+/-) plants showed methylation levels intermediate between homozygous *ddm1* (-/-) and the wild type, indicating that hypomethylation due to DDM1-deficiency shows some degree of persistence (**Figure 7A**). Whereas heterozygous Pol IV-pathway mutants accumulated wild-type levels of 5S rDNA-derived siRNAs, the overexpression of siRNAs characteristic of *ddm1* persisted in *DDM1* (+/-) (**Figure 7B**). Therefore, persistent DNA hypomethylation in *DDM1* (+/-) correlates with persistent overaccumulation of 5S rDNA-derived siRNAs. This suggests that in wild-type plants, DNA methylation rather than the DDM1 protein itself, represses production of these siRNAs, potentially at the level of their transcript precursors.

To test for epistasy between the Pol IV pathway and maintenance methylation via DDM1, I crossed *nrpd1a*, *rdr2*, and *dcl3* with *ddm1*. F2 plants homozygous for both *ddm1* and these Pol IV pathway mutants were then self-fertilized for two generations and compared to *ddm1* and wild-type lines propagated in parallel (**Figure 8**).

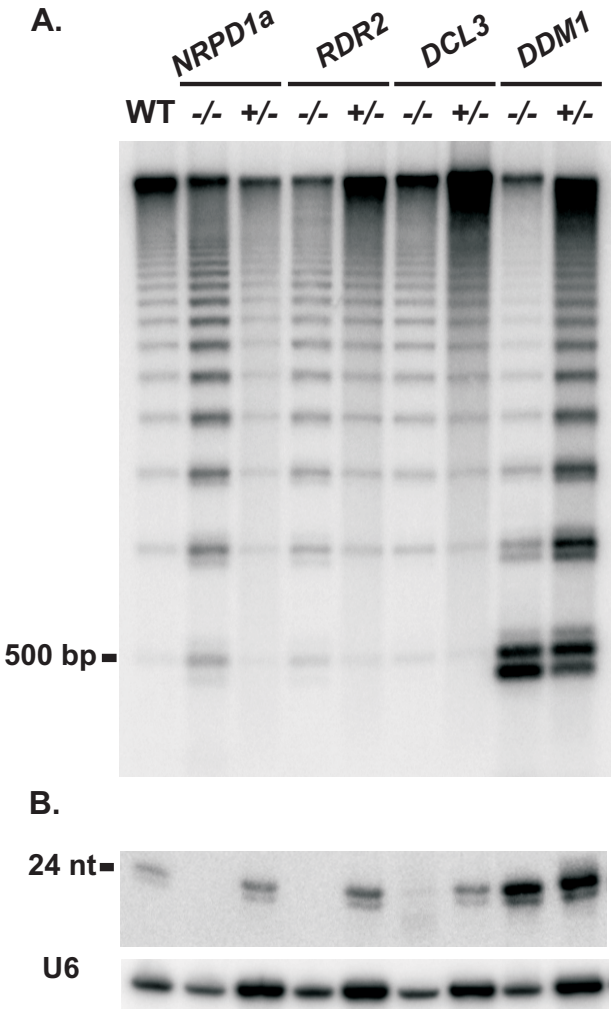


Figure 7. 5S rDNA methylation and siRNA accumulation in outcrosses of Pol IV pathway and DDM1-deficient mutants. (A) DNA blot hybridization analysis of cytosine methylation was performed using *Hpa* II, a methylation-sensitive restriction enzyme. DNA from *nrpd1a*, *rdr2*, *dcl3* and *ddm1* (-/-) mutant lines were compared to the F1 outcrosses (+/-) of these lines to the wild type (WT). The membrane was hybridized with a probe representing *Arabidopsis* 5S rDNA repeats (Appendix, **Figure A3**). **(B)** Blot hybridization of low molecular weight RNA from the samples analyzed in **Panel A**, probed for 5S rDNA-derived siRNAs (At157).

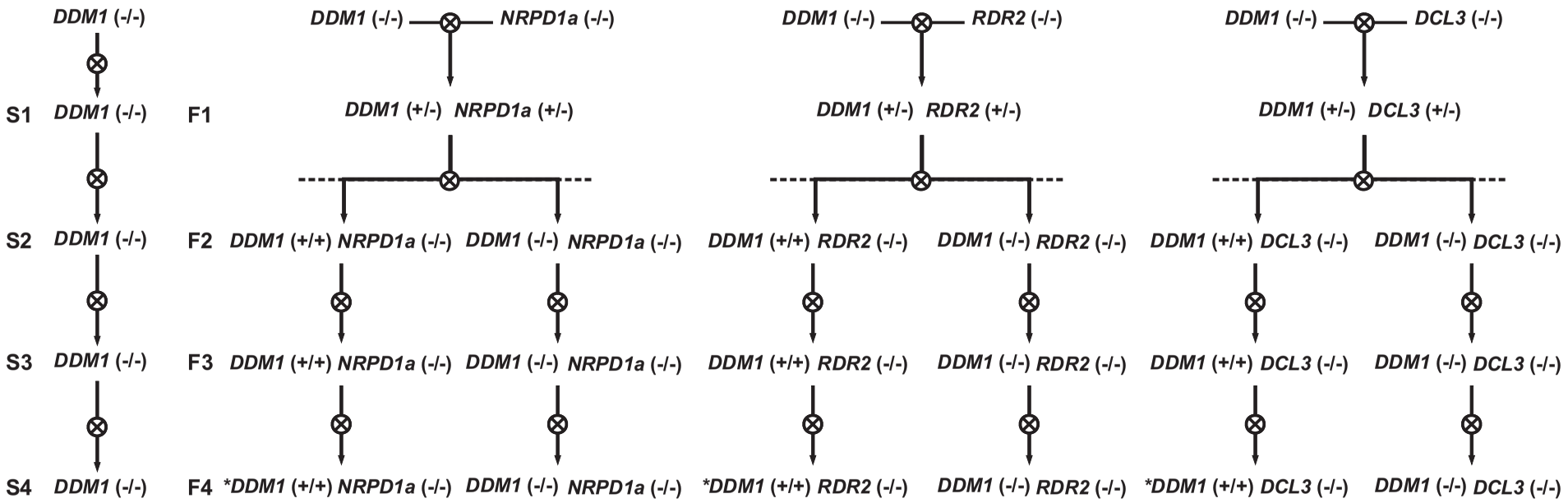


Figure 8. The protocol used to obtain double mutant combinations deficent for *DDM1* and *NRPD1a*, *RDR2* or *DCL3*. Only relevant progeny are shown. The genotypes are indicated as wild type (+/+), heterozygote (+/-), and homozygous recessive (-/-) without regard for linkage of genes. **DDM1* (+/+) lines were used as sibling controls for the double mutants.

Sibling *DDM1* (+/+) plants in the F2 generation, which were homozygous for mutations in the Pol IV pathway, were designated *DDM1** and used as additional controls. DNA blot hybridization analysis showed that 5S rDNA repeats in each double mutant line were less methylated than in either *ddm1* alone or their *DDM1** siblings (**Figure 9A**). Therefore, with respect to methylation of the *Hpa II* sites tested, none of the Pol IV pathway-deficient mutants are epistatic to *ddm1*. Analysis of the corresponding 5S rDNA-derived siRNAs showed, as expected, that their signal was abolished in *nrpd1a* and *rdr2* mutant lines. Interestingly, the signal from the aberrant pattern of siRNAs observed in **DDM1 dcl3*, is enhanced in *ddm1 dcl3* double mutant plants (**Figure 9B**). This is consistent with the hypothesis that DDM1 limits the biogenesis of dsRNA precursors, which are processed by DCLs downstream of Pol IVa and RDR2. It is unlikely that enhanced DCL3 activity or 24 nt siRNA stability explain the overexpression effect, since it also occurred in *dcl3 ddm1*.

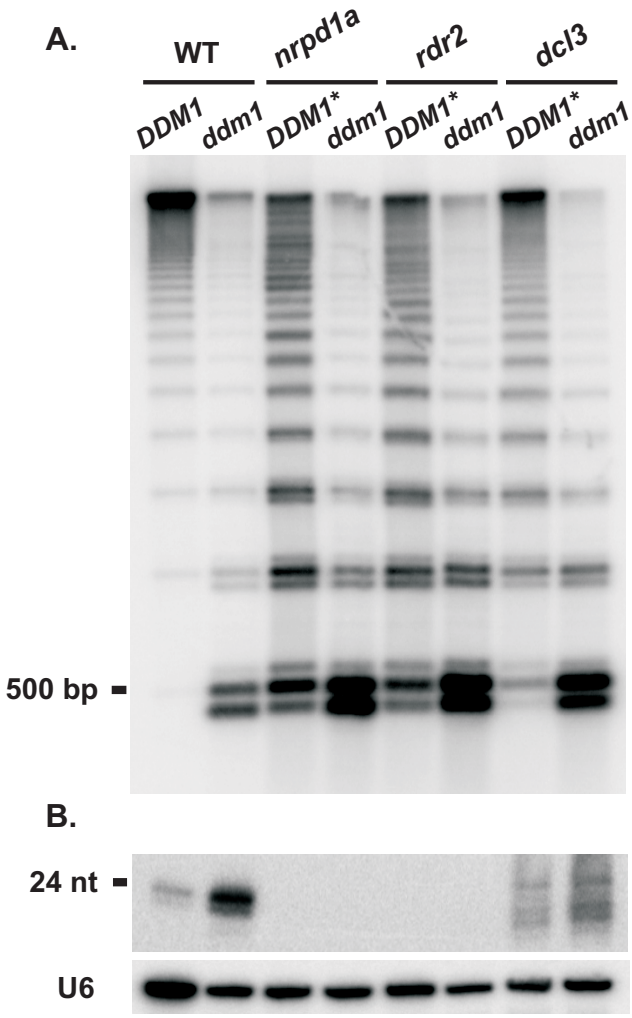


Figure 9. Epistasis analysis of 5S rDNA hypomethylation. (A) DNA blot hybridization analysis of cytosine methylation was performed using *Hpa* II, a methylation-sensitive restriction enzyme. DNA isolated from leaves of double mutants *ddm1 nrpd1a*, *ddm1 rdr2*, and *ddm1 dcl3* were compared to *DDM1** sibling controls and the *ddm1* parent line -- the F4 generation of lines described in **Figure 8**. The membrane was hybridized with a probe representing *Arabidopsis* 5S rDNA repeats (Appendix, **Figure A3**). **(B)** Blot hybridization of low molecular weight RNA from the samples analyzed in **Panel A**, probed for 5S rDNA repeat-derived siRNAs (At157).

6.2. Phenotypes of *ddm1* hybrids with mutants deficient in the Pol IV pathway

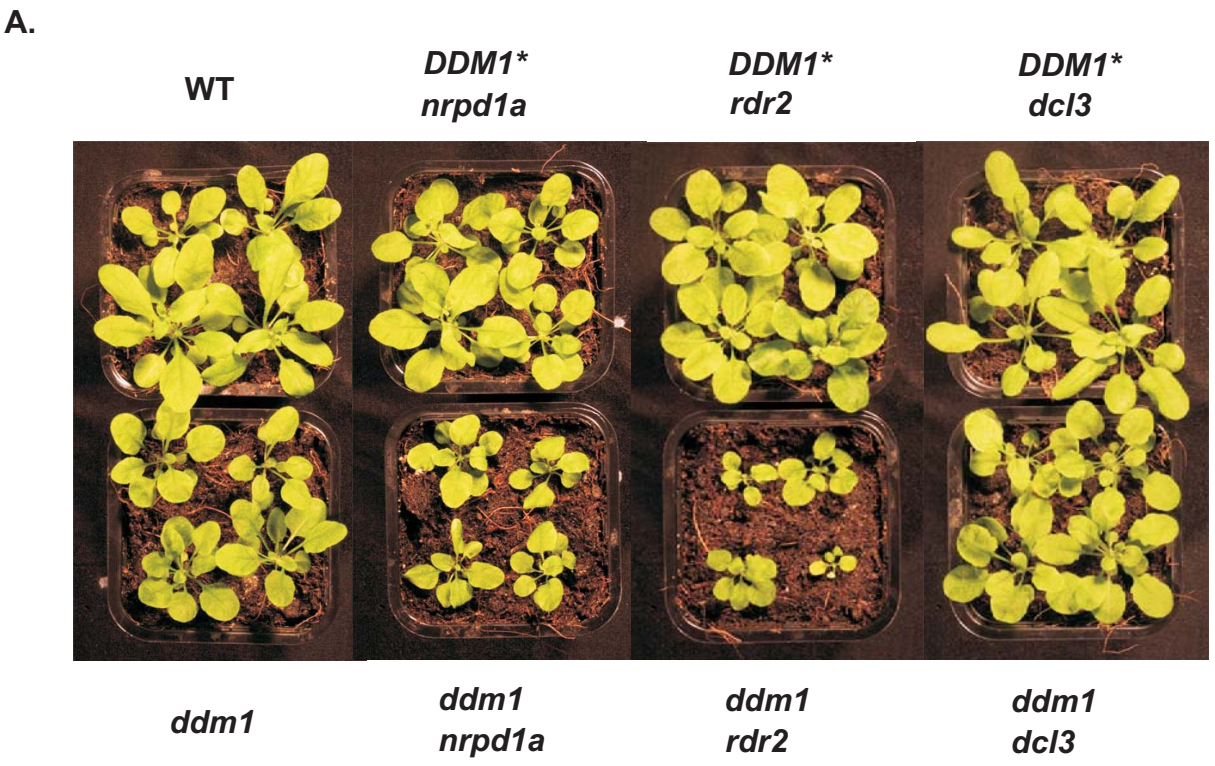
Although the single mutants *nrpd1a*, *rdr2* and *dcl3* show hypomethylation of some DNA repeats, their effects on growth and development are rather mild. To date, the only significant phenotype described for the *nrpd1a*, *rdr2* and *dcl3* mutants, is a delayed onset of flowering (Chan et al., 2004). The *ddm1-2* mutant used in this dissertation shows developmental phenotypes after successive self-pollination (Kakutani et al., 1996), but no severe manifestations of these phenotypes were observed in the *ddm1-2* generation that I used for crosses. To find out if novel or enhanced phenotypes appear in the double mutants, I compared F4 generations of plants from my DNA methylation analysis (**Figure 8**) under the same growth conditions and in the same experiment.

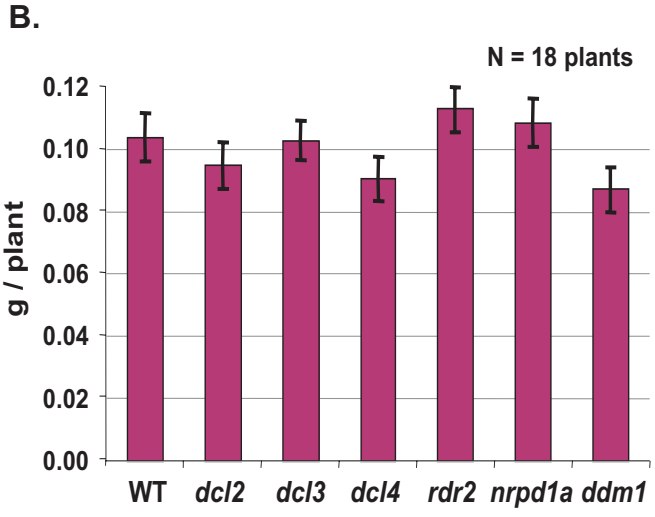
None of the single mutants showed striking developmental abnormalities (**Figure 10A** and data not shown). However, the aerial parts of *dcl4* and *ddm1* plants weighed slightly less than those of the wild type (**Figure 10B**). The appearance and average weight of *ddm1 dcl3* double mutants were indistinguishable from that of *ddm1* (**Figures 10A**/**C**). In contrast *ddm1 nrpd1a* and *ddm1 rdr2* were small in appearance and weighed ~50% of the *ddm1* or sibling *DDM1** (+/+) lines (**Figure 10C**).

6.3. Developmental phenotypes of quadruple *ddm1 d234* mutants

In Chapter 5, I found redundancy and compensation in the DCL family for the production of siRNAs. Here, I sought to test whether such redundancy might help explain the similar growth phenotype of *ddm1 dcl3* compared to *ddm1*. I introduced multiple *dcl*-mutants into the *ddm1* background by successive genetic crosses. This resulted in a quadruple mutant *ddm1 d234*, which is deficient for *DDM1*, *DCL2*, *DCL3* and *DCL4* (pedigree, **Figure A2**). Another genotype, *DDM1* d234*, which is a triple mutant *d234* (-/-) but wild type for *DDM1* (+/+), was also obtained from these crosses. Progeny from these *ddm1 d234* and *DDM1* d234* parent plants were grown under comparable conditions along with control lines.

As in the previous experiment, the aerial portion of each plant was weighed and the averages calculated. Although the effect is more subtle than for *ddm1 nrpd1a* and *ddm1 rdr2* lines, *ddm1 d234* is clearly less vigorous than *DDM1* d234* or any of the other controls (**Figure 11A/B**). This supports my proposition that growth deficiency occurs when a key step in the Pol IV pathway is impaired in a *DDM1*-deficient background. The normal growth phenotype of *ddm1 dcl3* might thus be explained by compensation from DCL2 and DCL4.





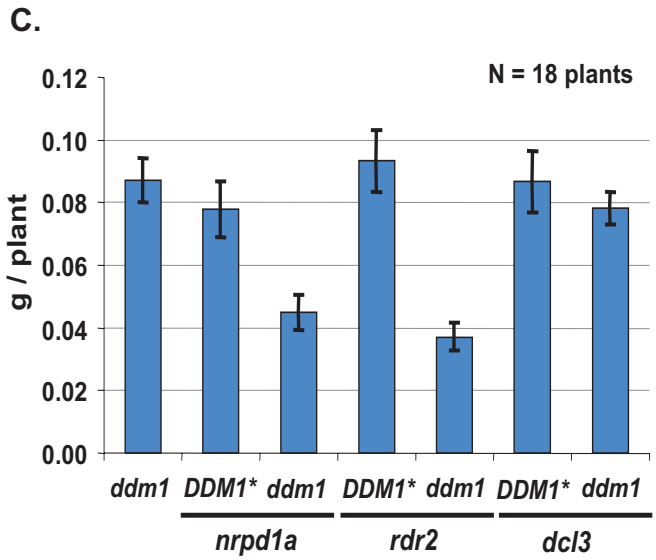
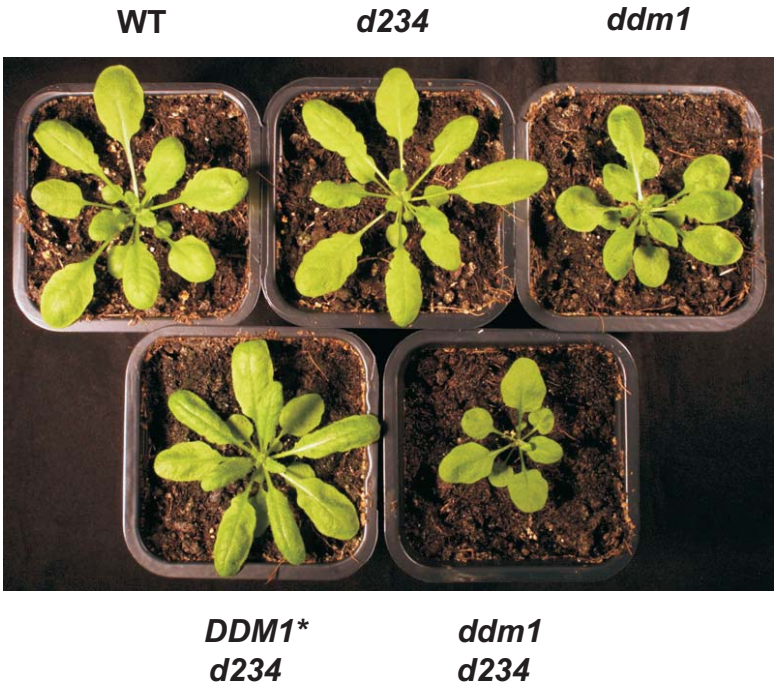


Figure 10. DDM1 and Pol IV-pathway deficient mutant phenotypes. (A) Rosette stage plants from wild type (WT), *ddm1* and double mutant combinations of *ddm1* with three Pol IV pathway deficient mutants (*nrpd1a*, *rdr2* and *dcl3*). *DDM1** (+/+) controls segregated from these crosses and were propagated to the F4 generation (see **Figure 8**). All strains were grown at the same time and under equivalent illumination. **(B)** Fresh weight measured for WT, *dcl*, *rdr2*, *nrpd1a* and *ddm1* mutants. The areal portions of 18 plants were individually weighed and the mean calculated. **(C)** Fresh weight of plants from the *ddm1* parent line were compared to double mutant combinations of Pol IV pathway mutants with *ddm1*. *DDM1** individuals homozygous for each RNA silencing deficient mutant are controls, as described above. The bars show standard errors calculated for the mean.

A.



B.

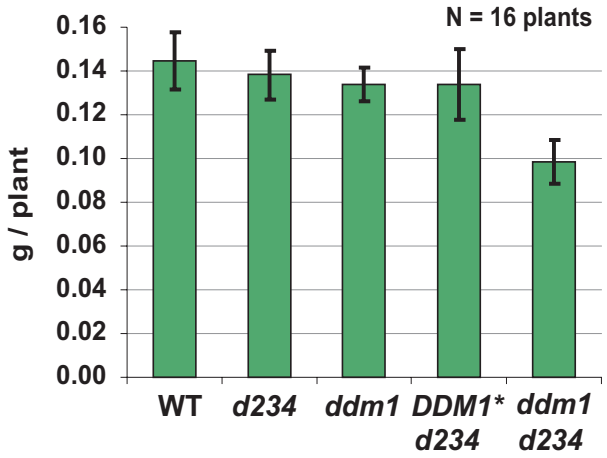


Figure 11. Growth deficiency in *ddm1 d234* quadruple mutant plants. (A) Rosette stage plants were analyzed from the following *Arabidopsis* strains: wild-type (WT), triple mutant *d234, ddm1,* and F3 progeny of a cross between *ddm1* and *dcl* mutants, including the triple mutant *DDM1* d234* and the quadruple mutant *ddm1 d234. DDM1* d234* is a sibling control line that resulted from the genetic crosses (Appendix, **Figure A2**). All plants were grown simultaneously and under equivalent illumination. **(B)** Fresh weight analysis of the plants shown in **Panel A**. The bars represent standard errors calculated from the mean.

6.4. Summary

Here, I presented evidence that DDM1 and the Pol IV pathway play distinct but intertwined roles in 5S rDNA methylation. Mutants deficient for the Pol IV pathway, *nrpd1a* and *rdr2*, showed wild-type levels of 5S rDNA methylation after outcrossing. In contrast, 5S rDNA hypomethylation induced by *ddm1* persisted after outcrossing. Hypomethylation in heterozygous *DDM1* (+/-) plants correlated with persistent overaccumulation of 5S rDNA-derived siRNAs, which was not specific to 24 nt species: alternate size-classes of siRNAs overaccumulated in the *ddm1 dcl3* double mutant compared to *dcl3*. These results are consistent with the hypothesis that CpG-methylation, maintained in part by DDM1, limits expression of precursors for 5S rDNA-derived siRNAs. Furthermore, I showed that double mutants deficient for DDM1 and the Pol IV pathway have less 5S rDNA methylation than corresponding single mutants (*ddm1*, *nrpd1a*, *rdr2*, or *dcl3*). This suggests that Pol IV-mediated and DDM1-mediated methylation of 5S rDNA are not part of the same pathway. Interestingly, the double mutants *ddm1 nrpd1a* and *ddm1 rdr2*, and quadruple mutant *ddm1 d234* showed growth deficiency not apparent in the *ddm1* parent lineage.

7. Biogenesis of siRNAs in geminivirus-infected *Arabidopsis*

In the previous two chapters, I showed that different DCLs mediate production of specific size classes of endogenous repeat-derived siRNAs, whose precursors are apparently Pol IV and RDR2-dependent. The following chapters examine viral siRNA biogenesis during infection of *Arabidopsis* with DNA viruses, which generate circular minichromosomes in the plant nucleus. The first virus, a geminivirus (covered in this chapter), replicates by a rolling circle mechanism (Hanley-Bowdoin et al., 2000). The second virus, a caulimovirus (covered in Chapter 8), replicates via nuclear transcription and then cytoplasmic reverse transcription of its genome (Pfeiffer and Hohn, 1983). Both viruses express transcripts that are exported and translated in the cytoplasm, but neither requires long dsRNA for its replicative cycle or encodes an RdRP activity (Hull, 2004).

RNA silencing can be induced by various geminiviruses (Kjemtrup et al., 1998; Turnage et al., 2002), which have also been shown to encode suppressors of silencing (Voinnet et al., 1999; Trinks et al., 2005). However, the mechanisms that produce dsRNA and siRNAs derived from geminiviruses are not well understood. Can known endogenous RNA silencing pathways, such as the abovementioned Pol IV pathway, account for dsRNA production during infection? What DCLs mediate viral siRNA biogenesis during infection? To answer these questions, I infected my panel of RNA silencing-deficient mutants with a model geminivirus and analyzed the resulting silencing phenomena.

7.1. Effects of RDR and DCL-deficiency on CaLCuV-mediated VIGS

For all experiments in this chapter, I used a recombinant form of the geminivirus *Cabbage Leaf Curl Virus* (CaLCuV), which infects *Arabidopsis* and other plants of the *Brassicaceae* family (Hill et al., 1998). The modified virus carries a 360 bp fragment of the *Arabidopsis Chlorata I* (*CHLI*) gene, and is therefore called CaLCuV::*ChlI* (Turnage et al., 2002) (**Figure 12B**). CaLCuV::*ChlI* has the advantage that one can simultaneously monitor effects on virus replication, symptomology and VIGS (Muangsan et al., 2004). CaLCuV::*ChlI* infection triggers silencing of the endogenous *CHLI* gene, which encodes a magnesium chelatase required for chlorophyll biosynthesis, and, hence, results in a bleached “chlorata” phenotype of upper rosette leaves of the plant (**Figure 12C**).

