GEMINIVIRUSES AS MODELS TO STUDY THE
ESTABLISHMENT AND MAINTENANCE
OF DNA METHYLATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Jamie N. Jackel, B.S.
Graduate Program in Molecular Genetics

The Ohio State University
2013

Dissertation Committee:
Dr. David M. Bisaro, Adviser
Dr. Deborah S. Parris, Co-adviser
Dr. Biao Ding
Dr. David Mackey
Dr. R. Keith Slotkin
Copyright by
Jamie N. Jackel
2013
ABSTRACT

RNA silencing refers to a set of mechanistically related and evolutionarily conserved processes including post-transcriptional gene silencing (PTGS) and transcriptional gene silencing (TGS). In plants, TGS is often accompanied by RNA-directed DNA methylation (RdDM). Both PTGS and TGS use specialized double stranded RNA binding proteins (DRBs) in conjunction with specific ribonucleases (Dicer-like proteins; DCL) to cleave 21, 22, or 24 nucleotide (nt) small interfering RNAs (siRNAs) from large double stranded RNAs (dsRNA). Cytoplasmic PTGS utilizes 21 and 22 nt siRNAs to target mRNA for degradation. The current model for TGS/RdDM involves a complex nuclear pathway where evolutionarily related forms of RNA polymerase II (Pol II), known as Pol IV and Pol V, transcribe DNA and work together with RNA dependent RNA polymerase 2 (RDR2) to produce long dsRNAs that are cleaved into 24 nt siRNAs by Dicer-like 3 (DCL3). These siRNAs associate with a complex that contains the ribonuclease argonaute 4 (AGO4) to recruit histone and DNA methyltransferases that subsequently target homologous DNA for methylation-mediated silencing. A feature that enhances PTGS and TGS defense is the ability of the silencing signal to spread from cell-to-cell and systemically throughout the plant. An antiviral role is well established, and the importance of PTGS as a host defense is clear from the fact that virtually all plant viruses encode proteins that target and suppress different aspects of this pathway. By contrast, while TGS has long been known to suppress potentially damaging DNA such as
transposons, it has only recently been shown to target geminivirus DNA for repressive methylation.

Geminiviruses are circular single-stranded DNA (ssDNA) viruses that replicate in the nucleus through a double-stranded DNA (dsDNA) intermediate that associates with histones to form minichromosomes. Geminiviruses depend on host cellular machinery for both replication and transcription, thus providing excellent models to study the epigenetic regulation of these processes. Unlike most plant viruses (which have RNA genomes) geminiviruses must combat both PTGS and TGS defense pathways in