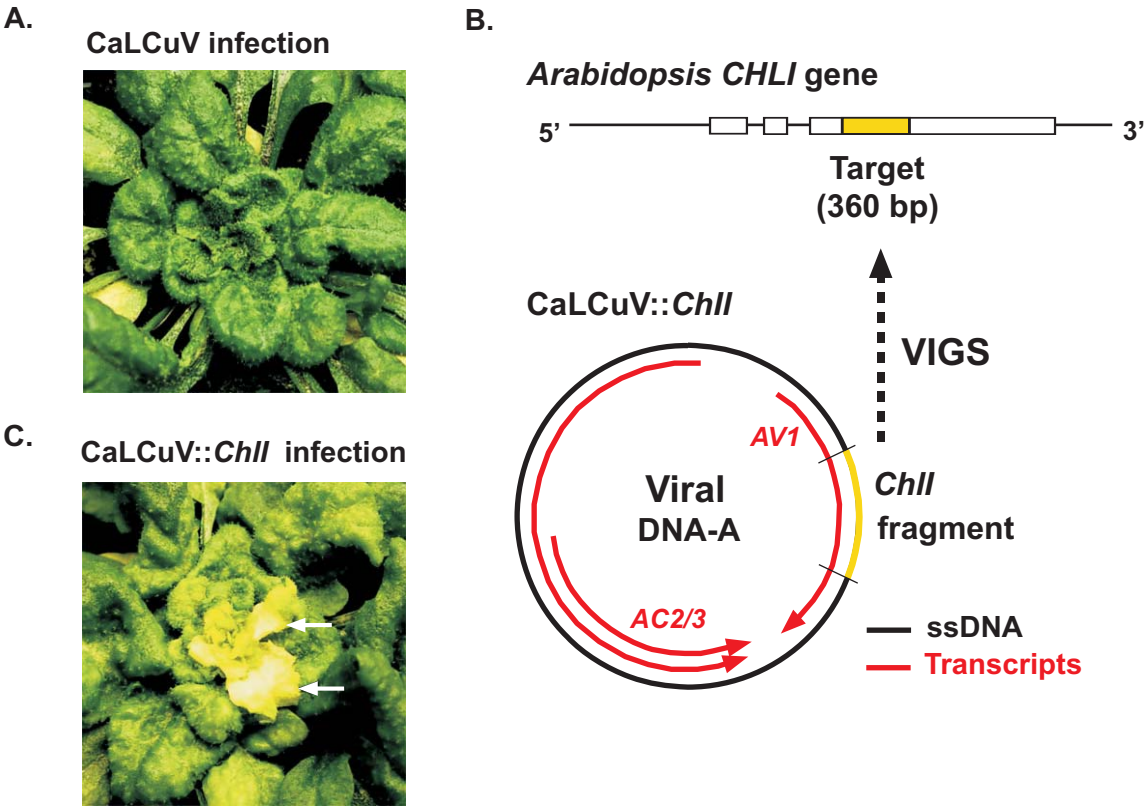


Figure 12. Virus-induced gene silencing (VIGS) targeting a gene in *Arabidopsis*. *Cabbage Leaf Curl Virus* (CaLCuV) infects plants in the *Brassicaceae* family. The viral genome is composed of two circular ssDNA molecules. **(A)** For inoculation, *Arabidopsis* was bombarded with 1 µm-dia. gold particles coated with infectious plasmids containing CaLCuV genomic DNA-A and DNA-B. This leads to systemic infection associated with crumpled leaf symptoms. **(B)** Inserting a 360 bp fragment of *CHLI,* a nuclear *Arabidopsis* gene required for chlorophyll biosynthesis, into CaLCuV DNA-A allows efficient VIGS of the target gene upon viral infection. **(C)** This construct, CaLCuV::*ChlI*, triggers a bleached "chlorata" phenotype in emerging rosette leaves (arrows), but not in lower leaves.

CaLCuV infection and VIGS were studied in the deficiency mutants *dcl2*, *dcl3*, *dcl4*, *rdr2* and *rdr6*. None of these mutants differed markedly from the wild type in the development or severity of CaLCuV symptoms (**Figure 13A**). Infection was also monitored by measuring the accumulation of CaLCuV genomic DNA in plants 28 days post-inoculation (dpi), when disease symptoms were well-developed. **Figure 13B** shows that, while viral DNA levels were slightly increased in *dcl4* and *rdr6*, the levels in *dcl2*, *dcl3*, and *rdr2* were comparable to wild type. The CaLCuV *AC2*/*AC3* transcript, which is transcribed from viral DNA-A (**Figure 12B**), was detected by RNA blot hybridization: its level was comparable in wild type, *dcl2*, *dcl3*, and *rdr2*, but elevated in *dcl4* and *rdr6* (**Figure 13C**).

In contrast to the modest effects of the mutants on virus infection, infected *dcl4* and *rdr6* plants exhibited almost no chlorata phenotype, while *dcl2*, *dcl3* and *rdr2* plants showed a chlorata phenotype limited to upper rosette leaves, like in the wild type (**Figure 13A**). The differences in extent of silencing were also observed by measurements of mRNA. While *CHLI* mRNA accumulation was reduced ca. 5-fold by infection of wild-type, *dcl2*, *dcl3* of *rdr2* plants, the reduction was only 2.5-fold in *dcl4* and *rdr6* plants (**Figure 13D**). Thus, while these mutants did not strongly affect CaLCuV::*ChlI* infection, *dcl4* and *rdr6* showed reduced efficiency for the establishment and/or maintenance of VIGS and the chlorata phenotype. An additional mutant, *hen1*, also showed a reduced chlorata phenotype compared to the wild type when infected with CaLCuV::*ChlI* (data not shown).

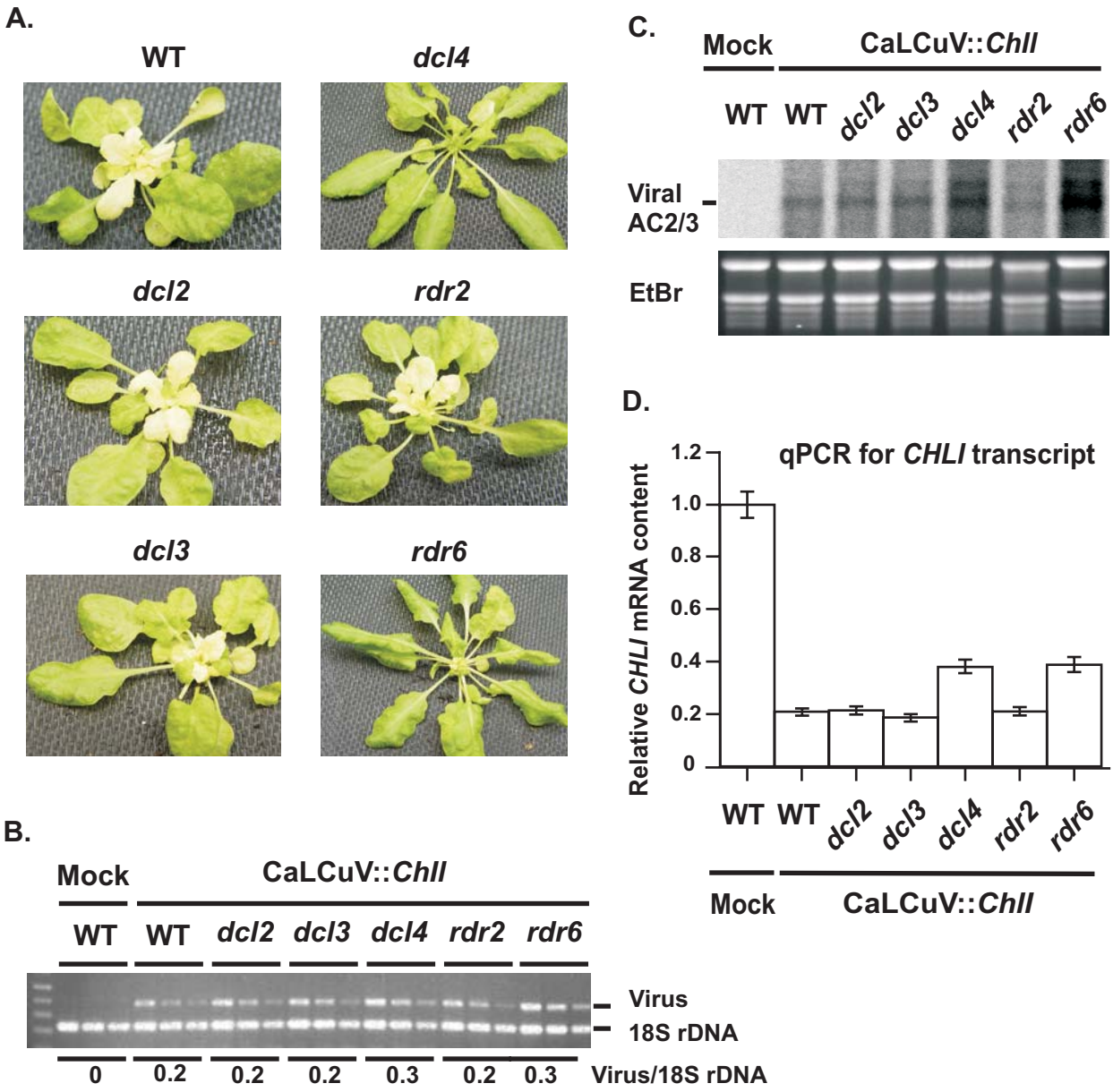


Figure 13. Analysis of VIGS triggered by CaLCuV::*ChlI* in *dcl* and *rdr*-mutants. (A) Wild type (WT), *dcl* and *rdr*-mutant plants 28 day post-inoculation (dpi) with CaLCuV::*ChlI*. **(B)** Viral titers were measured by semi-quantitative PCR on serial dilutions (5-fold each) of DNA from pools of plants. 18S rDNA amplification was used as an internal loading control. **(C)** RNA blot hybridization performed on column-purified total RNA (8 µg/lane). The membrane was hybridized with a random-labeled DNA probe for the CaLCuV-derived *AC2/3* transcript (see **Figure 12B**). Ethidium bromide (EtBr) staining is an RNA loading and quality control. **(D)** Quantitative RT-PCR (qPCR) was performed on cDNA synthesized from the same RNA as used for **Panel C**. RNA values are expressed as the ratio relative to WT mock infected, normalized for the *ACT2* mRNA standard.

7.2. Biogenesis of CaLCuV::*ChlI* siRNAs in mutants deficient in the Pol IV, *trans*-acting siRNA and miRNA pathways

To determine whether *RDR2*, *RDR6*, *NRPD1a* and *AGO4*, which encode components of endogenous RNA silencing pathways (**Figure 2B/C**), are required for viral siRNA production, I infected mutants deficient for these genes with CaLCuV::*ChlI* and measured accumulation of viral siRNAs by RNA blot hybridization. Probes for the viral *ChlI* insert detected siRNAs that migrated as 21, 22 and 24 nt bands in RNA from infected plants, but not in RNA from uninfected plants (**Figure 14A/B**). The pattern of viral siRNAs obtained with infected wild-type, *rdr2* and *rdr6* plants was qualitatively identical (**Figure 14A**). The more intense signal of *rdr6* bands correlated with enhanced viral transcript accumulation in that sample (**Figure 13C**). Similar patterns of viral siRNAs were also obtained for the different *Arabidopsis* genotypes using an antisense probe for *ChlI* (**Figure 14A**), and probes for two other regions of the CaLCuV::*ChlI* genome (data not shown). Next, I analyzed siRNAs from infected *nrpd1a* and *ago4* plants, mutants deficient for upstream and downstream steps in the Pol IV pathway, respectively. *nrpd1a* and *ago4* showed the same pattern of sense and antisense viral siRNA accumulation as the wild type (**Figure 14B**). Together, these results indicate that RDR2, RDR6, NRPD1a and AGO4 are not required for the biogenesis of siRNAs derived from CaLCuV::*ChlI*.

Non-viral smRNAs were used as controls to verify the functional deficiencies of *rdr2*, *rdr6*, *nrpd1a* and *ago4* (**Figure 14A/B**). As expected (Reinhart et al., 2002; Xie et al., 2004), accumulation of miR173 was not affected in any of these mutants. Further control hybridizations confirmed that accumulation of siR1003 was blocked in *rdr2* and *nrpd1a*, reduced in *ago4*, but unaffected in *rdr6*, whereas siR255 was blocked in *rdr6*, but not in *rdr2*, *nrpd1a* or *ago4* (Xie et al., 2004; Allen et al., 2005).

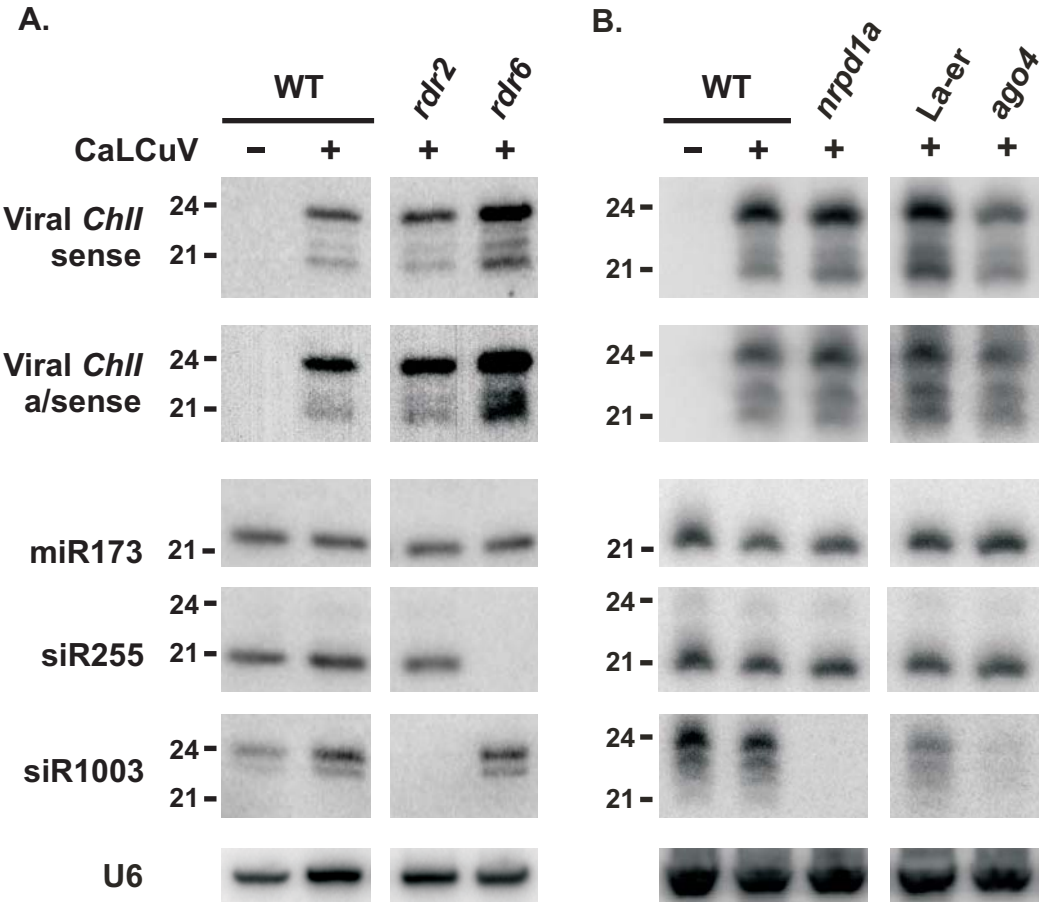


Figure 14. CaLCuV siRNA biogenesis in mutants deficient for the Pol IV and *trans*-acting siRNA pathways. Low molecular weight RNA was isolated from whole plants 28 dpi with CaLCuV::*ChlI* and analyzed by blot hybridization. Synthetic 21 and 24 nt RNA oligonucleotides were used as size markers. Membranes were successively hybridized with DNA oligo probes for viral regions and endogenous smRNAs (Appendix **Table A1**). **(A)** CaLCuV-derived siRNAs detected in mock (-) and infected (+) wild-type (WT), *rdr2* and *rdr6* plants. Endogenous smRNA controls were miR173 (*RDR*-independent), siR255 (*RDR6*-dependent) and siR1003 (*RDR2*-dependent).**(B)** *nrpd1a* and *ago4*, mutants deficient for the Pol IV pathway, were analyzed using the same probes as in **Panel A.** La-*er* is the wild type for *ago4.*

DCL1, HEN1 and HYL1 are required for the biogenesis or stabilization of *Arabidopsis* miRNAs (Reinhart et al., 2002; Vazquez et al., 2004a; Yu et al., 2005). To assess whether these components of the miRNA pathway are required for the biogenesis of CaLCuV::*ChlI* siRNAs, I infected the hypomorphic mutants *dcl1-8* and *dcl1-9*, and T-DNA insertion mutants *hen1* and *hyl1* with the virus. Relative to their respective wild-type reference ecotypes, the *dcl1-8*, *dcl1-9, and hyl1* mutations did not affect the size or abundance of viral siRNAs (**Figure 15A**). Sense and antisense siRNA accumulation, detected from the viral *ChlI* insert and viral *AC4* gene regions, were equivalent in this respect. As expected (Reinhart et al., 2002; Vazquez et al., 2004a), the 22 nt miR173 was not detected in *dcl1-8* or *dcl1-9*, and was only barely detected in *hyl1*.

The abundance of 21 nt viral siRNAs was reduced and additional ~23-25 nt siRNAs were detected in *hen1*. I confirmed this effect for miR173 (**Figure 15A**), which had been described earlier (Park et al., 2002; Li et al., 2005). In the case of miRNAs, these additional bands have been shown to result from 3'-oligouridylation (Li et al., 2005). HEN1 is a 2'-O-methyltransferase that blocks 3'-oligouridylation by transferring methyl groups to the 2'-hydroxyl group of the 3'-end of miRNAs (Yu et al., 2005; Yang et al., 2006). My analysis of CaLCuV-derived siRNAs suggested that their 3'-ends are also protected in wild-type plants and oligouridylated in *hen1*. CaLCuV-derived siRNAs were therefore tested for sensitivity to β-elimination, an assay that is diagnostic of RNA 3'-end modifications (Yu et al., 2005). The pattern of CaLCuV-derived siRNAs extracted from wild-type plants and subjected to β-elimination was consistent with protection of their 3'-ends by methylation (data not shown).

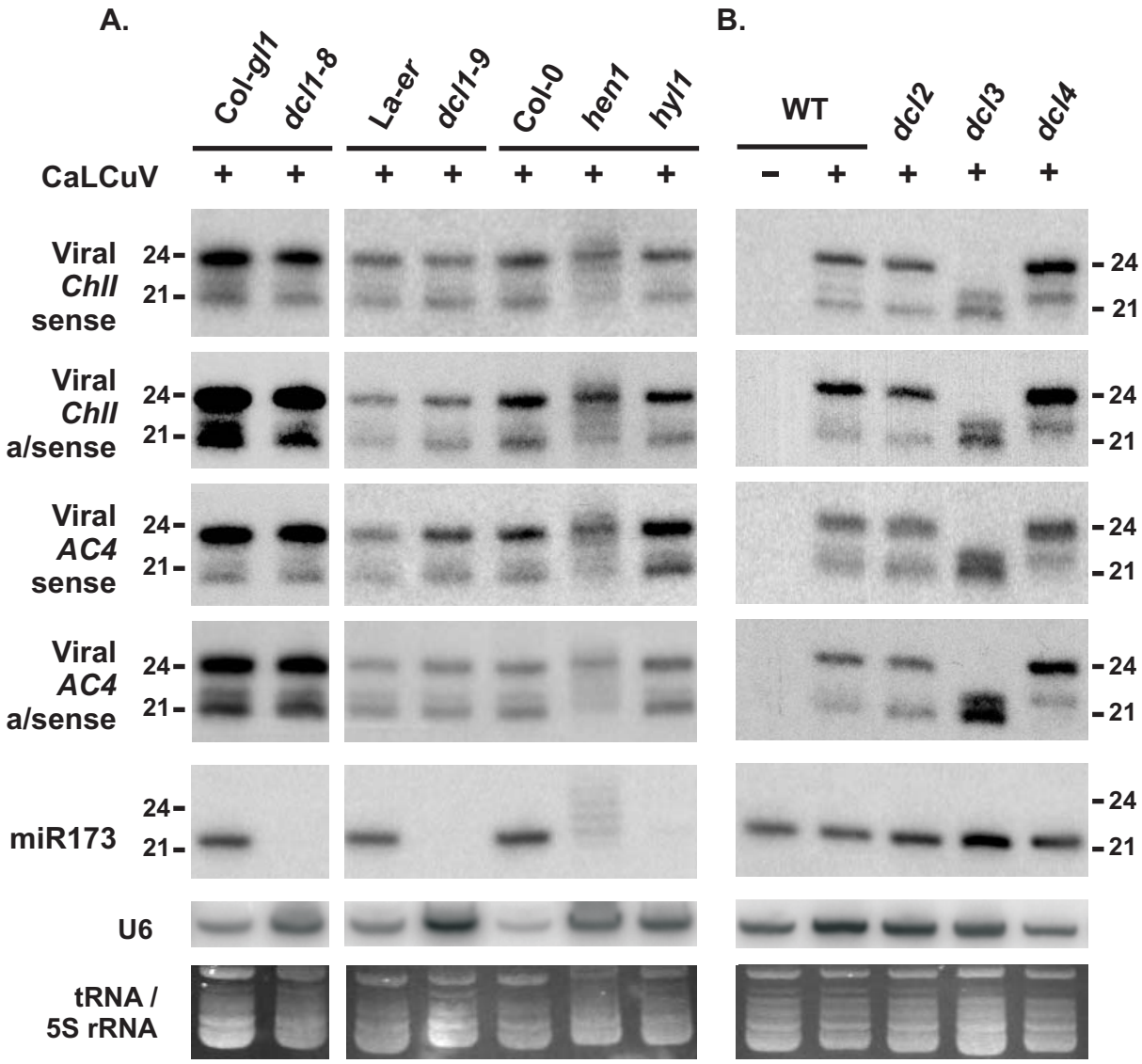


Figure 15. CaLCuV siRNA biogenesis in miRNA pathway and DCL-deficient mutants. Low molecular weight RNA was isolated from whole plants 28 dpi with CaLCuV::*ChlI* and analyzed by blot hybridization. Synthetic 21 and 24 nt RNA oligos were used as size markers. Membranes were successively hybridized with DNA oligonucleotide probes for the viral *ChlI*-fragment and viral *AC4* gene; miR173 is an endogenous control whose accumulation is abolished in *dcl1*mutants. **(A)** Analysis of mutants affecting miRNA biogenesis -- *dcl1-8*, *dcl1-9*, *hyl1*, *hen1* -- compared to their wild type (WT) controls: Col-*gl1*, La-*er* and Col-0. **(B)** Analysis of infected WT and DCL-deficient mutants. U6 snRNA and tRNA/5S rRNA are loading controls.

7.3. Biogenesis of CaLCuV::*ChlI* siRNAs in *DCL*-deficient mutants

Chapter 5 showed that DCL family members have redundant functions in the biogenesis of siRNAs derived from DNA repeats. Here, I tested the same *dcl2*, *dcl3*, and *dcl4* mutant strains, infected with CaLCuV::*ChlI*, for their effects on the accumulation of siRNAs derived from the viral *ChlI* insert and *AC4* regions (**Figure 15B**). Relative to wild type, 24 nt siRNA accumulation was unaffected in *dcl2*, slightly enhanced in *dcl4*, but not detected in *dcl3*; accumulation of 22 nt siRNA was slightly enhanced in *dcl3* and *dcl4*, but not detected in *dcl2*; and accumulation of 21 nt siRNAs was unaffected in *dcl2*, slightly enhanced in *dcl3*, but not detected in *dcl4*. These results are consistent with the interpretation that DCL4 is required for 21 nt siRNA formation, DCL2 is required for 22 nt siRNA formation, and DCL3 is required for 24 nt siRNA formation. These DCL / size-class correspondences were observed for viral siRNAs of both sense and antisense polarities (**Figure 15B**). Accumulation of the miR173 control was not significantly affected by *dcl2*, *dcl3* or *dcl4* mutations, as expected (Xie et al., 2004; Xie et al., 2005).

To test for compensation between *Arabidopsis* DCLs effecting biogenesis of viral siRNAs, I infected the double and triple *dcl*-mutants described in Chapter 5 with CaLCuV::*ChlI*. **Figure 16A** shows that that disease symptoms following infection with CaLCuV::*ChlI* were similar to the wild type in the mutant combinations *d23*, *d24*, *d34* and *d234*, as were the virus titers (data not shown). While *d23* exhibited chlorata bleaching in upper rosette leaves, the mutant combinations *d24*, *d34*, and *d234* -- all which contain the *dcl4* mutation -- did not (**Figure 16A**). All infected *dcl*-mutant combinations showed reduced accumulation of endogenous *CHLI* mRNA (**Figure 16B**) as well as residual accumulation of virus-derived *ChlI* siRNAs (**Figure 16C**). As expected from my previous analysis of single *dcl*-mutants, individual size classes of 21, 24 or 22 nt viral siRNAs accumulated in *d23*, *d24* and *d34*, respectively. Surprisingly, the *d234* still showed residual viral 21 nt siRNA accumulation, suggesting that DCL1 can compensate for the absence of DCL2, DCL3 and DCL4. Additionally, viral smRNA analyzed from infected *d234* showed slower migrating species of sense and antisense polarities, which might be partially processed dsRNA intermediates of siRNA biogenesis (**Figure 16C**).

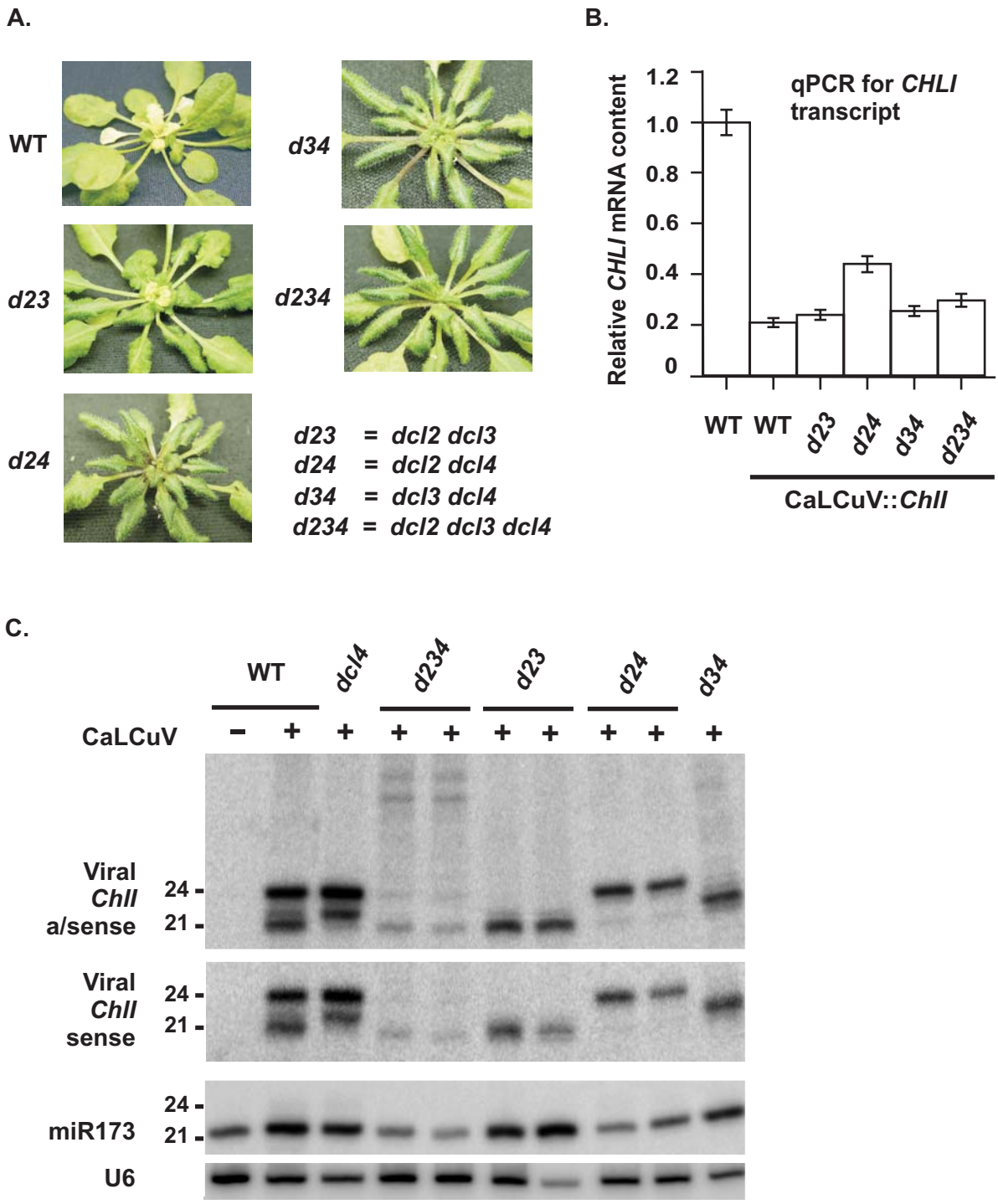


Figure 16. *Arabidopsis* DCLs are partially redundant for VIGS and siRNA biogenesis. **(A)** Phenotype of wild type (WT), double mutants *d23*, *d24*, *d34* and the triple mutant *d234*, 28 dpi with CaLCuV::*ChlI*. **(B)** Quantitative RT-PCR (qPCR) to measure *CHLI* mRNA levels in CaLCuV::*ChlI*-infected double and triple mutants, performed as in **Figure 13D**. **(C)** Low molecular weight RNA blot hybridization analysis of CaLCuV::*ChlI*-infected WT, *dcl4*, triple and double *dcl*-mutants. Replicate sample pools are shown for *d234*, *d23* and *d24* plants.

7.4. Summary

The biogenesis of CaLCuV-derived siRNAs does not appear to require DCL1 or HYL1, which are key components of the miRNA pathway, nor NRPD1a, RDR2 or RDR6, which are key components of the Pol IV and *trans*-acting siRNA pathways. In contrast, HEN1 is required for normal accumulation of 21, 22 and 24 nt size classes of CaLCuV siRNA species and appears to be important for methylation of their 3'-ends. DCL2, DCL3 and DCL4 mediate production of specific size classes of CaLCuV siRNAs. Alternative DCLs may compensate for the absence of others in terms of siRNA biogenesis and VIGS affecting *CHLI* mRNA accumulation. The chlorata phenotype, induced by CaLCuV::*ChlI* infection in wild-type plants, required *RDR6*, *DCL4* and *HEN1* but not other RNA silencing-related genes tested. Although DCL1 was thought mainly to function in miRNA biogenesis, 21 nt viral siRNAs in the triple mutant *d234* suggests that DCL1 can compensate for deficiency in the other three DCLs to mediate production of siRNAs from viral dsRNA.

8. Biogenesis of siRNAs in caulimovirus-infected *Arabidopsis*

The virion structure and life cycle of caulimoviruses differ considerably from that of geminiviruses. This chapter focuses on the biogenesis of siRNAs derived from an extensively studied caulimovirus, CaMV (Hohn, 1999). The circular dsDNA genome of CaMV is transcribed into a more-than-genome length, 35S transcript in the nucleus by the host RNA polymerase II (Rothnie et al., 1994). A shorter, 19S RNA transcript, which is coterminal with the 35S RNA, is also produced during CaMV infection. The 35S RNA functions as both messenger and pregenomic RNA in viral replication. These functions are controlled by the 600 nt-long leader sequence that folds into an elongated hairpin (Hemmings-Mieszczak et al., 1997; Pooggin et al., 1998). The 35S RNA leader contains three major stem sections, designated I, II, III (**Figure 17**, next page). The presence of these structures in one region of the same RNA molecule offered an exceptional opportunity to investigate the effect of viral RNA structure on siRNA biogenesis.

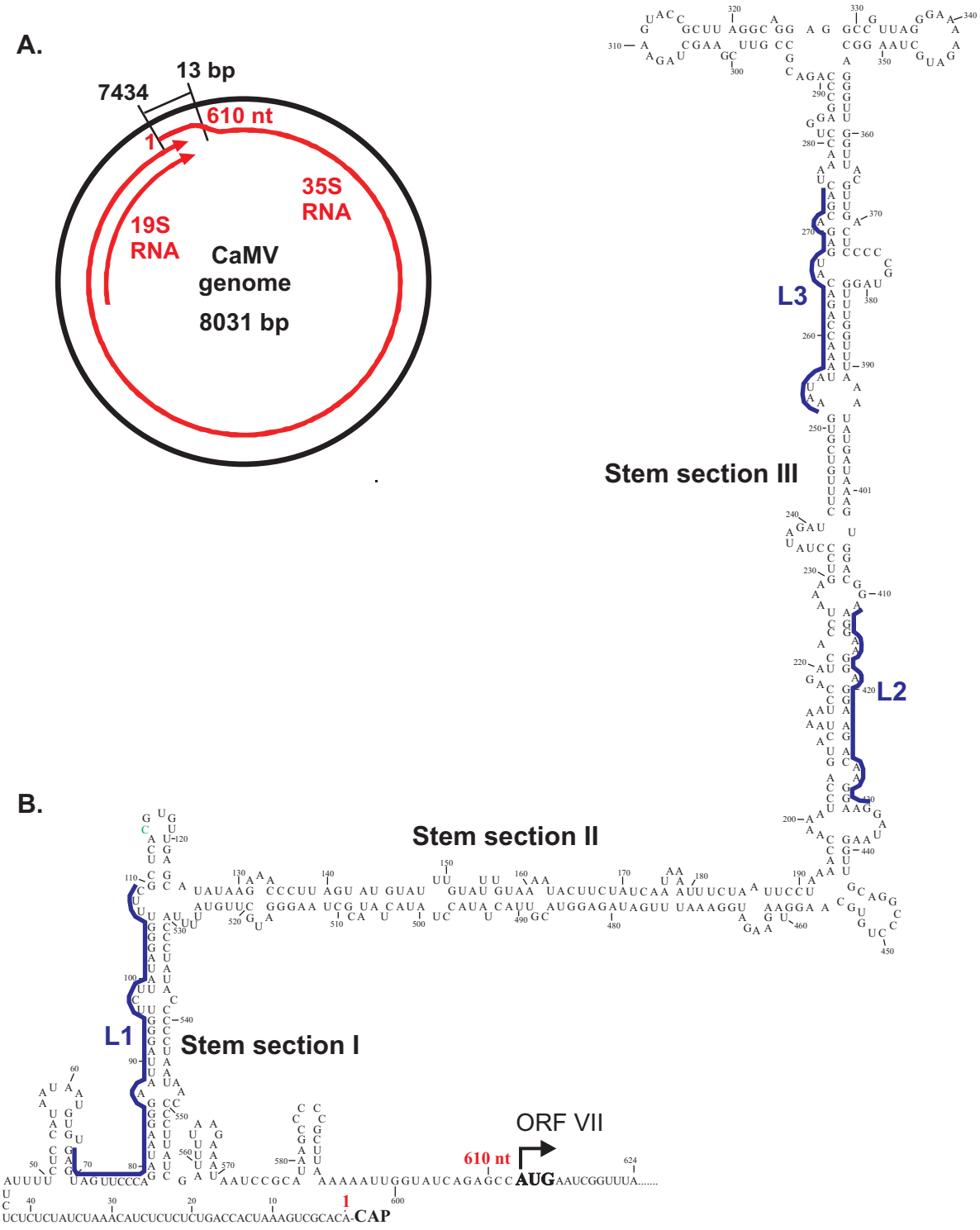


Figure 17. The 35S RNA transcript of Cauliflower Mosaic Virus (CaMV). (A) The 35S and 19S RNAs (red) are transcribed from minichromosomes formed by the circular, 8031 bp CaMV genome in the nucleus of infected plant cells. Basepairs 7434 to 13 are the template for the 35S RNA leader, which corresponds to transcript nucleotides 1 to 610 **(B)** The 35S RNA leader folds into a structure with three stem sections (I, II and III). A 40 nt region, L1, was used to design sense and antisense probes for RNA blot hybridization, whereas the L2 and L3 regions represent the most frequently sequenced viral siRNAs from CaMV-infected wild-type *Arabidopsis* (see **Section 8.2**). The diagram was adapted from Pooggin *et al.* (1998).

8.1. Analysis of siRNAs derived from the 35S RNA, Stem section I

To detect siRNAs representing specific regions of the 35S RNA leader, sense and antisense oligonucleotide probes (L1) were designed to cover the 30 nt-long, Stem section I (**Figure 17**). Mutants and appropriate wild-type strains were infected with CaMV and analyzed by RNA blot hybridization. L1 probes detected 21, 22 and 24 nt siRNAs in the RNA of infected wild type plants (**Figure 18A**). The patterns of L1-specfic viral siRNAs obtained for infected wild-type, *rdr2* and *rdr6* plants were qualitatively identical. Similar patterns of viral siRNAs were also obtained for these genotypes using an antisense L1 probe (**Figure 18A**). Next, siRNAs from infected *nrpd1a* and *ago4* plants were analyzed. *nrpd1a* and *ago4* showed the same pattern of sense and antisense viral siRNA accumulation as the wild type (**Figure 18A**). Hybridization using probes for other regions of the 35S RNA leader (**Figure 17**), as well as coding regions of the CaMV genome, showed no variation in the abovementioned mutants compared to the wild-type pattern (data not shown). Together, these results indicate that RDR2, RDR6, NRPD1a and AGO4 are not required for the biogenesis of CaMV siRNAs.

Different viral siRNA size classes showed reduced accumulation in specific *dcl*-mutants (**Figure 18B**). Relative to wild type, 24 nt siRNA accumulation was unaffected in *dcl1-9*, *dcl2* or *dcl4* but not detected in *dcl3*; accumulation of 22 nt siRNAs was unaffected in *dcl1-9*, *dcl3* or *dcl4*, but was greatly reduced in *dcl2*; and accumulation of 21 nt siRNAs was unaffected in *dcl2*, *dcl3* or *dcl4*, but was reduced in *dcl1-9*. These results are consist with the interpretation that DCL1 contributes to 21 nt siRNA formation, DCL2 contributes to 22 nt siRNA formation, and DCL3 is required for 24 nt siRNA formation. In contrast, L1-specific siRNA accumulation in *hen1* suggests that the three siRNA size classes require HEN1 for their stability or proper processing, but not for their biogenesis *per se*. Supporting a putative role for DCL1 in 21 nt viral siRNA biogenesis, the mutant deficient for its binding partner, HYL1, showed reduced 21 nt relative to 22 nt siRNA accumulation: i.e., rather than the 21 nt siRNA signal being stronger than that of 22 nt siRNAs (wild-type Col-0), the two size classes were equally abundant in *hyl1* (**Figure 18B**).

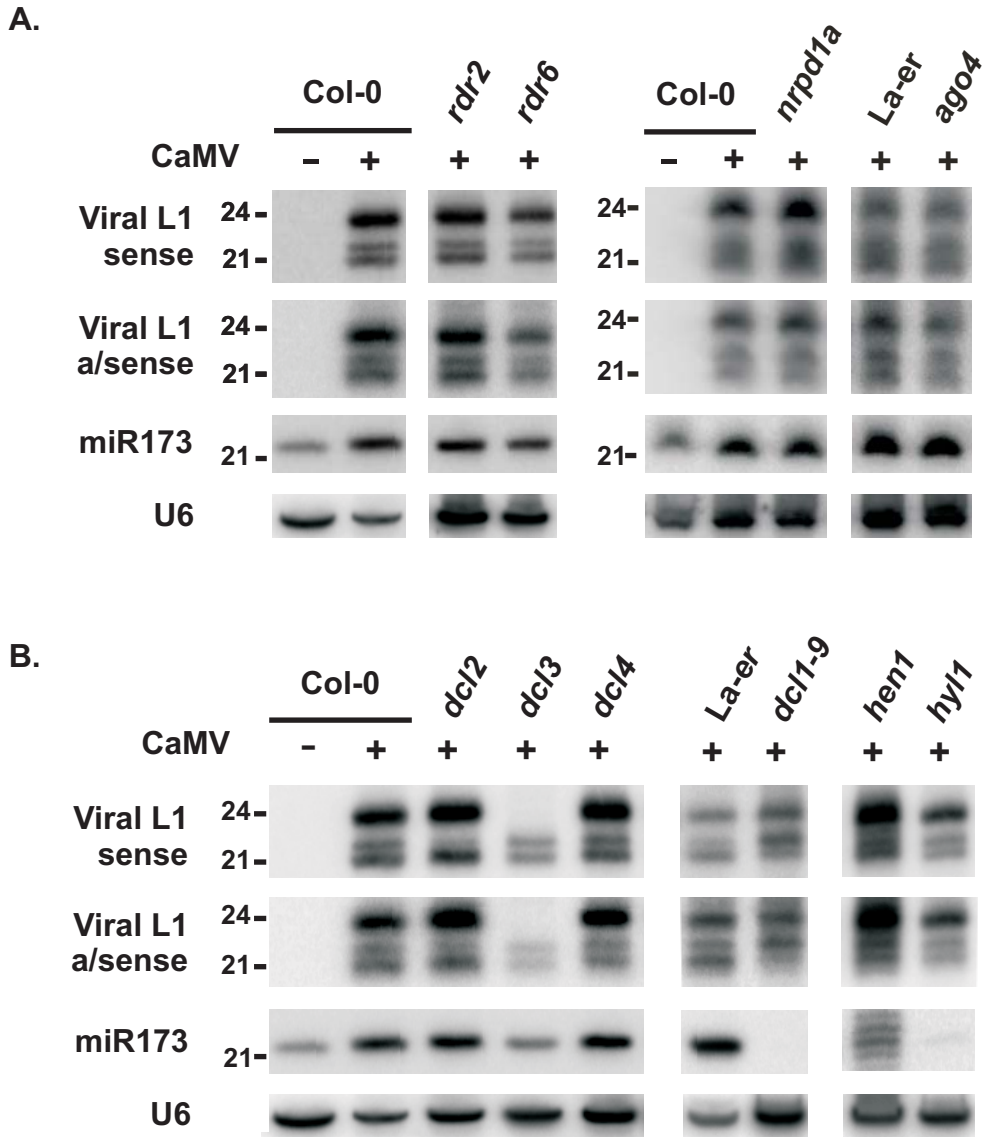


Figure 18. Analysis of siRNAs from Stem Section I of the CaMV 35S RNA leader. Low molecular weight RNA was isolated from whole plants 28 dpi with CaMV and analyzed by blot hybridization. Synthetic 21 and 24 nt RNA oligonucletides were used as size markers. Membranes were successively hybridized with DNA oligonucletide probes for viral and endogenous smRNAs (Appendix **Table A1**). **(A)** L1-specific siRNAs and controls detected in infected *rdr2, rdr6, nrpd1a* and *ago4* mutants. La-*er* is the wild type for *ago4*, and Col-0 the wild-type for all other mutants. **(B)** L1-specific siRNAs and controls detected in infected *dcl*, *hen1* and *hyl1* mutants. La-*er* is the wild type for *dcl1-9*, and Col-0 the wild type for all other mutants.

8.2. Sequencing of smRNAs from CaMV-infected *Arabidopsis*

High-throughput sequencing was used to evaluate the contribution of DCL1 to biogenesis of siRNAs representing the entire CaMV 35S RNA transcript. Because *dcl1-8* and *dcl1-9* are partial loss-of-function mutants, those strains were not chosen for sequencing. Instead, smRNAs from infected wild-type and triple mutant *dcl2 dcl3 dcl4* (*d234*) plants were sequenced. The smRNA population remaining in a *d234* background would likely be the result of DCL1 processing. The fraction of smRNAs approximately 16-28 nt in length was isolated from RNA of six infected rosette-stage plants of each genotype using polyacrylamide gel electrophoresis. The 5' and 3' ends of these smRNAs were ligated to oligonucleotide adapters and these products were used as templates to synthesize cDNAs, and subsequently dsDNA containing smRNA-derived inserts, which were sequenced using Solexa technology (**Figure 19A**) (Solexa, 2006).

The sequencing data were processed *in silico* to identify smRNAs inserts with intact 5' and 3' adapter sequences on both ends. This yielded 127'499 unique smRNA inserts for wild-type and 126'952 for *d234* plants, which were matched to the *Arabidopsis* and CaMV genomes using a *Basic Local Alignment Search Tool* (BLAST) analysis. Many sequences in the raw datasets were too short (<<20 nt) for the BLAST method to be representative or were not identical to sequences in the reference genomes. These sequences as well as a small fraction of large (>>25 nt) sequences, which exceeded the quality limit of the sequencing method, were discarded. The remaining "clean" sequence datasets contain ca. 50'000, 20-25 nt long, sequences that exactly matched the *Arabidopsis* or CaMV genome; these datasets were used in all subsequent analyses. Approximately 53% of the wild-type dataset and 34% of the *d234* dataset represented CaMV-derived siRNAs. This modest reduction in the viral siRNA population in *d234* seems reasonable, given that DCL1 is still expressed in *d234* plants.

The size distribution of endogenous *Arabidopsis* smRNAs sequenced from infected wild-type plants (**Figure 19B**) was qualitatively similar to that obtained earlier by cloning (see **Figure 4A** and Xie et al., 2004). Endogenous smRNAs from the infected wild type were predominantly 24 nt with a smaller peak at 21 nt. The abundance of the 21 nt size class increased and that of the 24 nt size class decreased in *d234*. Similarly, infected wild-type plants accumulated major 21 *and* 24 nt size classes of CaMV siRNAs, and there was a pronounced shift to the 21 nt size class in infected *d234* plants (**Figure 19C**).

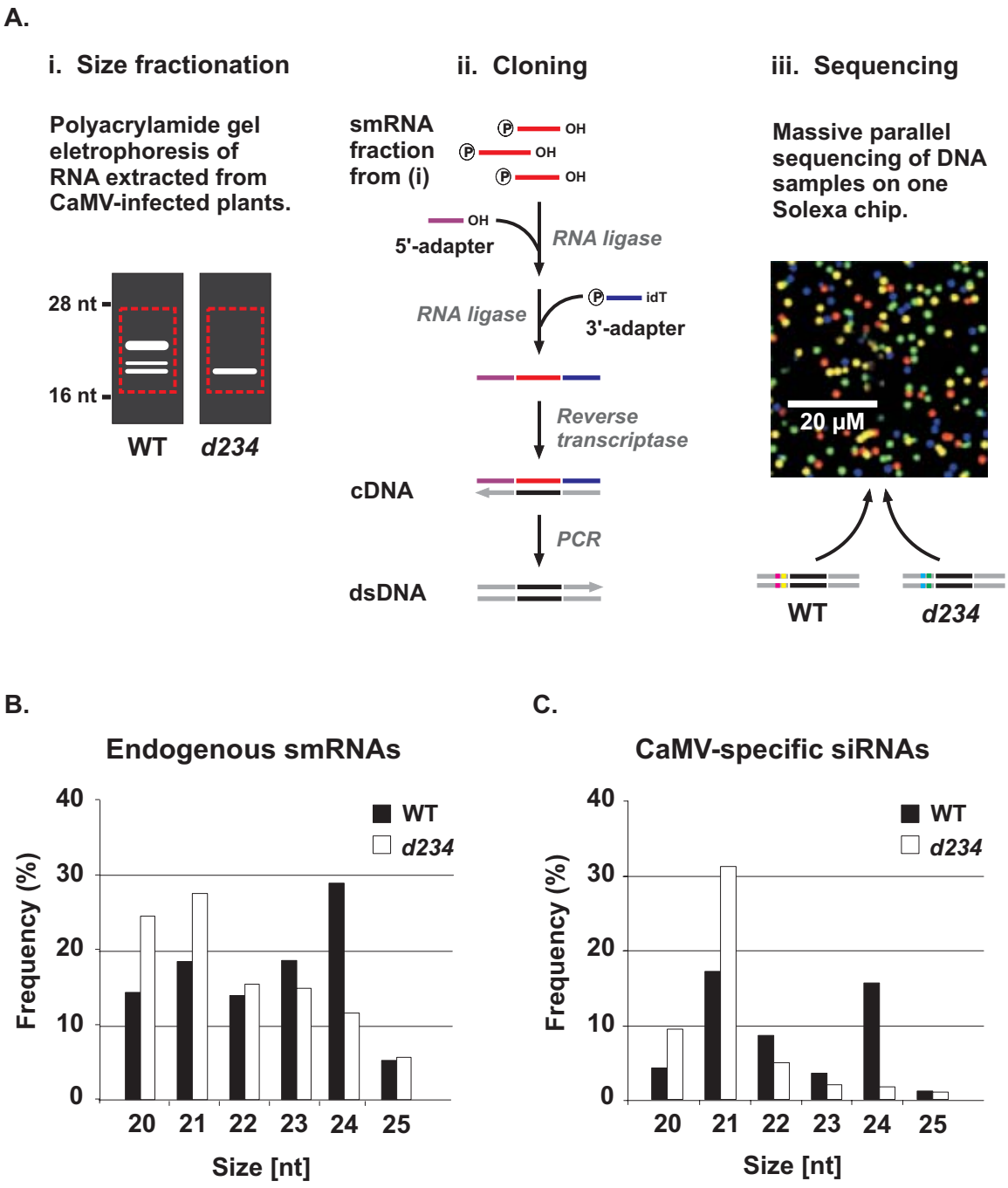


Figure 19. Sequencing of smRNAs from CaMV-infected *Arabidopsis*. (A) RNA was extracted from pools of CaMV-infected, wild-type (WT) and *dcl2 dcl3 dcl4* (*d234*) *Arabidopsis* (i) The ~16-28 smRNA fraction was isolated from total RNA by polyacrylamide gel electrophoresis. (ii) Oligonucleotide adapters were ligated to the 5' and 3'-ends of smRNAs. Reverse transcription was used to synthesize cDNA, and polymerase chain reaction (PCR) performed to amplify dsDNA containing smRNA sequence inserts. (iii) The two DNA samples, generated with slightly different 5'-adapters, were sequenced using Solexa technology. **(B)** Size distribution of endogenous smRNAs sequenced from CaMV-infected WT and *d234* plants. **(C)** Size distribution of CaMV-derived siRNAs sequenced from WT and *d234*. In **Panels B/C,** only sequences that perfectly matched the plant or viral genomes were considered. Frequency was calculated relative to the total number of perfectly matched, 20-25 nt smRNAs.

The most striking feature of CaMV-derived siRNAs was their biased distribution when mapped to the CaMV genome (8031 bp). An overwhelming majority (94%) of viral siRNAs in wild-type plants mapped to genomic coordinates 7400-8031 bp, corresponding to the 35S RNA leader (**Figure 20A**). The majority of siRNAs (88%) represent the sense (+) orientation of the 35S transcript. A higher resolution map shows that siRNAs clustered at three particular regions in the leader, which appear to be "hot spots" for siRNA accumulation (**Figure 20B**). The 5' hot spot (L1) corresponds to leader Stem Section I, whose contribution to siRNA accumulation was analyzed in **Section 8.1**. The 3' hot spot (L2) has recently been reported by Moissiard and Voinnet (2006) and matches the *descending* arm of leader Stem section III. Finally, a novel hot spot (L3) corresponds to the *ascending* arm of leader Stem section III. Thus, the three major hot spots for CaMV siRNA accumulation (L1, L2, and L3) correspond to regions within Stem sections I and III of the 35S RNA transcript leader (see **Figure 17B**).

Although the number of viral siRNAs in the total *d234* dataset was ~50% of that in the wild type, 90% of viral species in *d234* were still derived from the leader and clustered at the same hot spots identified in wild-type plants (**Figure 20C**). The siRNAs sequenced from *d234* were almost exclusively in the 20-22 nt size range; 24 nt viral siRNAs were practically absent. Taken together, these data suggest that in the absence of other DCLs, DCL1 is sufficient to establish the characteristic pattern of siRNA biogenesis representing the 35S RNA leader.

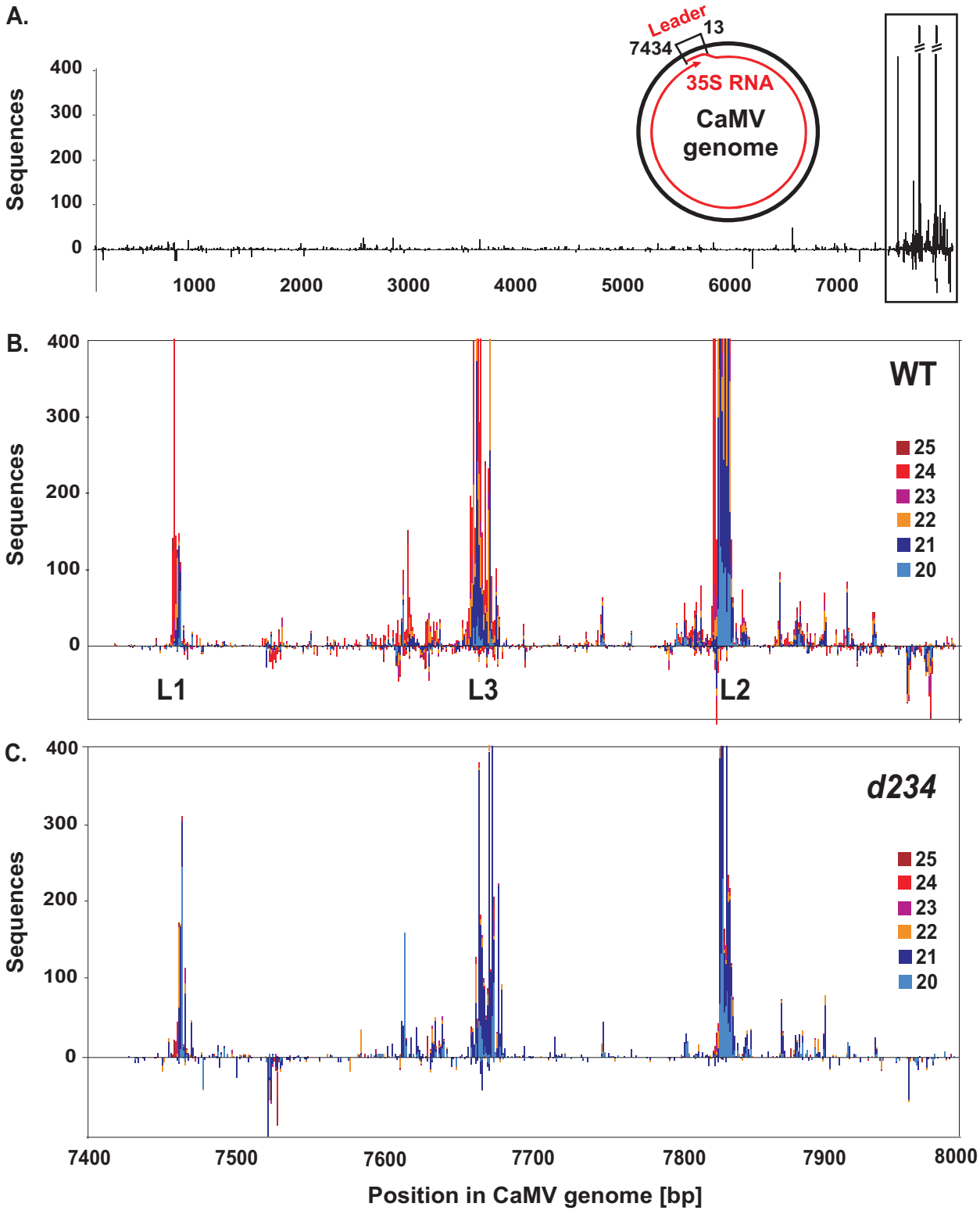


Figure 20. Genome distribution of CaMV-derived siRNAs. (A) Map of viral siRNAs sequenced from CaMV-infected WT *Arabidopsis*. The x-axis represents the genome of CaMV (8031 bp), and the y-axis shows the total number of sequences whose 5'-ends align to a particular position; bars beneath the x-axis show antisense sequences. The black box indicates the CaMV genomic coordinates 7400-8000 bp, roughly corresponding to the 35S RNA leader region. **(B)** High resolution map of leader siRNAs sequenced from WT plants: the L1, L2 and L3 regions are hotspots for smRNA accumulation. **(C)** High resolution map of leader siRNAs sequences from *d234* triple mutant plants. Colors in **Panels A/B** show totals of 20, 21, 22, 23, 24 and 25 nt siRNAs.

8.3. Analysis of siRNAs derived from the 35S RNA, Stem section III

The great majority of siRNAs sequenced from leader Stem section III fell into two ~27 nt windows on its ascending (L3) and descending arms (L2), regions which are indicated in **Figure 17B**. Because L2 has been studied in parallel work by Moissiard and Voinnet (2006), I focused my analysis on the novel hotspot, L3. In the wild type, of 1771 L3-derived sequences identified, 32% were 21 nt long and 46% were 24 nt long. This distribution was dramatically shifted in *d234* plants, in which 69% of the 628 sequences identified in the 27 nt window were 21 nt long and 6% where 24 nt long. The representation of particular L3-derived sequences in the wild-type and *d234* data are shown schematically in **Figure 21A**.

The DCL requirements for L3-derived siRNA production were tested using oligonucleotide probes for the most frequent L3 sequence, with samples from wild-type, *dcl2*, *dcl3* and *dcl4* plants (**Figure 21B**). Both sense and antisense probes detected 21, 22 and 24 nt siRNAs in the RNA of virus-infected wild type plants. The accumulation of the 22 and 24 nt siRNAs (sense and antisense) depended on DCL2 and DCL3, respectively. While sense 21 nt siRNAs accumulated in the *dcl4* mutant, corresponding antisense 21 nt siRNAs were barely detectable (**Figure 21B**). For comparison, I also tested accumulation of siRNAs from the L2 region. My results suggest that DCL4 contributes to antisense, 21 nt siRNA biogenesis from L2 and L3 regions; however, some 21 nt signal is still detected in *dcl4* (**Figure 21B**). The effect of *dcl4* on accumulation of antisense L1-specific siRNAs was less striking (**Figure 18B**). Although antisense siRNAs corresponding to L3 were sequenced, their abundance was very low in the datasets, e.g., 10-13 replicate sequences (0.7**%**) from wild-type plants. This may indicate that the sequencing approach -- perhaps the ligation step -- is biased against these sequences, or that antisense siRNAs are less stable *in planta*.

The *d234* sequence dataset contained abundant 21 nt siRNAs from the L3 viral region (**Figure 21A**). This indicated, as did my results for the L1 region, that DCL1 contributes to the biogenesis of 21 nt CaMV siRNAs. To test this hypothesis, I examined the accumulation of L3-specific siRNAs associated with CaMV infection of quadruple *dcl*-mutants (materials described in **Section 5.3**). RNA blot hybridization confirmed that L3-specific, 21 nt siRNA titers were lower in quadruple *dcl*-mutants, than in *d234* (**Figure 21C**). This is consistent with the idea that DCL1 directly mediates production of ~21 nt siRNAs from the L3 region of the 35S RNA leader or related dsRNA species. Since the *dcl1* alleles in the quadruple mutants are hypomorphic, a complete absence of 21 nt siRNA signal would not be expected. Interestingly, the overall signal detected for viral 21-24 nt siRNAs was lower in *dcl1*-9 than in either the wild type or *d234* samples.

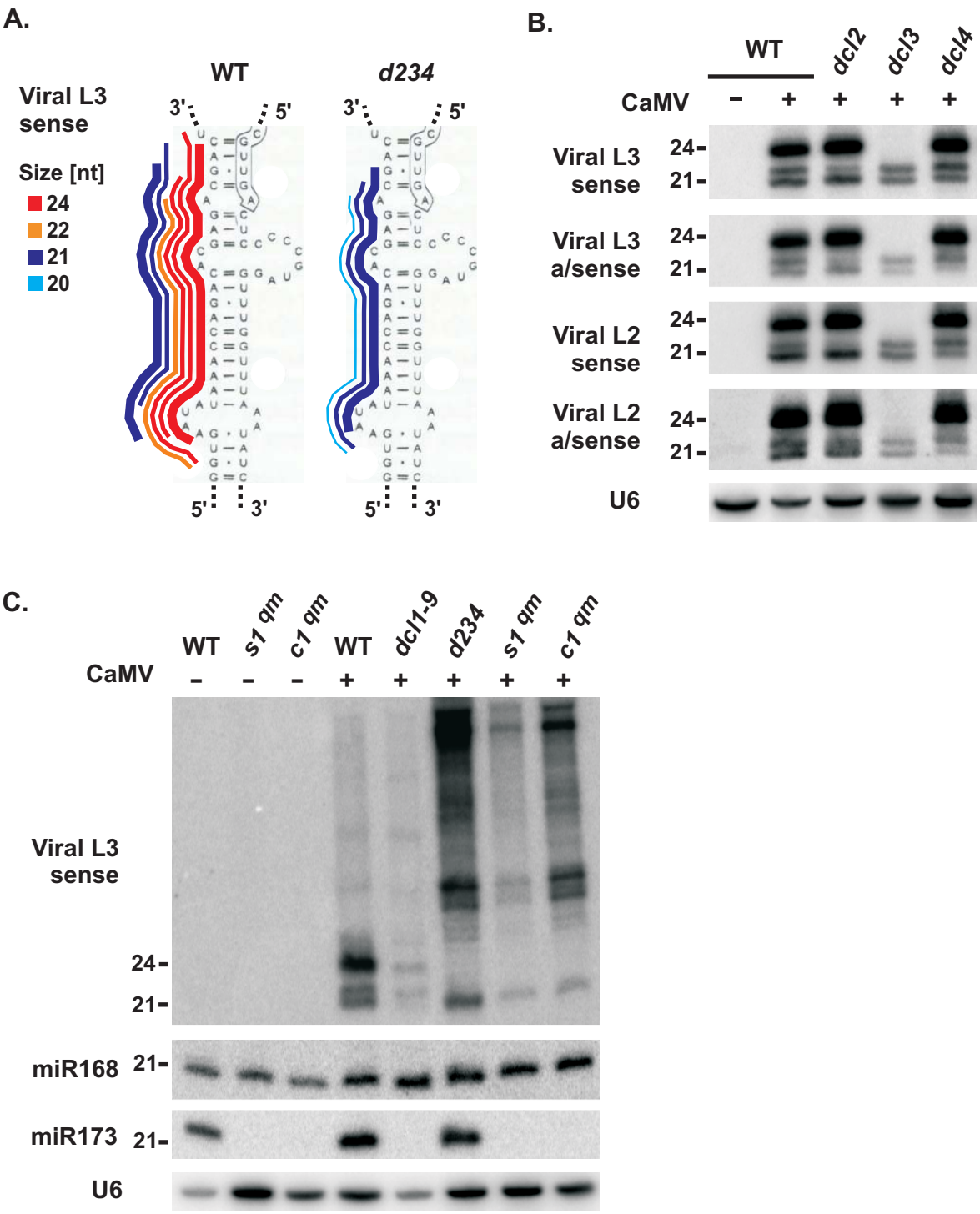


Figure 21. Analysis of siRNAs from Stem Section III of the CaMV 35S RNA. (A) The stem structure in the L3 region of the 35S RNA leader. Each colored bar represents an L3-specific siRNA that was sequenced at least 50 times: The heaviest bars represent >250, intermediate bars 100-250, and lightest bars 50-99 replicate sequences. Probes were designed to detect siRNAs corresponding to the regions L3 and L2 (see **Figure 17B**). **(B)** L3- and L2-specific siRNAs detected in *dcl2*, *dcl3* and *dcl4* samples **(C)** L3-specific siRNAs detected from quadruple *dcl*-mutant samples: *dcl1-8 dcl2 dcl3 dcl4* ($s1^{qm}$) and *dcl1-9 dcl2 dcl3 dcl4* ($c1^{qm}$). miR168, which accumulates in *dcl1-8* and *dcl1-9,* and miR173, which is sensitive to these mutations, are controls for the quadruple mutants.

This could mean that DCL1 facilitates processing of the 35S RNA leader by other DCLs. Furthermore, CaMV-infected triple and quadruple *dcl*-mutants showed increased accumulation of higher molecular weight RNAs homologous to L3 sense and antisense probes (shown here for sense). These are potentially dsRNA intermediates that overaccumulate due to multiple DCL-deficiency (**Figure 21C**).

8.4. Summary

Genome-scale sequencing of smRNAs from CaMV-infected plants identified three hotspots for siRNA accumulation matching the 35S RNA leader. These hotspots are also found in CaMV-infected, triple mutant *d234* plants. Accumulation of the 22-24 nt fraction of these siRNAs depended on the same DCLs as did siRNAs representing endogenous 5S rDNA repeats and CaLCuV::*ChlI*. In contrast, biogenesis of CaMV-derived 21 nt siRNAs seems to be primarily DCL1-dependent, and partially DCL4-dependent, with the latter being particularly apparent for antisense 21 nt siRNA species. This last finding is difficult to verify for the whole CaMV genome, because antisense siRNAs are underrepresented in the sequencing datasets. Given the overall reduction of 21-24 nt viral siRNAs in the *dcl1-9* mutant, these results are consistent with the hypothesis that DCL1 processes the structured 35S RNA leader, perhaps giving rise to substrates for DCL2, DCL3 and DCL4.

9. Discussion

9.1. Tandem repeat-derived siRNAs and the Pol IV pathway.

Repeated DNA motifs, one to thousands of bases in length, are a characteristic of all eukaryotic genomes (Heslop-Harrison, 2000). Variation in simple genomic repeats can cause genetic disease (Mirkin, 2007), whereas longer repeats are often transposable element insertions or tandem repeat arrays, some of which have conserved biological functions (Kumar and Bennetzen, 1999; Heslop-Harrison et al., 2003; Eickbush and Eickbush, 2007). The first part of this dissertation dealt with the relationship between small RNA pathways and two major types of genomic tandem repeat in *Arabidopsis*: the 180 bp centromeric satellite and 5S rDNA pericentromeric repeats (Campell et al., 1992; Cloix et al., 2000; Heslop-Harrison et al., 2003). Cytosines in these tandem repeats are heavily methylated (Martinez-Zapater et al., 1986; Vongs et al., 1993), a phenomenon also reported for tandem repeats of other plants (Goldsbrough et al., 1981; Fulnecek et al., 1998; Kovarik et al., 2000). My work explored the function of siRNAs in RNA-directed DNA methylation (RdDM) of *Arabidopsis* tandem repeats.

Centromeres are sites for spindle attachment to chromosomes, allowing kinetochore assembly, and are thus required for faithful segregation of chromosomes in cell division (Dawe and Henikoff, 2006). Budding yeast centromeres are defined by specific, short DNA sequences, whereas the functional relevance of satellite sequences in multicellular eukaryotes is not clear (Morris and Moazed, 2007). In plants, these centromeric repeats vary greatly between species, both in sequence and copy number (Jiang et al., 2003; Ma et al., 2007). Indirect evidence suggests that RNA transcribed from satellite repeats contributes to plant centromere function. For example, centromere-encoded 40-200 nt RNAs are integral components of the maize kinetochore (Topp et al., 2004). In addition, centromeric siRNAs were detected in RNA from *Arabidopsis* and *Oryza punctata* using probes for satellite repeat units of those plants (May et al., 2005; Zhang et al., 2005b). These studies raise the possibility that repeat-derived transcripts and/or siRNAs participate in plant centromere function, as reported earlier for fission yeast (Hall et al., 2003; Volpe et al., 2003; Ekwall, 2004). That said, a role for these siRNAs or the Pol IV pathway in cytosine methylation of 180 bp repeats is not supported by my data (Chapter 5) or that from two other groups (Kanno et al., 2005b; May et al., 2005).

The second type of *Arabidopsis* tandem repeat, 5S rRNA genes, are transcribed by RNA polymerase III to generate 120 nt, 5S rRNA molecules (Douet and Tourmente, 2007). The 5S rRNA is a structurally conserved, integral component of the large ribosomal

subunit in all organisms and appears to be essential for ribosome function (Ammons et al., 1999; Szymanski et al., 2003; Kiparisov et al., 2005). The functional significance of 5S rDNA methylation, however, is still not understood: unlike protein-coding genes in *Arabidopsis*, whose expression level tends to be negatively correlated with flanking promoter and 3'-end methylation (Zhang et al., 2006b; Zilberman et al., 2007), the steady state level of 5S rRNA is not affected by reduced 5S rDNA methylation (Mathieu et al., 2002; Vaillant et al., 2007a). However, analysis of nucleotide polymorphisms has shown that "major" and "minor" 5S rRNA transcripts exist, and that the contribution of minor 5S rRNA species to the total transcript pool increases with decreasing 5S rDNA methylation (Cloix et al., 2002; Mathieu et al., 2003). This indicates that DNA methylation could regulate the differential expression of specific 5S rDNA repeat units (Vaillant et al., 2007a).

9.1.1. Biogenesis of siRNAs derived from *Arabidopsis* tandem genomic repeats.

In Chapter 5, I showed that the 180 bp satellite and 5S rDNA repeats are each a source of multiple, overlapping siRNA clones. These siRNAs are predominantly ~24 nt in length and their accumulation requires NRPD1a and RDR2, like siRNAs derived from *AtSN1* and *FWA* repeats (Hamilton et al., 2002; Xie et al., 2004; Onodera et al., 2005; Chan et al., 2006a). No tandem repeat-derived siRNAs were detected in RNA from *nrpd1a* or *rdr2* mutants, which is consistent with recent genome-scale sequencing of smRNA isolated from *nrpd1a* and *rdr2* (Lu et al., 2006; Zhang et al., 2007). However, genetic requirements for tandem (180 bp and 5S rDNA) and dispersed (*AtSN1*) repeat siRNA accumulation are different from those of inverted repeat siRNA accumulation. Zhang *et al.* (2007) documented 21-24 nt siRNAs derived from an endogenous, inverted repeat called *IR71*; accumulation of these siRNAs did not require NRPD1a or RDR2. Similarly, siRNA expression from a transcribed inverted repeat transgene was unaffected by NRPD1a-deficiency (Kanno et al., 2005b). Thus, production of tandem repeat-derived siRNAs depends on the Pol IV pathway, while that of inverted repeat-derived siRNAs does not.

I propose that the difference between tandem and inverted repeat siRNA biogenesis reflects a requirement for Pol IVa and RDR2 in the synthesis of RNA complementary to tandem repeat transcripts (**Figure 22A**). Production of siRNAs from inverted repeat-derived transcripts would not require this process, because they fold back upon themselves to form extended regions of dsRNA (Svoboda and Di Cara, 2006; Zhang et al., 2007). Before the discovery of a Pol IV function in endogenous RNA silencing, similar models were proposed based on observations using inverted repeat transgenes (Stam et al., 1998; Mette et al., 1999; Smith et al., 2000; Zilberman et al., 2004).

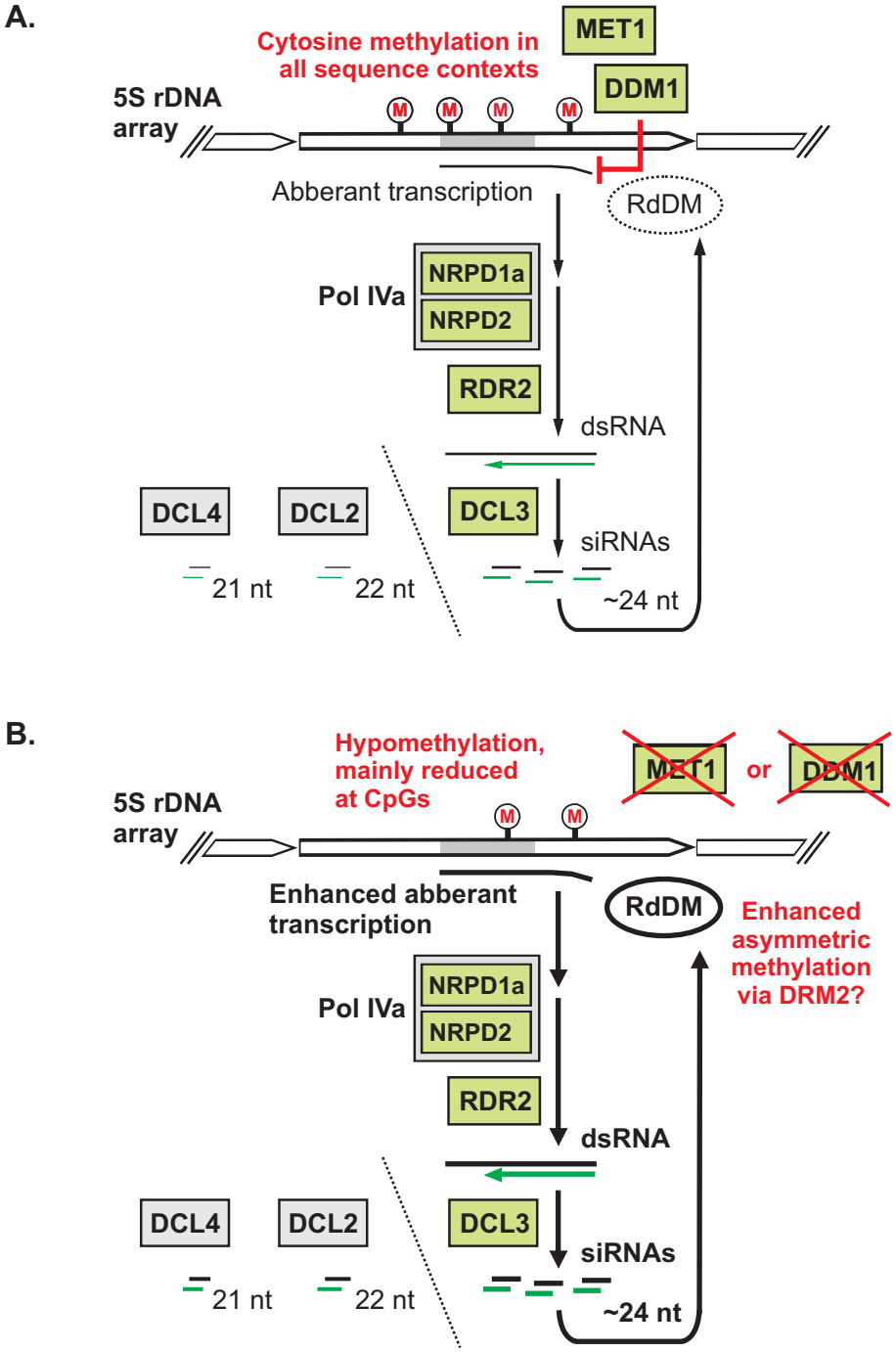


Figure 22. Model for Pol IV pathway-mediated methylation of 5S rDNA repeats. (A) *Arabidopsis* 5S rDNA repeats of are heavily methylated in all sequence contexts. Abberant (e.g., 210 nt) 5S rDNA-derived transcripts accumulate at low levels, being repressed by MET1/DDM1. These transcripts are potential subtrates for dsRNA biogenesis mediated by Pol IVa and RDR2. DCL3 could subsequently process this dsRNA into ~24 nt siRNAs, which are thought to guide RdDM. **(B)** When 5S rDNA repeats are hypomethylated (i.e., in *ddm1* or *met1*), abberant 5S rDNA transcription is enhanced, correlating with increased accumulation of 24 nt siRNAs. RdDM of 5S rDNA is also potentially enhanced. Dense lines and bold labels indicate increase production of molecules compared to **Panel A**. DCL2 and DCL4 can compensate for DCL3 in this pathway, producing alternate siRNA size classes.

My genetic dissection of siRNA biogenesis suggests that the Pol IV pathway is generally required for production of siRNAs from tandem genomic repeats, whether highly (5S rDNA) or weakly (180 bp) transcribed (May et al., 2005; Vaillant et al., 2007a).

In both cases, tandem repeat-derived siRNA accumulation showed a hierarchical dependence on three DCLs: the predominant 23-24 nt siRNAs in wild-type plants appear to be DCL3 products, whereas aberrant accumulation of 21 and 22 nt species in the *dcl3* mutant requires DCL4 and DCL2, respectively. Apparently, alternative DCL proteins compensate for DCL3 deficiency, mediating production of siRNAs from repeat-derived dsRNA that is normally processed by DCL3 in the wild type (**Figure 22**). In this respect, 180 bp and 5S rDNA-derived siRNA accumulation showed the same basic pattern of DCL-dependencies reported for other endogenous repeats, like *AtSN1* and *IR71* (Hamilton et al., 2002; Xie et al., 2004; Gasciolli et al., 2005; Henderson et al., 2006). Therefore, the DCL step marks a point of convergence between endogenous tandem (5S rDNA and 180 bp), dispersed (*AtSN1*) and inverted repeat (*IR71*) siRNA biogenesis in *Arabidopsis*.

Little is known about precursors upstream of dsRNA and siRNA production in the Pol IV pathway. However, an important clue is the overaccumulation of 24 nt, 5S rDNA-derived siRNAs in *met1* and *ddm1* mutants, described here and by Onodera et al. (2005). This overaccumulation cannot be explained by a specific perturbation in DCL3, because I found that the same effect occurred in *ddm1 dcl3* double mutant plants, where different siRNA size classes and DCLs are implicated. However, these observations might be explained by enhanced production of 5S rDNA-derived transcripts in *met1* and *ddm1* plants. MET1 and DDM1 maintain CpG-methylation and heterochromatin of tandem genomic repeats in wild-type plants (Vongs et al., 1993; Gendrel et al., 2002). The mature 120 nt 5S rRNA transcript terminates upstream of the cluster of siRNAs that I described in Chapter 5, but an extended 210 nt transcript was recently identified in *met1* and *ddm1* mutants (Vaillant et al., 2006), which could be a precursor for siRNA biogenesis.

I propose that the CpG-methylated state of 5S rDNA limits expression of aberrant, read-through transcripts, which are substrates for dsRNA production via Pol IVa and RDR2. The dsRNA could then be processed by DCL3 into 24 nt siRNAs that guide RdDM to cytosines in 5S rDNA repeats (**Figure 22A**), as proposed by two current models (Pontes et al., 2006; Matzke et al., 2007). Increased aberrant transcription of 5S rDNA in *met1* and *ddm1* compared to the wild type (Vaillant et al., 2006) would thus account for the overaccumulation of 5S rDNA-specific siRNAs observed in those mutants (**Figure 22B**). Maintenance of methylation in CpNpG sites, which requires CMT3, appears not to influence biogenesis of 5S rDNA-derived siRNAs: these species did not overaccumulate in the *cmt3* mutant, nor did 210 nt 5S rDNA transcripts accumulate in that strain (Onodera et

al., 2005; Vaillant et al., 2006). However, there is a clear distinction between 5S rDNA and many other repeats in *Arabidopsis*, because siRNAs derived from 180 bp, *AtSN1* and *FWA* repeats do not overaccumulate in *met1* or *ddm1* (Lippman et al., 2003; Lippman et al., 2004; Onodera et al., 2005; Chan et al., 2006a).

Vaillant *et al.* (2006) and (2007) showed that *ddm1*, *met1*, *cmt3* and *ago4* mutants release silencing of minor 5S rRNA transcripts, indicating that DNA methylation and (perhaps) the Pol IV pathway regulate expression of specific 5S rDNA repeat units. However, the accumulation of minor 5S rRNA transcripts needs to be analyzed in mutants deficient for NRPD1a, RDR2 and the four DCLs to confirm that the Pol IV pathway mediates the silencing hypothesized by Vaillant *et al.* (2007). In addition, the level of aberrant 5S rDNA transcripts should be measured in mutants deficient in that pathway. My model in **Figure 22** postulates that aberrant 210 nt transcripts derived from 5S rDNA repeats are templates for NRPD1a and RDR2-mediated production of dsRNAs, which would be processed into siRNAs by DCL3 (and alternatively, by other DCLs). It thus predicts that 210 nt or other aberrant transcripts will overaccumulate in *nrpd1a*, *rdr2*, *dcl3* and the triple mutant *d234*, whereas RNA silencing affecting minor 5S rRNAs might be compromised in those strains, as was observed in *ago4*.

9.1.2. Function of repeat-derived siRNAs in 5S rDNA methylation.

Despite their constitutive expression in wild-type plants, the bulk of 5S rDNA-derived siRNAs are not required to maintain cytosine methylation of 5S rDNA repeats: quadruple *dcl*-mutants, which expressed much lower levels of siRNAs than the wild type, showed no changes in *Hpa* II-site methylation in 5S rDNA repeats. In contrast, these sites were hypomethylated in mutants deficient for Pol IVa subunits or RDR2, which act upstream of DCL3 in the production of siRNAs (Chapter 5). Although these differences in 5S rDNA methylation at *Hpa* II sites were modest, they were reproduced in different laboratories, using independently propagated materials and different mutant alleles, and therefore cannot be explained by transgenerational variation amongst *dcl*, *rdr2* and *nrpd1a* lines (Kanno et al., 2005b; Onodera et al., 2005; Pontes et al., 2006). Furthermore, I did not detect changes in 5S rDNA methylation in two *dcl3* lineages after propagating them three generations by self-fertilization. However, it cannot be ruled out that 5S rDNA methylation is mediated by the small amount of siRNAs present in quadruple *dcl*-mutants. DCL1 loss-of-function mutants are lethal and biogenesis of siRNAs via residual DCL1 activity is not abolished in the *dcl1*-alleles that I used (Chapter 5, Schauer et al., 2002; Bouche et al., 2006). Furthermore, a small reduction in 5S rDNA methylation was found in *dcl3* using the

restriction enzyme *Hae* III, which reports on cytosine methylation in an asymmetric sequence context (Pontes et al., 2006). Methylation of *Hae* III and other such sites of 5S rDNA arrays should therefore be tested in triple and quadruple *dcl*-mutants to better understand the function of siRNAs in methylating asymmetric cytosines.

Studies of transgenes targeted by RNA silencing demonstrated that RdDM initiates methylation at cytosines in all sequence contexts (Pelissier et al., 1999; Aufsatz et al., 2002; Cao et al., 2003). However, propagation of CpG, CpNpG and asymmetric methylation over cell divisions depends on at least three different DNA methyltransferases (Finnegan and Kovac, 2000). These enzymes have specialized functions in maintenance and *de novo* DNA methylation (Matzke et al., 2005). In *Arabidopsis*, protection of 5S rDNA arrays from *Hpa* II digestion appears to be largely due to CpG-methylation, i.e., of the inner cytosine in the sequence CCGG; to a lesser extent, the outer cytosine of CCGG, a CpNpG site, is also methylated (Vongs et al., 1993; Fulnecek et al., 2002). Once established, DNA methylation in these two types of symmetric site is maintained by MET1 and CMT3, respectively (Bartee et al., 2001; Lindroth et al., 2001; Saze et al., 2003). In contrast, the *de novo* methyltransferase DRM2, which has been linked to RdDM and the Pol IV pathway (Cao et al., 2003; Onodera et al., 2005), is crucial for methylation of asymmetric sites (Cao and Jacobsen, 2002). I assayed a combination of CpG and CpNpG sites using *Hpa* II, and found that 5S rDNA is -- at least in part -- methylated via the Pol IV pathway. Given the above considerations, my working hypothesis is that the low level of siRNA biogenesis in quadruple *dcl*-mutants is sufficient to program RdDM of *Hpa* II sites (methylation largely maintained by MET1 and CMT3) but not sufficient to sustain methylation of asymmetric sites, which requires siRNAs generated by the Pol IV pathway to perpetually reestablish cytosine methylation after cell division.

9.1.3 Genetic interactions of DDM1 and the Pol IV pathway in 5S rDNA methylation.

Analysis of double mutants deficient for both DDM1 and the Pol IV pathway (Chapter 6) showed that while DCL3 deficiency was not associated with reduced 5S rDNA methylation of *Hpa* II sites, the combination of DCL3 and DDM1 deficiencies had an additive effect. That is, *dcl3 ddm1* plants showed even more hypomethylation relative to wild type than did *ddm1* plants. Moreover, the *ddm1 nrpd1a* and *ddm1 rdr2* combinations also showed more hypomethylation than *ddm1*. Thus, DDM1 and Pol IV pathway functions are not epistatic with respect to methylation in *Hpa* II sites of 5S rDNA repeats: two distinct processes seem to account for this methylation. These findings are parsimonious with my model, in which maintenance of CpG-methylation via MET1/DDM1, and RdDM via the Pol IV

pathway are overlapping but distinct aspects of 5S rDNA methylation (**Figure 22**). Following this view, DCL3 function in RdDM -- and that of its putative siRNA products -- would be particularly apparent when maintenance methylation is impaired (in *met1* and *ddm1*) or ineffectual (at asymmetric sites).

Based on their observations of *FWA* gene regulation, Chan *et al.* (2006a) proposed that RdDM is reinforced by pre-existing DNA methylation in a feedback system that enhances the stability of silencing. My work indicates that 5S rDNA genes are subject to a different mechanism, in which siRNA accumulation via the Pol IV pathway is enhanced by reduced CpG-methylation. Hundreds of distinct 5S rRNA genes showing heterogeneous transcriptional activity are found on three different *Arabidopsis* chromosomes (Campell et al., 1992; Cloix et al., 2000; Cloix et al., 2002). To understand how the Pol IV pathway influences particular 5S rRNA gene arrays, one could compare large-scale smRNA sequencing data to the sequences of different 5S rDNA repeat units, and thereby assess their relative contributions to the siRNA pool (Rajagopalan et al., 2006; Kasschau et al., 2007). This could, in turn, be compared to cytosine methylation profiles for particular repeat units with similar sequence polymorphism. One could thus test whether RNA silencing has disparate effects on different 5S rDNA repeat units, depending on their respective levels of CpG-methylation.

9.1.4 Speculation: Regulatory overlap between CpG-methylation and the Pol IV pathway.

The reduced weight of double mutants deficient for DDM1 and steps in the Pol IV pathway (Chapter 6) suggests that these factors regulate genes that impact plant growth. The relatively mild phenotypes in individual DDM1 and Pol IV pathway-deficient mutants imply that such regulation would involve both silencing pathways. Interestingly, several studies have uncovered effects of DNA methylation changes and RNA silencing on the plant response to abiotic and biotic stress (Stokes et al., 2002; Borsani et al., 2005; Katiyar-Agarwal et al., 2006; Sunkar et al., 2007). In particular, Stokes *et al.* (2002) recovered an epiallele called *bal* from self-fertilized lineages of *ddm1*. The *bal* phenotype, which includes growth deficiency similar to what I observed in the double mutant lines, is linked to hypomethylation and overexpression on an *R*-like gene, and may be due to constitutive expression of pathogen response genes (Stokes et al., 2002; Stokes and Richards, 2002). Theoretically, DNA methylation changes may be induced under stress conditions as an adaptive response (Jablonka and Lamb, 1989; Finnegan, 2002; Arnholdt-Schmitt, 2004; Madlung and Comai, 2004). In fact, common stress conditions, such as pathogen infection

and low temperature, do cause hypomethylation of protein-coding genes and DNA repeat elements in *Arabidopsis* and tobacco (Takeda et al., 2001; Wada et al., 2004; Pavet et al., 2006; Boyko et al., 2007; Choi and Sano, 2007). Although speculative, the type of mechanism that I proposed for 5S rDNA methylation (**Figure 22**) could regulate protein coding genes after stress exposure and ensuing genomic hypomethylation by reestablishing epigenetic marks via RdDM. Precedents for RdDM triggered in the wake of global hypomethylation include the *superman* and *bonsai* epialleles, which cause developmental phenotypes linked to reorganization of methylation patterns in the *Arabidopsis* genome (Jacobsen and Meyerowitz, 1997; Saze and Kakutani, 2007).

9.2 Viral siRNA biogenesis during DNA virus infection

Understanding the RNA silencing response of plants to model viruses, such as CaLCuV and CaMV, is economically significant because related viruses like *Africa cassava mosaic virus* (*Geminiviridae*), and *Banana streak virus* (*Caulimoviridae*) cause severe crop losses in Africa, Asia and Latin America (Agrios, 1997; Harper et al., 2002; Legg and Fauquet, 2004). A key feature of DNA virus infection in plants, is the accumulation of multiple size classes of viral siRNAs. Having noticed similarities between these siRNA species and the tandem repeat-derived siRNAs described above, I analyzed the same *Arabidopsis* mutant panel after inoculation with DNA viruses, in order to dissect the genetic requirements for viral siRNA biogenesis (Chapters 7 and 8). The mechanisms of DNA virus replication allowed me to test whether viral siRNA biogenesis requires specific host RDRs, without interference from a viral RdRP like those required for RNA virus replication (Dalmay et al., 2000; Hull, 2004). In addition, DNA viruses that infect mammals have been shown to express miRNAs (Pfeffer et al., 2004; Sullivan et al., 2005), but it is not clear whether stem-loop hairpin structures from plant virus transcripts can be processed like miRNAs. The CaMV 35S RNA provides a case study for the role of DCL1 and viral transcript structure in viral siRNA biogenesis.

9.2.1 *Arabidopsis* DCLs determine the size class of viral siRNAs

Production of smRNA duplexes requires upstream substrates that form dsRNA or hairpin RNA molecules suitable for DCL activities, which then generate different size classes of smRNAs (Matzke et al., 2001; Ambros and Chen, 2007). In *Arabidopsis*, repeat-associated siRNA biogenesis is thought to require Pol IVa and RDR2 for production of dsRNA that is subsequently processed by DCL3, whereas biogenesis of *trans*-acting siRNAs is thought to require RDR6 to produce dsRNA for processing by DCL4 (Pontes et al., 2006; Vaucheret, 2006). I did not find an obvious contribution of the putative RNA polymerases RDR2, RDR6 or POL IVa to DNA virus-derived siRNA biogenesis. Recent results suggest that mutants deficient in RDR1, RDR3, RDR4, and RDR5 do not show altered accumulation of CaLCuV::*ChlI* or CaMV siRNAs (personal communication, F. Vazquez and M. Pooggin). Furthermore, although RDR1 is inducible by virus infection in *Arabidopsis*, it was not found to be required for biogenesis of siRNAs derived from most viruses that have been tested (Yu et al., 2003; Xie et al., 2004; Deleris et al., 2006). Therefore, in contrast to repeat-associated and *trans*-acting siRNAs, it seems likely that no member of the RDR family is individually required for viral siRNA biogenesis.

DNA virus infection results in 21-24 nt siRNA production matching viral protein coding and non-coding regions. These viral siRNAs accumulate in the same DCL-determined manner as the equivalent size classes of endogenous species, while not utilizing upstream parts of endogenous siRNA pathways. DCL2 appears to mediate production of 22 nt siRNAs, DCL3 production of ~24 nt siRNAs and DCL4 production of 21 nt siRNAs (model, **Figure 23**). This suggests that *Arabidopsis* DCLs are versatile and size-specific enzymes, which process various dsRNA substrates. Furthermore, it supports a "branched" but generic pathway model, in which RNA substrates from different sources are funneled through the same DCL step(s), culminating with HEN1 methylating smRNA 3'-ends, as illustrated in my work for DNA viruses, and shown earlier for endogenous smRNAs (Boutet et al., 2003; Yu et al., 2005; Yang et al., 2006).

The relative contributions of DCL1 and DCL4 to viral siRNA biogenesis differed between CaLCuV and CaMV. During CaLCuV infection, accumulation of 21 nt viral siRNAs depended primarily on DCL4, whereas during CaMV infection the accumulation of 21 nt viral siRNAs was largely independent of DCL4. My data suggest that DCL1 mediates siRNA production from the CaMV 35S RNA leader in wild-type plants, and from CaLCuV dsRNA in triple mutant *d234* plants. In contrast to DNA viruses, dsRNA from cytoplasmic RNA viruses appears to be primarily targeted by DCL4 and DCL2, but rarely DCL3, resulting in a pattern of ~21-22 nt siRNA accumulation (Deleris et al., 2006; Fusaro et al., 2006; Diaz-Pendon et al., 2007). Thus, RNA substrates produced during infection with diverse viruses are differentially targeted by subsets of all four *Arabidopsis* DCLs. A possible explanation for abundant 24 nt viral siRNAs during DNA virus infection is the replication and transcription of their genomes in the nucleus, where DCL3 seems to localize and function (Xie et al., 2004; Pontes et al., 2006).

Whereas RNA viruses must produce sense and antisense copies of their genome to replicate, only sense transcripts are required for DNA virus replication (Hull, 2004). Therefore, Voinnet (2001) proposed two mechanisms for the production of viral dsRNAs from DNA viruses: (i) the synthesis of "unintended", overlapping sense/antisense transcripts around the circular viral genome, or (ii) the activity of host-encoded RdRPs. Geminivirus and caulimovirus infections both give rise to sense and antisense siRNAs (Chapters 7 and 8), which implies the existence of longer dsRNA precursors. I observed higher molecular weight RNAs of both sense and antisense polarity in *d234* plants infected with the DNA viruses, which may represent longer dsRNA intermediates. Given that accumulation of DNA virus-derived siRNAs does not appear to be dependent on known RDRs or Pol IVa, the above results support the first mechanism proposed by Voinnet, namely, that overlapping RNA polymerase II transcripts are the source of dsRNA for viral

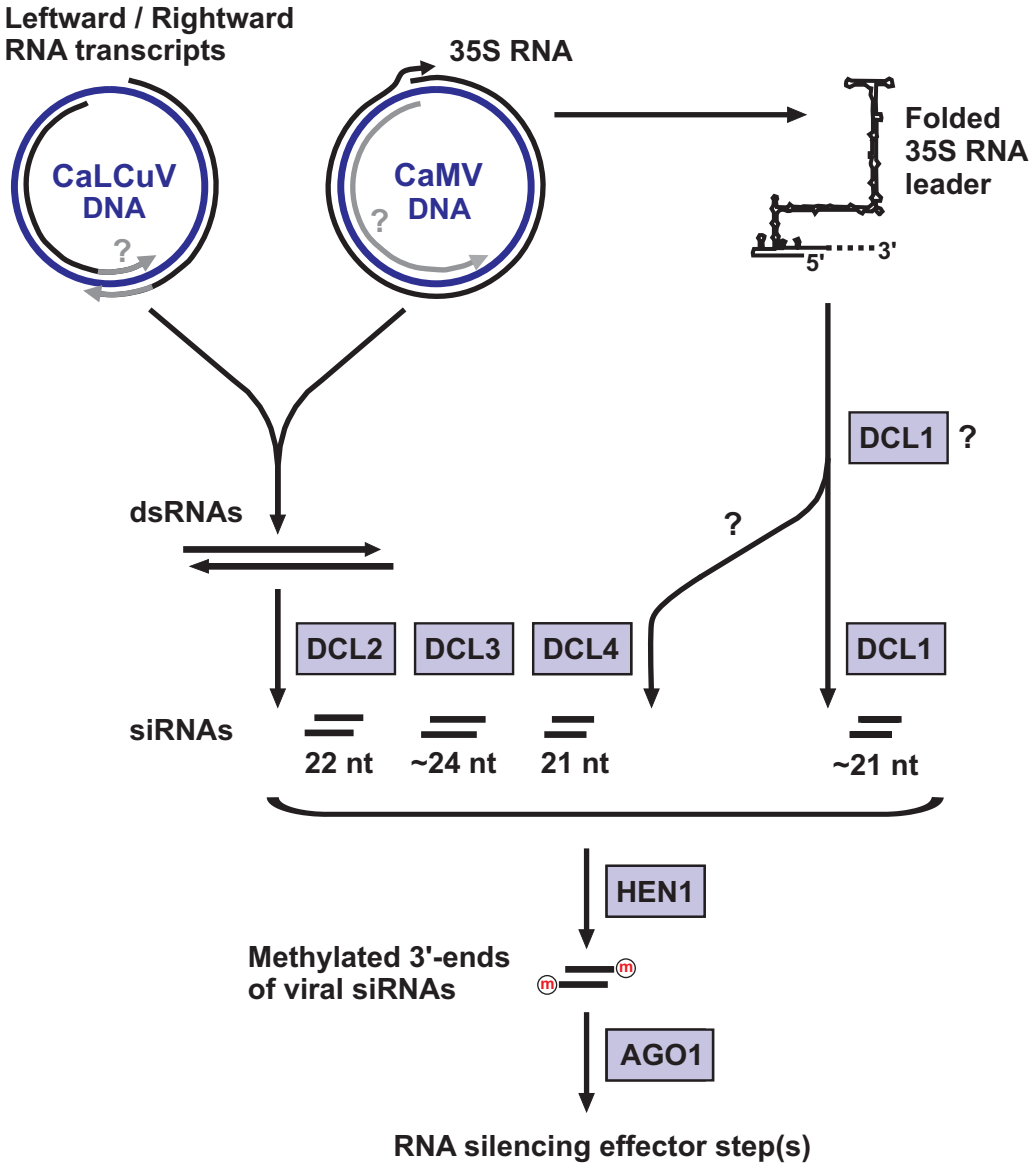


Figure 23. Model for DNA virus-derived siRNA biogeneis. Accumulation of CaLCuV and CaMV-derived siRNAs does not require any particular known RDR in *Arabidopsis*. However, overlapping sense / antisense transcription from the circular viral genome could be an RDR-independent mechanism for viral dsRNA biogenesis. Production of sense and antisense viral siRNAs is apparently mediated by DCL2, DCL3 and DCL4 during DNA virus infection. Biogenesis of 21 nt CaMV siRNAs appears to be mediated by DCL1 and/or DCL4, depending on the genomic region from which they arise. The folded CaMV 35S RNA leader is a potential substrate for DCL1. Alternatively, dsRNA molecules derived from the same genomic region could be a source of DCL1 products. Based on the accumulation of viral siRNAs in *hen1* mutant plants, their 3'-ends are likely methylated by HEN1 in the wild type. Regions of overlapping sense / antisense transcription in the diagram are purely schematic.

siRNA biogenesis (**Figure 23**). It should be pointed out, however, that this conclusion is based on the assumption that there is no substantial redundancy within the RDR family for viral siRNA biogenesis. The latter concern could be resolved, at least in part, by testing double and triple mutants deficient for RDR1, RDR2 and RDR6, which have established functions in RNA silencing (Ding and Voinnet, 2007). In addition, the model could be supported by detection and cloning of read-through transcripts in CaLCuV-infected plants.

There is evidence that overlapping sense and antisense transcripts are produced during DNA virus infection in plants. Bidirectional RNA polymerase II (Pol II) promoters in geminiviruses normally generate converging left- and rightward transcripts (Hanley-Bowdoin et al., 2000; Shivaprasad et al., 2005). Thus, viral transcripts could extend beyond their overlapping polyadenylation signals prior to cleavage and polyadenylation. Degradation products of read-through transcripts have been detected for the geminivirus promoter region (Shivaprasad et al., 2005). Although bi-directional transcription has not been reported for caulimoviruses, the strong CaMV enhancers of 35S and 19S RNA promoters (Driesen et al., 1993) could drive Pol II transcription in the antisense direction. Indeed, bidirectional transcription driven by the CaMV 35S core promoter was documented for a reporter system in *Arabidopsis* (Xie et al., 2001a).

In conclusion, some enzymatic steps, including production of specific siRNA size classes by four DCL proteins, and siRNA methylation by HEN1, appear to be shared by viral and endogenous siRNA biogenesis pathways, whereas upstream steps of their biogenesis are different.

9.2.2 A role for DCL1 in processing viral RNAs

DCL1 was believed to function exclusively in the production of miRNAs from imperfect stem-loop hairpin precursors (Reinhart et al., 2002; Finnegan et al., 2003; Kurihara and Watanabe, 2004). My findings provide strong evidence that DCL1 also processes certain longer viral dsRNAs into siRNAs. I detected sense and antisense 21 nt siRNAs derived from the CaLCuV::*ChlI* genome in the triple mutant *d234*, which is deficient for DCL2, DCL3, and DCL4. Processing of CaLCuV dsRNA in *d234* appears to be less efficient than in infected wild-type plants, because larger viral RNA intermediates of sense and antisense polarity accumulated in addition to 21 nt siRNAs in infected *d234* plants. In a recent study using a *d234* mutant similar to mine, Henderson *et al.* (2006) observed accumulation of 21 nt siRNAs derived from the endogenous inverted repeat *IR71*, dispersed repeat *AtSN1* and tandem repeat *TR2558*. They interpreted this as evidence that DCL1 can process endogenous long hairpin RNAs and dsRNA (Henderson et al.,

2006). Our findings are consistent and support the hypothesis that DCL1 can process a diverse set of dsRNA or long hairpin RNAs of endogenous and viral origin. However, the abovementioned siRNAs were only observed as a weak signal in *d234* lines, while DCL3 and DCL4 appear to be primarily responsible for siRNA production from the corresponding substrates in wild-type plants (Gasciolli et al., 2005; Xie et al., 2005; Bouche et al., 2006; Henderson et al., 2006). This raises the question, would DCL1 ever be a "default" DCL for processing viral precursors, even when other DCLs are not mutated?

The structured 35S RNA leader of CaMV seems to be exactly such a case. Almost all viral siRNAs (94%) sequenced from infected wild-type *Arabidopsis* matched the leader, and a similar fraction of viral siRNAs (90%) were sequenced from that region using *d234* plants (Chapter 8). Furthermore, a large portion of leader-derived siRNAs are 21 nt long in the wild type, and show reduced accumulation in *dcl1-9*. Finally, alignment of leader-derived sequences to the CaMV genome (Chapter 8) identified three major hotspots for siRNA accumulation, which persist in the alignment of siRNAs sequenced from *d234*. Thus, the 35S RNA leader region is the principal source for viral siRNA biogenesis, and DCL1 appears to be sufficient for this processing, although one cannot rule out involvement of other, unidentified proteins with RNase III activity.

Whether plant DNA viruses, like some of their animal counterparts (Pfeffer et al., 2005; Sullivan et al., 2005), code for true miRNAs is an important question: the targeting of host transcripts by virus-encoded miRNAs, perhaps in order to reprogram host gene expression, would have implications for both viral and host evolution (Ding and Voinnet, 2007). Plant miRNAs function by targeting endogenous transcripts for cleavage (Llave et al., 2002). Interestingly, 21 nt smRNA species derived from a region of the CaMV 35S leader (annotated L2 in my work), were recently implicated in cleavage of an *Arabidopsis* mRNA called *RCC1* (Moissiard and Voinnet, 2006). The biological significance of this cleavage -- either for virus or host -- is unclear, since plants deficient for *RCC1* were neither more nor less susceptible to CaMV infection (Moissiard and Voinnet, 2006). But this is the first candidate for a virus-encoded miRNA that would regulate plant host gene expression and potentially effect the plant-pathogen interaction.

Although my study is the first to describe the sequencing of viral siRNAs from DNA virus-infected plants, smRNAs have been cloned from plants infected with RNA viruses. For example, analysis of tobacco infected with *Cymbidium ringspot tombusvirus* (CymRSV) led to the discovery of hotspots for viral siRNA accumulation, which showed non-random distribution along the RNA genome and correlated with structured RNA regions (Molnar et al., 2005). Genomic hotspots for viral siRNAs were also observed in *Brassica juncea* infected with Turnip mosaic virus (Ho et al., 2007). Together, these

studies show that hotspots for viral siRNA accumulation are not specific to CaMV infection. A tempting conjecture is that folded structures in the 35S RNA leader -- or similar structures of RNA virus genomes -- are recognized as miRNA precursor-like substrates by DCL1 and its orthologs in other plants. In the CaMV case, this could be tested using mutated constructs that produce infectious viruses with modified base-pairing in stem sections of the 35S RNA leader (Pooggin et al., 1998).

Moissiard and Voinnet (2006) argue, based on the overall lower titer of CaMV siRNAs detected in infected *dcl1-9* compared to the wild type, that DCL1 is not directly involved in processing leader-derived siRNAs, but facilitates their biogenesis by other DCLs. This hypothesis is not consistent with my data: (i) The abundance of 21 nt viral siRNAs was reduced compared to 22 / 24 nt signals from two different regions of the leader in infected *dcl1-9*. (ii) Infected quadruple *dcl*-mutants accumulate less 21 nt leader siRNAs than the *d234* triple mutant. (iii) The global pattern of 21 nt siRNAs matching the 35S RNA leader was qualitatively similar in *d234* and wild-type plants. These findings suggests that DCL1, rather than simply facilitating siRNA biogenesis by other DCLs, actually mediates production of a 21 nt class of viral siRNAs. Significantly, the predominant size class of *bona fide* miRNAs in *Arabidopsis* (Rajagopalan et al., 2006), which are thought to be DCL1 products (Kurihara and Watanabe, 2004), is also 21 nt.

A model to explain CaMV-derived smRNA biogenesis need not exclude a direct role for DCL1. In the endogenous miRNA pathway (**Figure 2A**), precursor transcripts appear to be first processed by DCL1, producing pre-miRNA hairpins, and then cut again by DCL1 to generate ~21 nt miRNAs (Kurihara and Watanabe, 2004). Thus, dual DCL1 functions in processing viral RNA might be expected. Based on these considerations, my model (**Figure 23**) incorporates multiple routes for CaMV-derived smRNA biogenesis: (i) direct processing of the 35S RNA by DCL1, (ii) subsequent processing of the 35S RNA leader by other DCLs, and (iii) an alternative route of dsRNA production by sense/antisense transcription about the circular viral genome, followed by processing by multiple DCLs. The third route reflects the fact that a significant portion of CaMV smRNAs are likely produced from dsRNA; sense and antisense strands were detected for each size class. Specific PCR amplification of sense and antisense cDNA species could be used to test this supposition, following methods applied to repeat-derived dsRNA in *Drosophila* and *Arabidopsis* (Aravin et al., 2001; Mette et al., 2005). CaMV-derived dsRNA should overaccumulate in *d234*, due to lack of processing by DCL2, DCL3 and DCL4.

In conclusion, viral siRNA biogenesis during plant DNA virus infection does not appear to require an RDR protein; instead, distinctive features of these viruses, including

overlapping sense/antisense transcription and folding of viral transcripts into secondary structures, might produce substrates for DCL activities.

9.2.3 Function of DNA virus-derived siRNAs

In current models for RNA silencing pathways in *Arabidopsis* (summarized, **Figure 2**), siRNAs program AGO-RISC complexes to cleave specific RNA transcripts, or presumably, guide histone modifications and DNA methylation to specific genomic sequences. My results show that the mutation *dcl4*, which causes deficiency for biogenesis of 21 nt siRNAs, impairs VIGS targeting the endogenous *CHLI* mRNA. This is consistent with the proposed function of DCL4 in RNA silencing induced by inverted-repeats (Dunoyer et al., 2005; Fusaro et al., 2006; Dunoyer et al., 2007), and its role in VIGS mediated by RNA viruses (Deleris et al., 2006). A partial loss-of-function mutant for AGO1 was previously reported to show delayed onset of VIGS targeting *CHLI* (Muangsan et al., 2004). Because AGO1 has an activity that can slice an mRNA target of miR165 (Baumberger and Baulcombe, 2005) and preferentially associates with 21 nt smRNAs (Qi et al., 2006), one can hypothesize that geminivirus-derived, 21 nt siRNAs program AGO1 to cleave the *CHLI* mRNA. Supporting this idea, viral siRNAs have been co-precipitated with AGO1 protein (Zhang et al., 2006a). Furthermore, siRNAs derived from CymRSV co-fractionate in two protein complexes that likely correspond to free AGO1, and partially or fully assembled RISC (Pantaleo et al., 2007). In conclusion, available evidence suggests that AGO1 functions in effector complexes that mediate VIGS and defense against viruses in *Arabidopsis*, which are programmed, at least in part, by 21 nt products of DCL4.

Using the CaLCuV::*ChlI* vector, I showed that establishment and/or maintenance of the chlorata VIGS phenotype requires DCL4 and RDR6. DCL1, DCL2, and DCL3 present in combinations or individually are apparently sufficient, in DCL4-deficient plants, to produce viral siRNAs and support reduction of *CHLI* transcript levels, but not to trigger the chlorata phenotype. One possible explanation is that DCL4-dependent, 21 nt siRNAs spread into the shoot apical meristem, from which viruses are usually excluded, to trigger *CHLI* silencing and the chlorata phenotype. Dunoyer *et al.* (2005) found that cell-to-cell spread of silencing triggered by inverted-repeat transgenes requires DCL4 and is correlated with the accumulation of 21 nt but not 24 nt siRNAs. They proposed that 21 nt siRNAs are short-range silencing signals and that long-range silencing signals require RDR6-mediated amplification. Although speculative, a cooperative role for DCL4 and RDR6 in VIGS spread is supported by observations of other groups using transgene silencing inducers (Dunoyer et al., 2005; Dunoyer et al., 2007; Smith et al., 2007).

It is widely reported that siRNA biogenesis pathways function to protect plants against infection by viruses (Ding and Voinnet, 2007). If this is true, blocking these pathways should increase plant susceptibility to infection, accumulation of viral RNA and/or viral symptom severity. Indeed, infection with an RNA virus, Cucumber Mosaic Virus (CMV), showed enhanced viral symptoms in *d234* compared to the wild type (Bouche et al., 2006; Fusaro et al., 2006; Diaz-Pendon et al., 2007). In addition, genomic RNA of the viruses TCV and CMV overaccumulated in infected double mutant *d24* plants compared to the wild type, implicating DCL2 and DCL4 products in antiviral defense (Deleris et al., 2006; Diaz-Pendon et al., 2007).

Surprisingly, mutations in RNA silencing-related genes – e.g., the triple mutant *d234* -- did not result in hypersusceptibility to CaLCuV::*ChlI* or CaMV infection (Chapters 7 and 8). Defense mediated by RNA silencing may be less efficient against DNA viruses, given that their genomes are not expected to be cleaved by siRNA-programmed RISCs. But a role for RNA silencing in defense against geminiviruses is implied by the existence of geminivirus-encoded silencing suppressors (Voinnet et al., 1999; Trinks et al., 2005), and the protection conveyed by transgenes expressing dsRNA homologous to geminiviruses (Pooggin and Hohn, 2004). More importantly, it is a feature of viral infection in plants, that silencing suppression can mask hypersusceptibility in RNA silencing-deficient mutants when compared to the wild type. For example, Diaz-Pendon *et al.* (2007) found the clearest difference in viral susceptibility when they compared wild-type and *d234* plants infected with a suppressor-deleted strain of CMV, and only a subtle difference using the suppressor-expressing CMV strain.

In future experiments, CaLCuV constructs deficient for the viral suppressor, AC2, could be used to test whether enhanced viral replication or symptoms occurs in *d234*. Until such data are available, a judgment on the function of DCL2, DCL3 and DCL4 in defense against geminivirus infection would be premature. In contrast, CaMV-infected *d234* plants accumulated higher titers of viral molecules, including the 35S RNA and viral coat protein, than infected wild-type or single *dcl*-mutant plants (Moissiard and Voinnet, 2006). That these differences were modest and not accompanied by enhanced CaMV symptoms, may be in part due to DCL1, which is present in *d234*. In conclusion, a combination of siRNA products from four DCLs may contribute to limit CaMV replication.

9.3 Concluding remarks

RNA silencing mediated by smRNAs has many functions in eukaryotes (Almeida and Allshire, 2005; Baulcombe, 2005). Understanding smRNA biogenesis is therefore of growing interest to diverse biological fields, including development, cell biology, epigenetics and virology (Meins et al., 2005; Ding and Voinnet, 2007; Lawrie, 2007).

I found that accumulation of specific size classes of endogenous and viral siRNAs depends on particular *Arabidopsis* DCLs: 22 nt siRNAs on DCL2, ~24 nt siRNAs on DCL3, and 21 nt siRNAs on DCL4 or DCL1. This is consistent with the finding that distinct activities produce 21 and 24 nt siRNAs in wheat germ extracts (Tang and Zamore, 2004), and suggests that each plant DCL processes dsRNA into a specific siRNA size class. The large repertoire of *Arabidopsis* siRNAs is probably typical for flowering plants, because the genomes of many angiosperms encode four or more DCLs (Margis et al., 2006).

In *Arabidopsis*, the Pol IV pathway is required for biogenesis of siRNAs derived from genomic tandem repeats, shown here for methylated 180 bp and 5S rDNA arrays. While DCL3 is an integral component of the Pol IV pathway, other members of the DCL family are partially redundant for production of repeat-derived siRNAs. The role of these siRNAs in 5S rDNA methylation seems to be sequence-context sensitive, being particularly important for asymmetric sites (Pontes et al., 2006). My working hypothesis is that methylation in symmetric sites does not require high siRNA titers, because it is independently maintained by CMT3 and MET1 in concert with DDM1. To quantify these differences, DNA from wild-type and quadruple *dcl*-mutant plants could be used for bisulfite sequencing analysis of 5S rDNA, as described by Fulnecek *et al.* (2002).

An epigenetic state maintained by MET1 and DDM1 -- perhaps hypermethylation of CpG sites -- represses accumulation of Pol IV-dependent siRNAs derived from 5S rDNA, suggesting a link between these distinct silencing mechanisms (**Figure 22**). The overlap of DDM1 and Pol IV contributions to DNA methylation may have biological functions: regulation of “minor” transcript titer in the 5S rRNA pool (Vaillant et al., 2007b), reestablishment of epigenetic marks after genomic hypomethylation (Saze and Kakutani, 2007), and control of transposable element activity (Lippman et al., 2003). In future experiments, it will be important to analyze single-copy genes in my double mutant series, to see whether DDM1 and Pol IV functions overlap to regulate their expression. One candidate is the R-like gene which is overexpressed in *bal* epialleles (Stokes et al., 2002); their phenotype is similar to that of my *ddm1 nrpd1a* and *ddm1 rdr2* lines.

In mammals, proteins with analogous function mediate DNA methylation of repeats, although there are also key differences from plant pathways. The mammalian

ortholog of DDM1, Lsh, is required for normal murine development and methylation of repetitive elements in the mouse genome (Geiman et al., 2001; De La Fuente et al., 2006). Although mice do not have a Pol IV pathway, a process involving 26-31 nt "piRNAs" and associated PIWI proteins silences transposable elements in testes (Aravin et al., 2007) and perhaps other tissues (Ro et al., 2007). Since the mechanism that targets Lsh to repeats is unknown (Huang et al., 2004), it would be interesting to test whether piRNA function or expression is linked to Lsh.

RNA silencing in various eukaryotes involves an RdRP (Wassenegger and Krczal, 2006). In fission yeast and *Tetrahymena*, the production of endogenous siRNAs by Dicer is physically and functionally coupled to RdRP proteins and dsRNA biogenesis (Colmenares et al., 2007; Lee and Collins, 2007). The plant Pol IV pathway, which is similar to RNAi in fission yeast (Pikaard, 2006), may involve such a coupling of RDR2 and DCL3. This would explain the bias of tandem repeat-derived siRNAs towards the 24 nt size class. But this coupling could not be absolute, since DCL3-mediated production of viral siRNAs occurs independently of upstream components from the Pol IV pathway.

During CaMV infection siRNAs accumulate from three major hotspots corresponding to two stem sections of the 35S RNA leader region. Data from *dcl*-mutants suggest that DCL1 mediates production of a significant fraction of these siRNAs, facilitating access for processing by other DCLs, whose siRNA products may limit CaMV replication (Chapter 8 and Moissiard and Voinnet, 2006). Conversely, viral siRNAs that regulate host gene expression may be amongst the abundant 35S RNA leader species (Moissiard and Voinnet, 2006). These preliminary investigations raise the intriguing possibility that expression of certain viral siRNAs (miRNAs?) enhance CaMV fitness, whereas other siRNAs derived from the same transcript restrict viral infection.

10. Acknowledgements

I would like to thank my adviser, Frederick Meins, Jr., and colleagues from his laboratory, who made this dissertation possible. Fred encouraged me to always look twice at experiments that “didn’t work,” and taught me that scientific communication -- conveying knowledge -- is just as important as generating data. I thank Azeddine Si-Ammour, Claudia Kutter and Franck Vazquez for their close collaboration over these years, and Estelle Arn, Monique Thomas, and Jérôme Ailhas for organizing the laboratory and introducing me to everything from planting seeds to bacterial transformation. Without Azeddine, Hampe Schöb and Zhenya Glazov, I would not have gotten my first project off the ground. My second research project emerged from a fruitful and ongoing scientific collaboration with Mikhail Pooggin and others in the Laboratory of Thomas Hohn. I owe Konstantina Boutsika the credit for showing me the proper way to cross *Arabidopsis*, and helping me impose order in my science. Michael Stadler, Michael Rebhan and Ed Oakeley introduced me to bioinformatics and fixed my Perl scripts. Alan Naylor and Sjoerd Van Eeden kept my IT running. Franck Vazquez, Fred Meins, Mikhail Pooggin, Thomas Hohn and Florian Mette gave detailed and critical comments on my experiments and this dissertation. The members of my thesis committee -- Fred Meins, Florian Mette, Thomas Hohn and Dirk Schübeler -- have been both patient and responsive during the materialization of this document. I thank Ortrun Mittelsten Scheid and Barbara Hohn for much helpful advice and support over these years. I thank the Friedrich Miescher Institute for Biomedical Research and the Botanical Institute of the Universität Basel. I am appreciative of my teachers from the Universität Basel, especially Prof. Thomas Boller for smoothing my transition into the Swiss academic system. I thank Stefan Höppner for his counsel and friendship from across the Rhine. My loving parents Marcia and Bill Blevins, and in-laws Xavier and Marie-Anne Muller, have always supported me, my career and growing family. This dissertation would not have seen the light of day without my dear wife Anne, whose steadfast support kept me focused. Finally, the births of my son Theodore and daughter Louise will forever be tied in space and time to this dissertation.

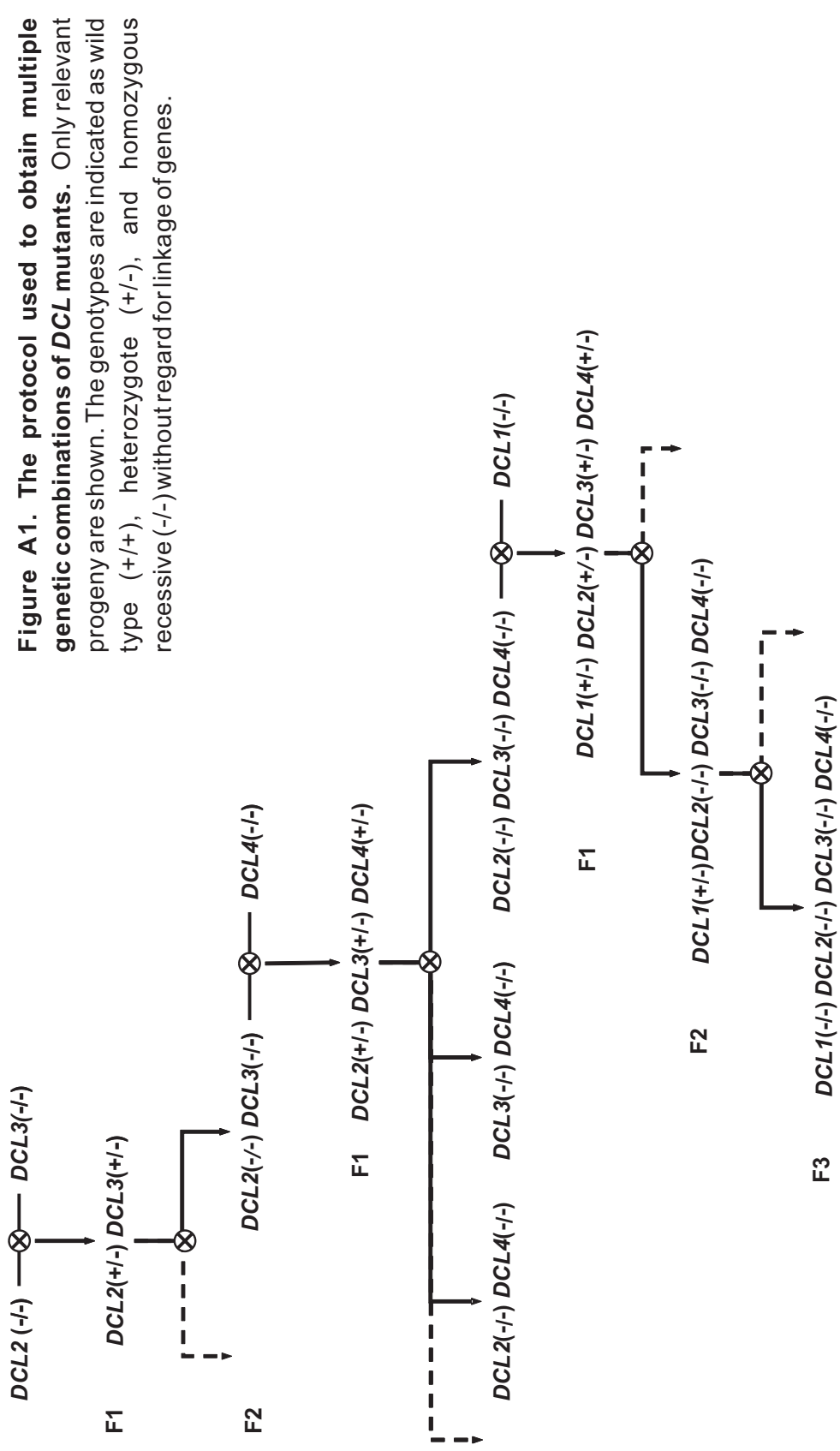


Figure A1. The protocol used to obtain multiple genetic combinations of *DCL* mutants. Only relevant progeny are shown. The genotypes are indicated as wild type (+/+), heterozygote (+/-), and homozygous recessive (-/-) without regard for linkage of genes.

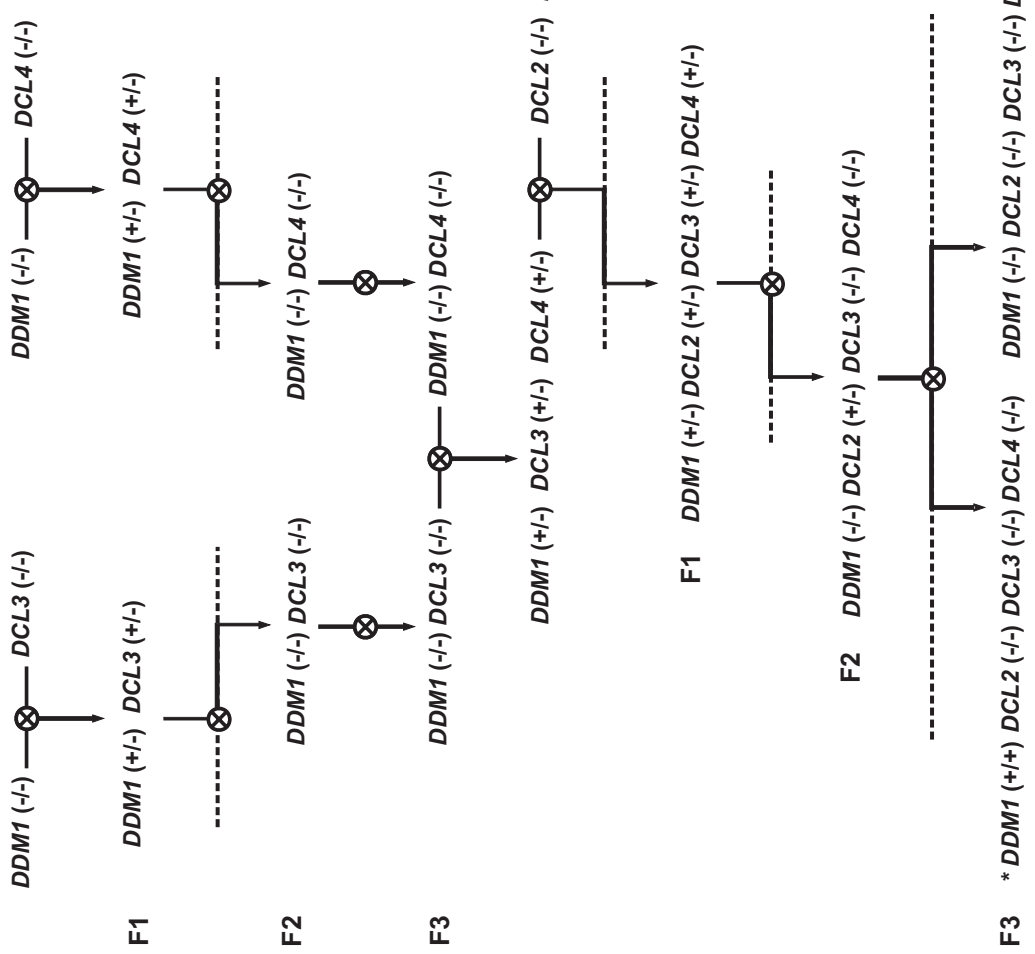


Figure A2. The protocol used to obtain a quadruple mutant combination deficient for *DDM1*, *DCL2*, *DCL3*, and *DCL4*. Only relevant progeny are shown. The genotypes are indicated as wild type (+/+), heterozygote (+/-), and homozygous recessive (-/-) without regard for linkage of genes.

Figure A3. Probe sequences for DNA blot hybridization

(i) 5S rDNA repeat (1x in pCRII plasmid):

- *Eco*RI restriction sites of the pCRII plasmid are shown in UPPERCASE letters.
- *Hpa*II restriction sites used for the DNA methylation analysis are in blue.
- Sequences corresponding to siRNAs cloned from *Arabidopsis* are shown in red.

GAATTCgcccttccctccgcgaaagagcaggacaatgtaacttctccattgacttttgtcgaccccaaattt
tgacctttaactactttttcgggcattttcgtgatttgagctatattacggacccaaaattacttgttcaagc
attgttttcgaatttttttcatgcatcaaagctcgttaagactagatggggtatccctacatagcgggtgggac
ccacggcgaatggttcatcaagtcttcaaaaaagaatatatacgattgcattgcatatactaacggatgtgat
cataccagcactaatgca**ccgg**atcccatcagaactccgcagttaagcgtgcttgggcgagagtagtactagg
atgggtgacctc**ccgg**gaagtcctcgtgttgcatccctcttttatgtttaacctttttttttttggttaaaac
tttatgactctataacttctataccgtgaggccaaacttggcatgtgatacctttcggaaagcccaaagaca
gaagggc**GAATTC**

(ii) 180-bp satellite repeats (2x in pSK plasmid):

- *Hind*III restriction sites from the satellite repeats are shown in UPPERCASE.
- *Hpa*II restriction sites used for the DNA methylation analysis are in blue.
- Sequences corresponding to siRNAs cloned from *Arabidopsis* are shown in red.

AAGCTTtcatggtgtagccaaagtccatatgagtctttggctttgtgtcttctaacaaggaaacactacttag
gcttataagat**Ccgg**ttgcgatttatgttcttatactcaatcatacacatgagatcaagtcatattcgactcc
aaaacactaaccaaccttcttcttgcttctca**AAGCTT**tcatggtgtagccaaagtccatatgagtctttggc
tttgtgtcttctaacaaggaaacactacttaggcttataagat**Ccgg**ttgcgatttatgttcttatactcaat
catacacatgagatcaagtcatattcgactccaaaacactaaccaaccttcttcttgcttctc**AAGCTT**

Table A1. Probes for RNA blot hybridization

Probe name	Detects	Sequence
CaLCuV AC2 s	viral transcript sense	5'-TGGAGGAAGATAGAACACCCGCAGTTC-3'
CaLCuV AC2 as	antisense polarity	5'-GAACTGCGGGTGTTCTATCTTCCTCCA-3'
CaLCuV AC4 s	viral transcript sense	5'-TGGTGATGTAATTCTTGACGGCATTGGTGTCT-3'
CaLCuV AC4 as	antisense polarity	5'-AGACACCAATGCCGTCAAGAATTACATCACCA-3'
CaLCuV *ChlI* s	viral transcript sense	5'-AATATGGTTGATCTTCCTTTGGGTGCAACAG-3'
CaLCuV *ChlI* as	antisense polarity	5'-CTGTTGCACCCAAAGGAAGATCAACCATATT-3'
CaMV L1 as	viral transcript sense	5'-GAAACCCTATAAGAACCCTAATTCCCTTATCTGGGAACTACTC-3'
CaMV L1 s	antisense polarity	5'-GAGTAGTTCCCAGATAAGGGAATTAGGGTTCTTATAGGGTTTC-3'
CaMV L2 as	viral transcript sense	5'-CTTCCTTGTCTTCCTCCTTC-3'
CaMV L2 s	antisense polarity	5'-GAAGGAGGAAGACAAGGAAG-3'
CaMV L3 as	viral transcript sense	5'-AGTCGTCTCGTGTCTGGTTTATAT-3'
CaMV L3 s	antisense polarity	5'-ATATAAACCAGACACGAGACGACT-3'
CaMV TAV as	viral transcript sense	5'-TCAACGGATTTGTTGATTCTTTACCAGGAGCCGTTTGCTCTGG-3'
CaMV TAV s	antisense polarity	5'-CCAGAGCAAACGGCTCCTGGTAAAGAATCAACAAATCCGTTGA-3'
miR173 as	guide strand	5'-GTGATTTCTCTCTGCAAGCGAA-3'
miR173* as	passenger strand	5'-CTTTCGCTTACACAGAGAATC-3'
siR1003 as	sense strand	5'-ATGCCAAGTTTGGCCTCACGGTCT-3'
siR1003 s	antisense strand	5'-AGACCGTGAGGCCAAACTTGGCAT-3'
siR255 as	guide strand	5'-TACGCTATGTTGGACTTAGAA-3'
siR255* as	passenger strand	5'-TATTCTAAGTCCAACATAGCG-3'
U6-I		5'-GGCCATGCTAATCTTCTCTGTATCGTT-3'
U6-II		5'-CCAATTTTATCGGATGTCCCCGAAGGGAC-3'

Table A2. Primer sequences for genotyping *Arabidopsis* strains

Primer name	Sequence	Predicted T_m (°C)
dcl1_geno_F1	5'-TAATGCGGGAGAGGATCAAGG-3'	56.4 ºC
dcl1_geno_R1	5'-AGTAGAAGAACCGCAGCTGAATC-3'	56.9 ºC
LB-pGV_CAF1_2	5'-GACACACACATCATCTCATTGATGCTTGG-3'	59.6 ºC
dcl2_geno_F2	5'-AACCCAAAGGTTGGATGGTTCCAGG-3'	61.4 ºC
dcl2_geno_R2	5'-AAGATATATCGGATTCCCATCCGAAAGTGG-3'	59.5 ºC
dcl3-1_geno_F1	5'-TGTCTATCGACTACAGAGAGATTTTGCATTGC-3'	59.7 ºC
dcl3-1_geno_R1	5'-TGCGGCAAATACACCCCAATGG-3'	60.7 ºC
nrpd1_geno_F1	5'-TGTTCATTCAGTTACAAGTCGTG-3'	53.2 ºC
nrpd1_geno_R1	5'-TTCACCTTTTTGATCCCTTGATC-3'	53.3 ºC
rdr2-1_geno_F1	5'-ATGGTGTCAGAGACGACGACG-3'	58.9 ºC
rdr2-1_geno_R1	5'-CAGAAGCGTCACCATTAACACAAC-3'	56.6 ºC
rdr2-2_geno_F1	5'-CTGATCGCGAGATTTCAGTTC-3'	53.6 ºC
rdr2-2_geno_R1	5'-AGAAGATTGGAGCAAGCTTCC-3'	55.1 ºC
LB-new_SALK	5'-GACTCTTGTTCCAAACTGGAACAACACTCAACC-3'	62.1 ºC
LBb1_SAIL	5'-TAGCATCTGAATTTCATAACCAATCTCGATACA-3'	60.0 ºC
dcl3-2_geno_F1	5'-GTCAGAGATCACTAAGAGGCTTATCAAAGC-3'	58.4 ºC
dcl3-2_geno_R1	5'-TAGTTTCAGGAATGCACTTCTAAGGACATCC-3'	59.6 ºC
dcl4_geno_F1	5'-GGCTGCACAGCTGATGATTACAA-3'	57.9 ºC
dcl4_geno_R1	5'-GCCGCTCGAGATCATCAGCAAAGGAAT-3'	62.6 ºC
GABI_8409A	5'-TCTCCATATTGACCATCATACTCATT-3'	53.7 ºC
ddm1-2_F1	5'-GCTGGAAGGGAAAGCTTAACAACC-3'	58.4 ºC
ddm1-2_R1	5'-ACACTGCCATCGATTCTGCAAACC-3'	60.1 ºC

File A1. Perl script for BLAST filtering and smRNA size distributions

```
#!/bin/perl

# INPUT: Requires a tab-delimited BLAST report, with smRNA labels
# in format [Condition-Genotype-ID-Value-Size]

my $current_smRNA = '';
my @fwd = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0);
my @rev = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0);

my %tab_size = ("15"=>0,
                "16"=>1,
                "17"=>2,
                "18"=>3,
                "19"=>4,
                "20"=>5,
                "21"=>6,
                "22"=>7,
                "23"=>8,
                "24"=>9,
                "25"=>10,
                "26"=>11,
                "27"=>12);
   while(<>) {
         if (/^(\w+)-(\w+)-(\d+)-(\d+)-(\d+)/) {

              chomp;
              my @hit = split(/\t/);
              my $value = $4;
              my $class = $5;

              if ($current_smRNA ne $3 && $class == $hit[3]) {

                  $current_smRNA = $3;

                  print "$_\t$class\t$value\n";

                  if ($hit[8] < $hit[9]) {

                     $fwd[$tab_size{"$class"}] =
                     $fwd[$tab_size{"$class"}] + $value;

                  }

                  else {

                     $rev[$tab_size{"$class"}] =
                     $rev[$tab_size{"$class"}] + $value;

                  }
              }
          }
      }

print "# Matches to the forward strand = ";
print join ("\t",@fwd);
print "\n";
print "# Matches to the reverse strand = ";
print join ("\t",@rev);
```

File A2. Perl script for Mapping siRNAs to the CaMV genome

```
#!/bin/perl

# INPUT: Requires tab-delimited BLAST report, processed by File A1.

# This hash serves to select the appropriate column of an
# array of arrays data structure (the @smRNA_matrix).
# The sign indicates sense or antisense orientation of the
# BLAST hit with respect to the database, whereas the
# number is the size class of the smRNA being counted.

my $bin_num = 8000;     # number of bins (height of matrix);
my $bin_width = 1;      # width (size in bp) of bins
my $offset = 0;   # offset from 0 in bins

my %graph =("+20"=>0,
            "+21"=>1,
            "+22"=>2,
            "+23"=>3,
            "+24"=>4,
            "+25"=>5,
            "-20"=>6,
            "-21"=>7,
            "-22"=>8,
            "-23"=>9,
            "-24"=>10,
            "-25"=>11);

# Initialization of the $bin_num by 12 @smRNA_matrix to 0 values:
# This matrix will store counts of smRNA starting in particular
# bins ($bin_width bp each) of the database seq, according to their
# size class as cloned (20-25 nt).

my @smRNA_matrix;

   for (my $i = $offset; $i < ($bin_num + $offset); $i++) {

   @smRNA_matrix[$i] = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0];

   }

# This routine reads the blast_hit_table.data file,
# extracts important fields from the tabular format into @hit
# and proceeds to populate the @smRNA_matrix with tallies of
# cloned smRNAs corresponding to bins with size classes and
# sense/antisense orientation being recorded.

while(<>) {

   chomp $_;
   my @hit = split(/\t/);
   $start = $hit[3];

   for (my $i = $offset; $i < ($bin_num + $offset); $i++) {

       my $bin = $i * $bin_width;

       my $strand_size = "$hit[8]"."$hit[5]";
```

```
        if ( ($start >= $bin) && ($start < ($bin + $bin_width)) ) {

        $smRNA_matrix[$i][$graph{$strand_size}] += $hit[6];

        }
    }
}

# This routine prints out the @smRNA_matrix in a tabular format
# that is easily converted to a graph by Excel, etc.

    for (my $i = $offset; $i < ($bin_num + $offset); $i++) {

        for (my $j = 0; $j < 12; $j++) {

             print "$smRNA_matrix[$i][$j]\t";
        }
        print "\n";
    }
```

12. References

Abouzid, A., Hiebert, E., and Strandberg, J. (1992). Cloning, identification and partial sequencing of a new geminivirus infecting Brassicaceae. Phytopathology **82,** 1070.

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.

Agrios, G.N. (1997). Plant Pathology. (San Diego: Academic Press).

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.

Al-Kaff, N.S., Covey, S.N., Kreike, M.M., Page, A.M., Pinder, R., and Dale, P.J. (1998). Transcriptional and posttranscriptional plant gene silencing in response to a pathogen. Science **279,** 2113-2115.

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.

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.

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.

Almeida, R., and Allshire, R.C. (2005). RNA silencing and genome regulation. Trends Cell Biol **15,** 251-258.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science **301,** 653-657.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol **215,** 403-410.

Ambros, V., and Chen, X. (2007). The regulation of genes and genomes by small RNAs. Development.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). A uniform system for microRNA annotation. Rna. **9,** 277-279.

Ammons, D., Rampersad, J., and Fox, G.E. (1999). 5S rRNA gene deletions cause an unexpectedly high fitness loss in Escherichia coli. Nucleic Acids Res **27,** 637-642.

Anandalakshmi, R., Pruss, G.J., Ge, X., Marathe, R., Mallory, A.C., Smith, T.H., and Vance, V.B. (1998). A viral suppressor of gene silencing in plants. Proc Natl Acad Sci U S A **95,** 13079-13084.

***Arabidopsis*_Genome_Initiative.** (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408,** 796-815.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science **316,** 744-747.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr Biol **11,** 1017-1027.

Arnholdt-Schmitt, B. (2004). Stress-induced cell reprogramming. A role for global genome regulation? Plant Physiol **136,** 2579-2586.

Aufsatz, W., Mette, M.F., Matzke, A.J., and Matzke, M. (2004). The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol **54,** 793-804.

Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, A.J., and Matzke, M. (2002). RNA-directed DNA methylation in Arabidopsis. Proc Natl Acad Sci U S A **99 Suppl 4,** 16499-16506.

Axtell, M.J., and Bartel, D.P. (2005). Antiquity of microRNAs and their targets in land plants. Plant Cell **17,** 1658-1673.

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.

Bartee, L., Malagnac, F., and Bender, J. (2001). Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes Dev **15,** 1753-1758.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. **116,** 281-297.

Baulcombe, D. (2004). RNA silencing in plants. Nature **431,** 356-363.

Baulcombe, D. (2005). RNA silencing. Trends Biochem Sci **30,** 290-293.

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.

Beclin, C., Berthome, R., Palauqui, J.C., Tepfer, M., and Vaucheret, H. (1998). Infection of tobacco or Arabidopsis plants by CMV counteracts systemic post-transcriptional silencing of nonviral (trans)genes. Virology **252,** 313-317.

Bendahmane, A., Kanyuka, K., and Baulcombe, D.C. (1999). The Rx gene from potato controls separate virus resistance and cell death responses. Plant Cell **11,** 781-792.

Bendahmane, A., Kohn, B.A., Dedi, C., and Baulcombe, D.C. (1995). The coat protein of potato virus X is a strain-specific elicitor of Rx1-mediated virus resistance in potato. Plant J **8,** 933-941.

Bender, J. (2004). DNA methylation and epigenetics. Annu Rev Plant Biol **55,** 41-68.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature **447,** 407-412.

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.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev **16,** 6-21.

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.

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.

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.

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.

Boyko, A., Kathiria, P., Zemp, F.J., Yao, Y., Pogribny, I., and Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res **35,** 1714-1725.

Brigneti, G., Voinnet, O., Li, W.X., Ji, L.H., Ding, S.W., and Baulcombe, D.C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in Nicotiana benthamiana. Embo J **17,** 6739-6746.

Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. Trends Genet **22,** 268-280.

Brzeski, J., and Jerzmanowski, A. (2003). Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J Biol Chem **278,** 823-828.

Campell, B.R., Song, Y., Posch, T.E., Cullis, C.A., and Town, C.D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of Arabidopsis thaliana. Gene **112,** 225-228.

Cao, X., and Jacobsen, S.E. (2002). Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. Curr Biol. **12,** 1138-1144.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol. **13,** 2212-2217.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. **16,** 2733-2742.

Cedar, H., Solage, A., Glaser, G., and Razin, A. (1979). Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI. Nucleic Acids Res **6,** 2125-2132.

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.

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.

Chan, S.W., Zhang, X., Bernatavichute, Y.V., and Jacobsen, S.E. (2006a). Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol **4,** e363.

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.

Chan, S.W., Henderson, I.R., Zhang, X., Shah, G., Chien, J.S., and Jacobsen, S.E. (2006b). RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in arabidopsis. PLoS Genet **2,** e83.

Choi, C.S., and Sano, H. (2007). Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. Mol Genet Genomics **277,** 589-600.

Church, G.M., and Gilbert, W. (1984). Genomic sequencing. Proc Natl Acad Sci U S A **81,** 1991-1995.

Cloix, C., Tutois, S., Mathieu, O., Cuvillier, C., Espagnol, M.C., Picard, G., and Tourmente, S. (2000). Analysis of 5S rDNA arrays in Arabidopsis thaliana: physical mapping and chromosome-specific polymorphisms. Genome Res **10,** 679-690.

Cloix, C., Tutois, S., Yukawa, Y., Mathieu, O., Cuvillier, C., Espagnol, M.C., Picard, G., and Tourmente, S. (2002). Analysis of the 5S RNA pool in Arabidopsis thaliana: RNAs are heterogeneous and only two of the genomic 5S loci produce mature 5S RNA. Genome Res **12,** 132-144.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature **399,** 166-169.

Collins, R.E., and Cheng, X. (2005). Structural domains in RNAi. FEBS Lett **579,** 5841-5849.

Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol Cell **27,** 449-461.

Covey, S.N., Al-Kaff, N.S., Lángara, A., and Turner, D.S. (1997). Plants combat infection by gene silencing. Nature **385,** 781-782.

Crete, P., Leuenberger, S., Iglesias, V.A., Suarez, V., Schob, H., Holtorf, H., van Eeden, S., and Meins, F. (2001). Graft transmission of induced and spontaneous post-transcriptional silencing of chitinase genes. Plant J **28,** 493-501.

Dalmay, T. (2005). Virus-induced gene silencing. In Plant Epigenetics, P. Meyer, ed (Oxford: Blackwell), pp. 223-243.

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.

Dawe, R.K., and Henikoff, S. (2006). Centromeres put epigenetics in the driver's seat. Trends Biochem Sci **31,** 662-669.

de Carvalho, F., Gheysen, G., Kushnir, S., Van Montagu, M., Inze, D., and Castresana, C. (1992). Suppression of beta-1,3-glucanase transgene expression in homozygous plants. Embo J **11,** 2595-2602.

De La Fuente, R., Baumann, C., Fan, T., Schmidtmann, A., Dobrinski, I., and Muegge, K. (2006). Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells. Nat Cell Biol **8,** 1448-1454.

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.

Di Serio, F., Schob, H., Iglesias, A., Tarina, C., Bouldoires, E., and Meins, F., Jr. (2001). Sense- and antisense-mediated gene silencing in tobacco is inhibited by the same viral suppressors and is associated with accumulation of small RNAs. Proc Natl Acad Sci U S A **98,** 6506-6510.

Diaz-Pendon, J.A., Li, F., Li, W.X., and Ding, S.W. (2007). Suppression of Antiviral Silencing by Cucumber Mosaic Virus 2b Protein in Arabidopsis Is Associated with Drastically Reduced Accumulation of Three Classes of Viral Small Interfering RNAs. Plant Cell.

Ding, S.W., and Voinnet, O. (2007). Antiviral Immunity Directed by Small RNAs. Cell **130,** 413-426.

Douet, J., and Tourmente, S. (2007). Transcription of the 5S rRNA heterochromatic genes is epigenetically controlled in Arabidopsis thaliana and Xenopus laevis. Heredity.

Driesen, M., Benito-Moreno, R.M., Hohn, T., and Futterer, J. (1993). Transcription from the CaMV 19 S promoter and autocatalysis of translation from CaMV RNA. Virology **195,** 203-210.

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. Epub 2005 Nov 1356.

Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. Nat Genet **39,** 848-856.

Eickbush, T.H., and Eickbush, D.G. (2007). Finely orchestrated movements: evolution of the ribosomal RNA genes. Genetics **175,** 477-485.

Ekwall, K. (2004). The roles of histone modifications and small RNA in centromere function. Chromosome Res **12,** 535-542.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001a). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev **15,** 188-200.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001b). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature **411,** 494-498.

Erickson, F.L., Holzberg, S., Calderon-Urrea, A., Handley, V., Axtell, M., Corr, C., and Baker, B. (1999). The helicase domain of the TMV replicase proteins induces the N-mediated defence response in tobacco. Plant J **18,** 67-75.

Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A **97,** 11650-11654.

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.

Finnegan, E.J. (2002). Epialleles - a source of random variation in times of stress. Curr Opin Plant Biol **5,** 101-106.

Finnegan, E.J., and Kovac, K.A. (2000). Plant DNA methyltransferases. Plant Mol Biol **43,** 189-201.

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.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature **391,** 806-811.

Fojtova, M., Van Houdt, H., Depicker, A., and Kovarik, A. (2003). Epigenetic switch from posttranscriptional to transcriptional silencing is correlated with promoter hypermethylation. Plant Physiol **133,** 1240-1250.

Franck, A., Guilley, H., Jonard, G., Richards, K., and Hirth, L. (1980). Nucleotide sequence of cauliflower mosaic virus DNA. Cell **21,** 285-294.

Fulnecek, J., Matyasek, R., and Kovarik, A. (2002). Distribution of 5-methylcytosine residues in 5S rRNA genes in Arabidopsis thaliana and Secale cereale. Mol Genet Genomics **268,** 510-517.

Fulnecek, J., Matyasek, R., Kovarik, A., and Bezdek, M. (1998). Mapping of 5-methylcytosine residues in Nicotiana tabacum 5S rRNA genes by genomic sequencing. Mol Gen Genet **259,** 133-141.

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.

Gammelgard, E., Mohan, M., and Valkonen, J.P. (2007). Potyvirus-induced gene silencing: the dynamic process of systemic silencing and silencing suppression. J Gen Virol **88,** 2337-2346.

Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Luedi, M., Shepherd, R.J., and Messing, J. (1981). The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. Nucleic Acids Res **9,** 2871-2888.

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.

Geiman, T.M., Tessarollo, L., Anver, M.R., Kopp, J.B., Ward, J.M., and Muegge, K. (2001). Lsh, a SNF2 family member, is required for normal murine development. Biochim Biophys Acta **1526,** 211-220.

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A. (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science **297,** 1871-1873.

Golden, T.A., Schauer, S.E., Lang, J.D., Pien, S., Mushegian, A.R., Grossniklaus, U., Meinke, D.W., and Ray, A. (2002). SHORT INTEGUMENTS1/SUSPENSOR1/CARPEL FACTORY, a Dicer homolog, is a maternal effect gene required for embryo development in Arabidopsis. Plant Physiol **130,** 808-822.

Goldsbrough, P.B., Ellis, T.H., and Cullis, C.A. (1981). Organisation of the 5S RNA genes in flax. Nucleic Acids Res **9,** 5895-5904.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu Rev Biochem **74,** 481-514.

Goodwin, J., Chapman, K., Swaney, S., Parks, T.D., Wernsman, E.A., and Dougherty, W.G. (1996). Genetic and biochemical dissection of transgenic RNA-mediated virus resistance. Plant Cell **8,** 95-105.

Grishok, A., Sinskey, J.L., and Sharp, P.A. (2005). Transcriptional silencing of a transgene by RNAi in the soma of C. elegans. Genes Dev **19,** 683-696.

Grunweller, A., and Hartmann, R.K. (2005). RNA interference as a gene-specific approach for molecular medicine. Curr Med Chem **12,** 3143-3161.

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.

Gutierrez, C. (2000). DNA replication and cell cycle in plants: learning from geminiviruses. Embo J **19,** 792-799.

Hall, I.M., Noma, K., and Grewal, S.I. (2003). RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. Proc Natl Acad Sci U S A **100,** 193-198.

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. Embo J. **21,** 4671-4679.

Hamilton, A.J., and Baulcombe, D.C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. **286,** 950-952.

Hammond-Kosack, K., and Jones, J.D. (2000). Responses to Plant Pathogens. In Biochemistry and Molecular Biology of Plants, B.B. Buchanan, W. Gruissem, and R.L. Jones, eds (Rockville: American Society of Plant Physiologists), pp. 1102-1156.

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. Epub 2004 Jan 1013.

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S., and Robertson, D. (2000). Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Biochem Mol Biol **35,** 105-140.

Harper, G., Hull, R., Lockhart, B., and Olszewski, N. (2002). Viral sequences integrated into plant genomes. Annu Rev Phytopathol **40,** 119-136.

Hemmings-Mieszczak, M., Steger, G., and Hohn, T. (1997). Alternative structures of the cauliflower mosaic virus 35 S RNA leader: implications for viral expression and replication. J Mol Biol **267,** 1075-1088.

Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. Nature **447,** 418-424.

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. Epub 2006 May 2014.

Herr, A.J., and Baulcombe, D.C. (2004). RNA silencing pathways in plants. Cold Spring Harb Symp Quant Biol **69,** 363-370.

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. Epub 2005 Feb 2003.

Heslop-Harrison, J.S. (2000). Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. Plant Cell **12,** 617-636.

Heslop-Harrison, J.S., Brandes, A., and Schwarzacher, T. (2003). Tandemly repeated DNA sequences and centromeric chromosomal regions of Arabidopsis species. Chromosome Res **11,** 241-253.

Hill, J.E., Strandberg, J.O., Hiebert, E., and Lazarowitz, S.G. (1998). Asymmetric infectivity of pseudorecombinants of cabbage leaf curl virus and squash leaf curl virus: implications for bipartite geminivirus evolution and movement. Virology **250,** 283-292.

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.

Hirochika, H., Okamoto, H., and Kakutani, T. (2000). Silencing of retrotransposons in arabidopsis and reactivation by the ddm1 mutation. Plant Cell **12,** 357-369.

Ho, T., Wang, H., Pallett, D., and Dalmay, T. (2007). Evidence for targeting common siRNA hotspots and GC preference by plant Dicer-like proteins. FEBS Lett **581,** 3267-3272.

Hohn, T. (1999). Plant Pararetroviruses -- Caulimoviruses: molecular biology. In Encyclopedia of Virology, A. Granoff and R.G. Webster, eds (San Diego: Academic Press), pp. 1281-1285.

Hohn, T., Akbergenov, R., and Pooggin, M. (2007). Production and Transport of the Silencing Signal in Transgenic and Virus-Infected Plant Systems. In Plant Cell Monogr, E. Waigmann and M. Heinlein, eds (Berlin: Springer-Verlag), pp. 127-157.

Huang, J., Fan, T., Yan, Q., Zhu, H., Fox, S., Issaq, H.J., Best, L., Gangi, L., Munroe, D., and Muegge, K. (2004). Lsh, an epigenetic guardian of repetitive elements. Nucleic Acids Res **32,** 5019-5028.

Huettel, B., Kanno, T., Daxinger, L., Aufsatz, W., Matzke, A.J., and Matzke, M. (2006). Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. Embo J **25,** 2828-2836.

Hull, R. (2004). Matthews' plant virology. (Amsterdam: Elsevier).

Hull, R., Harper, G., and Lockhart, B. (2000). Viral sequences integrated into plant genomes. Trends Plant Sci **5,** 362-365.

Hunter, C., Sun, H., and Poethig, R.S. (2003). The Arabidopsis heterochronic gene ZIPPY is an ARGONAUTE family member. Curr Biol. **13,** 1734-1739.

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.

Hutvagner, G., Mlynarova, L., and Nap, J.P. (2000). Detailed characterization of the posttranscriptional gene-silencing-related small RNA in a GUS gene-silenced tobacco. Rna. **6,** 1445-1454.

Ingelbrecht, I., Van Houdt, H., Van Montagu, M., and Depicker, A. (1994). Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. Proc Natl Acad Sci U S A **91,** 10502-10506.

Jablonka, E., and Lamb, M.J. (1989). The inheritance of acquired epigenetic variations. J Theor Biol **139,** 69-83.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416,** 556-560.

Jacobsen, S.E., and Meyerowitz, E.M. (1997). Hypermethylated SUPERMAN epigenetic alleles in arabidopsis. Science **277,** 1100-1103.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. Genes Dev **12,** 1714-1725.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat Genet **22,** 94-97.

Ji, L.H., and Ding, S.W. (2001). The suppressor of transgene RNA silencing encoded by Cucumber mosaic virus interferes with salicylic acid-mediated virus resistance. Mol Plant Microbe Interact **14,** 715-724.

Jiang, J., Birchler, J.A., Parrott, W.A., and Dawe, R.K. (2003). A molecular view of plant centromeres. Trends Plant Sci **8,** 570-575.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol **17,** 379-384.

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.

Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. (2006). MicroRNAs and Their Regulatory Roles in Plants. Annu Rev Plant Biol **57**.

Jones, J.D., and Dangl, J.L. (2006). The plant immune system. Nature **444,** 323-329.

Jones, L., Ratcliff, F., and Baulcombe, D.C. (2001). RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Curr Biol **11,** 747-757.

Jones, L., Hamilton, A.J., Voinnet, O., Thomas, C.L., Maule, A.J., and Baulcombe, D.C. (1999). RNA-DNA interactions and DNA methylation in post-transcriptional gene silencing. Plant Cell **11,** 2291-2301.

Kakutani, T., Jeddeloh, J.A., and Richards, E.J. (1995). Characterization of an Arabidopsis thaliana DNA hypomethylation mutant. Nucleic Acids Res **23,** 130-137.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci U S A. **93,** 12406-12411.

Kang, B.C., Yeam, I., and Jahn, M.M. (2005). Genetics of plant virus resistance. Annu Rev Phytopathol **43,** 581-621.

Kankel, M.W., Ramsey, D.E., Stokes, T.L., Flowers, S.K., Haag, J.R., Jeddeloh, J.A., Riddle, N.C., Verbsky, M.L., and Richards, E.J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. Genetics **163,** 1109-1122.

Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., and Matzke, A.J. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr Biol **14,** 801-805.

Kanno, T., Aufsatz, W., Jaligot, E., Mette, M.F., Matzke, M., and Matzke, A.J. (2005a). A SNF2-like protein facilitates dynamic control of DNA methylation. EMBO Rep **6,** 649-655.

Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J. (2005b). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet. **37,** 761-765. Epub 2005 May 2029.

Kasschau, K.D., and Carrington, J.C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. Cell **95,** 461-470.

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.

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 U S A **103,** 18002-18007.

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.

Kidner, C.A., and Martienssen, R.A. (2005). The role of ARGONAUTE1 (AGO1) in meristem formation and identity. Dev Biol **280,** 504-517.

Kilby, N.J., Leyser, H.M., and Furner, I.J. (1992). Promoter methylation and progressive transgene inactivation in Arabidopsis. Plant Mol Biol **20,** 103-112.

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W.J., Koornneef, M., and Kakutani, T. (2007). Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. Plant J **49,** 38-45.

Kiparisov, S., Petrov, A., Meskauskas, A., Sergiev, P.V., Dontsova, O.A., and Dinman, J.D. (2005). Structural and functional analysis of 5S rRNA in Saccharomyces cerevisiae. Mol Genet Genomics **274,** 235-247.

Kjemtrup, S., Sampson, K.S., Peele, C.G., Nguyen, L.V., Conkling, M.A., Thompson, W.F., and Robertson, D. (1998). Gene silencing from plant DNA carried by a Geminivirus. Plant J **14,** 91-100.

Klahre, U., Crete, P., Leuenberger, S.A., Iglesias, V.A., and Meins, F., Jr. (2002). High molecular weight RNAs and small interfering RNAs induce systemic

posttranscriptional gene silencing in plants. Proc Natl Acad Sci U S A **99,** 11981-11986.

Kobayashi, K., Tsuge, S., Stavolone, L., and Hohn, T. (2002). The cauliflower mosaic virus virion-associated protein is dispensable for viral replication in single cells. J Virol **76,** 9457-9464.

Korf, I., Yandell, M., and Bedell, J.A. (2003). BLAST. (Sebastopol: O'Reilly).

Kovarik, A., Koukalova, B., Lim, K.Y., Matyasek, R., Lichtenstein, C.P., Leitch, A.R., and Bezdek, M. (2000). Comparative analysis of DNA methylation in tobacco heterochromatic sequences. Chromosome Res **8,** 527-541.

Kumar, A., and Bennetzen, J.L. (1999). Plant retrotransposons. Annu Rev Genet **33,** 479-532.

Kunz, C., Schob, H., Leubner-Metzger, G., Glazov, E., and Meins, F., Jr. (2001). Beta-1, 3-glucanase and chitinase transgenes in hybrids show distinctive and independent patterns of posttranscriptional gene silencing. Planta **212,** 243-249.

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.

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.

Kurihara, Y., Inaba, N., Kutsuna, N., Takeda, A., Tagami, Y., and Watanabe, Y. (2007). Binding of tobamovirus replication protein with small RNA duplexes. J Gen Virol **88,** 2347-2352.

Kutter, C., Schob, H., Stadler, M., Meins, F., Jr., and Si-Ammour, A. (2007). MicroRNA-Mediated Regulation of Stomatal Development in Arabidopsis. Plant Cell.

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.

Lam, E., Kato, N., and Lawton, M. (2001). Programmed cell death, mitochondria and the plant hypersensitive response. Nature **411,** 848-853.

Lawrie, C.H. (2007). MicroRNAs and haematology: small molecules, big function. Br J Haematol **137,** 503-512.

Lecellier, C.H., and Voinnet, O. (2004). RNA silencing: no mercy for viruses? Immunol Rev **198,** 285-303.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell **75,** 843-854.

Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna **12,** 589-597.

Lee, S.R., and Collins, K. (2007). Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. Nat Struct Mol Biol **14,** 604-610.

Legg, J.P., and Fauquet, C.M. (2004). Cassava mosaic geminiviruses in Africa. Plant Mol Biol **56,** 585-599.

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.

Lindbo, J.A., and Dougherty, W.G. (1992). Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. Virology **189,** 725-733.

Lindbo, J.A., Silva-Rosales, L., Proebsting, W.M., and Dougherty, W.G. (1993). Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. Plant Cell **5,** 1749-1759.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science **292,** 2077-2080.

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,** E67. Epub 2003 Dec 2022.

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.

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.

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.

Lu, R., Martin-Hernandez, A.M., Peart, J.R., Malcuit, I., and Baulcombe, D.C. (2003). Virus-induced gene silencing in plants. Methods **30,** 296-303.

Luo, J., and Hall, B.D. (2007). A multistep process gave rise to RNA polymerase IV of land plants. J Mol Evol **64,** 101-112.

Ma, J., Wing, R.A., Bennetzen, J.L., and Jackson, S.A. (2007). Plant centromere organization: a dynamic structure with conserved functions. Trends Genet **23,** 134-139.

Madlung, A., and Comai, L. (2004). The effect of stress on genome regulation and structure. Ann Bot (Lond) **94,** 481-495.

Makeyev, E.V., and Bamford, D.H. (2002). Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. Mol Cell. **10,** 1417-1427.

Malagnac, F., Bartee, L., and Bender, J. (2002). An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. Embo J **21,** 6842-6852.

Mallory, A.C., and Vaucheret, H. (2006). Functions of microRNAs and related small RNAs in plants. Nat Genet **38 Suppl,** S31-36.

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.

Martinez-Zapater, J.M., Estelle, M., and Somerville, C. (1986). A high repeated DNA sequence in Arabidopsis thaliana. Mol. Gen. Genet. **204,** 417–423.

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.

Mathieu, O., Yukawa, Y., Sugiura, M., Picard, G., and Tourmente, S. (2002). 5S rRNA genes expression is not inhibited by DNA methylation in Arabidopsis. Plant J **29,** 313-323.

Mathieu, O., Jasencakova, Z., Vaillant, I., Gendrel, A.V., Colot, V., Schubert, I., and Tourmente, S. (2003). Changes in 5S rDNA chromatin organization and transcription during heterochromatin establishment in Arabidopsis. Plant Cell **15,** 2929-2939.

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.

Matzke, M., Matzke, A.J., and Kooter, J.M. (2001). RNA: guiding gene silencing. Science **293,** 1080-1083.

Matzke, M., Kanno, T., Huettel, B., Daxinger, L., and Matzke, A.J. (2007). Targets of RNA-directed DNA methylation. Curr Opin Plant Biol.

Matzke, M., Kanno, T., Huettel, B., Jaligot, E., Mette, M.F., Kreil, D.P., Daxinger, L., Rovina, P., Aufsatz, W., and Matzke, A.J. (2005). RNA-directed DNA Methylation. In Plant Epigenetics, P. Meyer, ed (Oxford: Blackwell), pp. 69-96.

Matzke, M.A., Primig, M., Trnovsky, J., and Matzke, A.J. (1989). Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. Embo J **8,** 643-649.

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.

Meins, F., Jr. (1996). Epigenetic modifications and gene silencing in plants. In Epigenetic mechanisms of gene regulation, V.M. Russo, R.; Riggs, A., ed (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 415-442.

Meins, F., Jr., Si-Ammour, A., and Blevins, T. (2005). RNA silencing systems and their relevance to plant development. Annu Rev Cell Dev Biol **21,** 297-318.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature **431,** 343-349.

Merai, Z., Kerenyi, Z., Kertesz, S., Magna, M., Lakatos, L., and Silhavy, D. (2006). Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. J Virol **80,** 5747-5756.

Mette, M.F., van der Winden, J., Matzke, M.A., and Matzke, A.J. (1999). Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. Embo J **18,** 241-248.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. Embo J **19,** 5194-5201.

Mette, M.F., Kanno, T., Aufsatz, W., Jakowitsch, J., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2002). Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. Embo J **21,** 461-469.

Mette, M.F., Aufsatz, W., Kanno, T., Daxinger, L., Rovina, P., Matzke, M., and Matzke, A.J. (2005). Analysis of double-stranded RNA and small RNAs involved in RNA-mediated transcriptional gene silencing. Methods Mol Biol **309,** 61-82.

Mirkin, S.M. (2007). Expandable DNA repeats and human disease. Nature **447,** 932-940.

Moazed, D., Buhler, M., Buker, S.M., Colmenares, S.U., Gerace, E.L., Gerber, S.A., Hong, E.J., Motamedi, M.R., Verdel, A., Villen, J., and Gygi, S.P. (2006). Studies on the mechanism of RNAi-dependent heterochromatin assembly. Cold Spring Harb Symp Quant Biol **71,** 461-471.

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.

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.

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.

Moriones, E., and Navas-Castillo, J. (2000). Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. Virus Res **71,** 123-134.

Morris, C.A., and Moazed, D. (2007). Centromere assembly and propagation. Cell **128,** 647-650.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell. **119,** 789-802.

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.

Muangsan, N., Beclin, C., Vaucheret, H., and Robertson, D. (2004). Geminivirus VIGS of endogenous genes requires SGS2/SDE1 and SGS3 and defines a new branch in the genetic pathway for silencing in plants. Plant J **38,** 1004-1014.

Nogueira, F.T., Madi, S., Chitwood, D.H., Juarez, M.T., and Timmermans, M.C. (2007). Two small regulatory RNAs establish opposing fates of a developmental axis. Genes Dev **21,** 750-755.

O'Donnell K, A., and Boeke, J.D. (2007). Mighty Piwis Defend the Germline against Genome Intruders. Cell **129,** 37-44.

Oakeley, E.J. (1999). DNA methylation analysis: a review of current methodologies. Pharmacol Ther **84,** 389-400.

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.

Palauqui, J.C., Elmayan, T., Pollien, J.M., and Vaucheret, H. (1997). Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. Embo J **16,** 4738-4745.

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.

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.

Park, Y.D., Papp, I., Moscone, E.A., Iglesias, V.A., Vaucheret, H., Matzke, A.J., and Matzke, M.A. (1996). Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. Plant J **9,** 183-194.

Parker, J.S., and Barford, D. (2006). Argonaute: A scaffold for the function of short regulatory RNAs. Trends Biochem Sci **31,** 622-630.

Pavet, V., Quintero, C., Cecchini, N.M., Rosa, A.L., and Alvarez, M.E. (2006). Arabidopsis displays centromeric DNA hypomethylation and cytological alterations of heterochromatin upon attack by pseudomonas syringae. Mol Plant Microbe Interact **19,** 577-587.

Pelissier, T., and Wassenegger, M. (2000). A DNA target of 30 bp is sufficient for RNA-directed DNA methylation. Rna. **6,** 55-65.

Pelissier, T., Thalmeir, S., Kempe, D., Sanger, H.L., and Wassenegger, M. (1999). Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucleic Acids Res. **27,** 1625-1634.

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.

Peters, A.H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. Curr Opin Cell Biol **17,** 230-238.

Pfeffer, S., Zavolan, M., Grasser, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., and Tuschl, T. (2004). Identification of virus-encoded microRNAs. Science **304,** 734-736.

Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grasser, F.A., van Dyk, L.F., Ho, C.K., Shuman, S., Chien, M., Russo, J.J., Ju, J., Randall, G., Lindenbach, B.D., Rice, C.M., Simon, V., Ho, D.D., Zavolan, M., and Tuschl, T. (2005). Identification of microRNAs of the herpesvirus family. Nat Methods **2,** 269-276.

Pfeiffer, P., and Hohn, T. (1983). Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell **33,** 781-789.

Pikaard, C.S. (2000). The epigenetics of nucleolar dominance. Trends Genet **16,** 495-500.

Pikaard, C.S. (2006). Cell Biology of the Arabidopsis Nuclear siRNA Pathway for RNA-directed Chromatin Modification. Cold Spring Harb Symp Quant Biol **71,** 473-480.

Pinto, Y.M., Kok, R.A., and Baulcombe, D.C. (1999). Resistance to rice yellow mottle virus (RYMV) in cultivated African rice varieties containing RYMV transgenes. Nat Biotechnol **17,** 702-707.

Poethig, R.S., Peragine, A., Yoshikawa, M., Hunter, C., Willmann, M., and Wu, G. (2006). The Function of RNAi in Plant Development. Cold Spring Harb Symp Quant Biol **71,** 165-170.

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.

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.

Pooggin, M., and Hohn, T. (2004). Fighting geminiviruses by RNAi and vice versa. Plant Mol Biol **55,** 149-152.

Pooggin, M., Shivaprasad, P.V., Veluthambi, K., and Hohn, T. (2003). RNAi targeting of DNA virus in plants. Nat Biotechnol **21,** 131-132.

Pooggin, M.M., Hohn, T., and Futterer, J. (1998). Forced evolution reveals the importance of short open reading frame A and secondary structure in the cauliflower mosaic virus 35S RNA leader. J Virol **72,** 4157-4169.

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.

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.

Ratcliff, F.G., Harrison, B.D., and Baulcombe, D.B. (1997). A Similarity Between Viral Defense and Gene Silencing in Plants. Science **276,** 1558-1560.

Ratcliff, F.G., MacFarlane, S.A., and Baulcombe, D.C. (1999). Gene silencing without DNA. rna-mediated cross-protection between viruses. Plant Cell **11,** 1207-1216.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev. **16,** 1616-1626.

Ro, S., Park, C., Song, R., Nguyen, D., Jin, J., Sanders, K.M., McCarrey, J.R., and Yan, W. (2007). Cloning and expression profiling of testis-expressed piRNA-like RNAs. Rna.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol **53,** 247-259.

Rothnie, H.M., Chapdelaine, Y., and Hohn, T. (1994). Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv Virus Res **44,** 1-67.

Ruiz, M.T., Voinnet, O., and Baulcombe, D.C. (1998). Initiation and maintenance of virus-induced gene silencing. Plant Cell **10,** 937-946.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning. A Laboratory Manual. (Woodbury: CSHL Press).

Saze, H., and Kakutani, T. (2007). Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1. Embo J.

Saze, H., Mittelsten Scheid, O., and Paszkowski, J. (2003). Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet **34,** 65-69.

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.

Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., and Sanger, H.L. (1993). RNA-directed RNA polymerase from tomato leaves. II. Catalytic in vitro properties. J Biol Chem **268,** 11858-11867.

Schiebel, W., Pelissier, T., Riedel, L., Thalmeir, S., Schiebel, R., Kempe, D., Lottspeich, F., Sanger, H.L., and Wassenegger, M. (1998). Isolation of an RNA-

directed RNA polymerase-specific cDNA clone from tomato. Plant Cell **10,** 2087-2101.

Schöb, H., Kunz, C., and Meins, F., Jr. (1997). Silencing of transgenes introduced into leaves by agroinfiltration: a simple, rapid method for investigating sequence requirements for gene silencing. Mol Gen Genet **256,** 581-585.

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.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S.A. (2002). A high-throughput Arabidopsis reverse genetics system. Plant Cell **14,** 2985-2994.

Shivaprasad, P.V., Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T., and Pooggin, M.M. (2005). Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. J Virol **79,** 8149-8163.

Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J.N., and Kooter, J.M. (2001). Transcriptional and posttranscriptional gene silencing are mechanistically related. Curr Biol **11,** 436-440.

Silhavy, D., and Burgyan, J. (2004). Effects and side-effects of viral RNA silencing suppressors on short RNAs. Trends Plant Sci **9,** 76-83.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol. **10,** 169-178.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard, C.S., and Baulcombe, D.C. (2007). An SNF2 Protein Associated with Nuclear RNA Silencing and the Spread of a Silencing Signal between Cells in Arabidopsis. Plant Cell.

Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M. (2000). Total silencing by intron-spliced hairpin RNAs. Nature **407,** 319-320.

Solexa, I. (2006). Protocol for Whole Genome Sequencing using Solexa Technology. Biotechniques**,** 1-2.

Sontheimer, E.J. (2005). Assembly and function of RNA silencing complexes. Nat Rev Mol Cell Biol **6,** 127-138.

Soosaar, J.L., Burch-Smith, T.M., and Dinesh-Kumar, S.P. (2005). Mechanisms of plant resistance to viruses. Nat Rev Microbiol **3,** 789-798.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. Embo J **21,** 6549-6559.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol **98,** 503-517.

Stam, M., Viterbo, A., Mol, J.N., and Kooter, J.M. (1998). Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: implications for posttranscriptional silencing of homologous host genes in plants. Mol Cell Biol **18,** 6165-6177.

Steimer, A., Amedeo, P., Afsar, K., Fransz, P., Mittelsten Scheid, O., and Paszkowski, J. (2000). Endogenous targets of transcriptional gene silencing in Arabidopsis. Plant Cell **12,** 1165-1178.

Stern-Ginossar, N., Elefant, N., Zimmermann, A., Wolf, D.G., Saleh, N., Biton, M., Horwitz, E., Prokocimer, Z., Prichard, M., Hahn, G., Goldman-Wohl, D., Greenfield, C., Yagel, S., Hengel, H., Altuvia, Y., Margalit, H., and Mandelboim, O. (2007). Host immune system gene targeting by a viral miRNA. Science **317,** 376-381.

Stokes, T.L., and Richards, E.J. (2002). Induced instability of two Arabidopsis constitutive pathogen-response alleles. Proc Natl Acad Sci U S A **99,** 7792-7796.

Stokes, T.L., Kunkel, B.N., and Richards, E.J. (2002). Epigenetic variation in Arabidopsis disease resistance. Genes Dev **16,** 171-182.

Sudarshana, M.R., Roy, G., and Falk, B.W. (2007). Methods for engineering resistance to plant viruses. Methods Mol Biol **354,** 183-195.

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A **102,** 152-157.

Sullivan, C.S., Grundhoff, A.T., Tevethia, S., Pipas, J.M., and Ganem, D. (2005). SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature **435,** 682-686.

Sunkar, R., Girke, T., and Zhu, J.K. (2005). Identification and characterization of endogenous small interfering RNAs from rice. Nucleic Acids Res **33,** 4443-4454.

Sunkar, R., Chinnusamy, V., Zhu, J., and Zhu, J.K. (2007). Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci **12,** 301-309.

Susi, P., Hohkuri, M., Wahlroos, T., and Kilby, N.J. (2004). Characteristics of RNA silencing in plants: similarities and differences across kingdoms. Plant Mol Biol **54,** 157-174.

Svoboda, P., and Di Cara, A. (2006). Hairpin RNA: a secondary structure of primary importance. Cell Mol Life Sci **63,** 901-908.

Szymanski, M., Barciszewska, M.Z., Erdmann, V.A., and Barciszewski, J. (2003). 5 S rRNA: structure and interactions. Biochem J **371,** 641-651.

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.

Takeda, S., and Paszkowski, J. (2006). DNA methylation and epigenetic inheritance during plant gametogenesis. Chromosoma **115,** 27-35.

Takeda, S., Sugimoto, K., Kakutani, T., and Hirochika, H. (2001). Linear DNA intermediates of the Tto1 retrotransposon in Gag particles accumulated in stressed tobacco and Arabidopsis thaliana. Plant J **28,** 307-317.

Talmor-Neiman, M., Stav, R., Klipcan, L., Buxdorf, K., Baulcombe, D.C., and Arazi, T. (2006). Identification of trans-acting siRNAs in moss and an RNA-dependent RNA polymerase required for their biogenesis. Plant J **48,** 511-521.

Tang, G., and Zamore, P.D. (2004). Biochemical dissection of RNA silencing in plants. Methods Mol Biol. **257,** 223-244.

Tariq, M., and Paszkowski, J. (2004). DNA and histone methylation in plants. Trends Genet **20,** 244-251.

Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y., and Paszkowski, J. (2003). Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. Proc Natl Acad Sci U S A **100,** 8823-8827.

Tijsterman, M., Ketting, R.F., and Plasterk, R.H. (2002). The genetics of RNA silencing. Annu Rev Genet **36,** 489-519.

Topp, C.N., Zhong, C.X., and Dawe, R.K. (2004). Centromere-encoded RNAs are integral components of the maize kinetochore. Proc Natl Acad Sci U S A **101,** 15986-15991.

Trinks, D., Rajeswaran, R., Shivaprasad, P.V., Akbergenov, R., Oakeley, E.J., Veluthambi, K., Hohn, T., and Pooggin, M.M. (2005). Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J Virol **79,** 2517-2527.

Turnage, M.A., Muangsan, N., Peele, C.G., and Robertson, D. (2002). Geminivirus-based vectors for gene silencing in Arabidopsis. Plant J **30,** 107-114.

Vaillant, I., Schubert, I., Tourmente, S., and Mathieu, O. (2006). MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in Arabidopsis. EMBO Rep **7,** 1273-1278.

Vaillant, I., Tutois, S., Cuvillier, C., Schubert, I., and Tourmente, S. (2007a). Regulation of Arabidopsis thaliana 5S rRNA Genes. Plant Cell Physiol **48,** 745-752.

Vaillant, I., Tutois, S., Cuvillier, C., Schubert, I., and Tourmente, S. (2007b). Arabidopsis thaliana 5S ribosomal RNA genes regulation. Plant Cell Physiol.

Valle, R.P., Skrzeczkowski, J., Morch, M.D., Joshi, R.L., Gargouri, R., Drugeon, G., Boyer, J.C., Chapeville, F., and Haenni, A.L. (1988). Plant viruses and new perspectives in cross-protection. Biochimie **70,** 695-703.

van Blokland, R., van der Geest, N., Mol, J.N., and Kooter, J.M. (1994). Transgene-mediated suppression of chalcone synthase expression in *Petunia hybrida* results from an increase in RNA turnover. The Plant Journal **6,** 861-877.

Vance, V., and Vaucheret, H. (2001). RNA silencing in plants--defense and counterdefense. Science **292,** 2277-2280.

Vanderschuren, H., Akbergenov, R., Pooggin, M.M., Hohn, T., Gruissem, W., and Zhang, P. (2007). Transgenic cassava resistance to African cassava mosaic virus is enhanced by viral DNA-A bidirectional promoter-derived siRNAs. Plant Mol Biol **64,** 549-557.

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.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev **20,** 759-771.

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. Epub 2004 May 1186.

Vazquez, F. (2006). Arabidopsis endogenous small RNAs: highways and byways. Trends Plant Sci **11,** 460-468.

Vazquez, F., Gasciolli, V., Crete, P., and Vaucheret, H. (2004a). The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. Curr Biol. **14,** 346-351.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P. (2004b). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Mol Cell. **16,** 69-79.

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. Epub 2004 Jan 2002.

Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. Trends Genet **17,** 449-459.

Voinnet, O. (2003). RNA silencing bridging the gaps in wheat extracts. Trends Plant Sci **8,** 307-309.

Voinnet, O. (2005). Induction and suppression of RNA silencing: insights from viral infections. Nat Rev Genet **6,** 206-220.

Voinnet, O., and Baulcombe, D.C. (1997). Systemic signalling in gene silencing. Nature **389,** 553.

Voinnet, O., Pinto, Y.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 U S A **96,** 14147-14152.

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D.C. (1998). Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. Cell **95,** 177-187.

Volpe, T., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C. (2003). RNA interference is required for normal centromere function in fission yeast. Chromosome Res **11,** 137-146.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993). Arabidopsis thaliana DNA methylation mutants. Science **260,** 1926-1928.

Wada, Y., Miyamoto, K., Kusano, T., and Sano, H. (2004). Association between up-regulation of stress-responsive genes and hypomethylation of genomic DNA in tobacco plants. Mol Genet Genomics **271,** 658-666.

Wang, X.H., Aliyari, R., Li, W.X., Li, H.W., Kim, K., Carthew, R., Atkinson, P., and Ding, S.W. (2006). RNA interference directs innate immunity against viruses in adult Drosophila. Science **312,** 452-454.

Wassenegger, M., and Krczal, G. (2006). Nomenclature and functions of RNA-directed RNA polymerases. Trends Plant Sci. **11,** 142-151.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell **76,** 567-576.

Watson, J.M., Fusaro, A.F., Wang, M., and Waterhouse, P.M. (2005). RNA silencing platforms in plants. FEBS Lett **579,** 5982-5987.

Weber, M., and Schubeler, D. (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Curr Opin Cell Biol **19,** 273-280.

Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G., and Waterhouse, P.M. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J. **27,** 581-590.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. Cell **78,** 1101-1115.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell **75,** 855-862.

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.

Wolffe, A.P. (1998). Chromatin: Structure & function. (San Diego: Academic Press).

Xie, M., He, Y., and Gan, S. (2001a). Bidirectionalization of polar promoters in plants. Nat Biotechnol **19,** 677-679.

Xie, Z., Fan, B., Chen, C., and Chen, Z. (2001b). An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. Proc Natl Acad Sci U S A. **98,** 6516-6521. Epub 2001 May 6515.

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. Epub 12005 Aug 12929.

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. Epub 2004 Feb 2024.

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.

Yi, H., Riddle, N.C., Stokes, T.L., Woo, H.R., and Richards, E.J. (2004). Induced and natural epigenetic variation. Cold Spring Harb Symp Quant Biol **69,** 155-159.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet **13,** 335-340.

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. Epub 2005 Aug 2130.

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.

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.

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.

Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. Cell **128,** 763-776.

Zhang, P., Vanderschuren, H., Futterer, J., and Gruissem, W. (2005a). Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. Plant Biotechnol J **3,** 385-397.

Zhang, W., Yi, C., Bao, W., Liu, B., Cui, J., Yu, H., Cao, X., Gu, M., Liu, M., and Cheng, Z. (2005b). The transcribed 165-bp CentO satellite is the major functional centromeric element in the wild rice species Oryza punctata. Plant Physiol **139,** 306-315.

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.

Zhang, X., Yuan, Y.R., Pei, Y., Lin, S.S., Tuschl, T., Patel, D.J., and Chua, N.H. (2006a). Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. Genes Dev **20,** 3255-3268.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J.R. (2006b). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell **126,** 1189-1201.

Zheng, X., Zhu, J., Kapoor, A., and Zhu, J.K. (2007). Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. Embo J **26,** 1691-1701.

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. Epub 2003 Jan 2009.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet **39,** 61-69.

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.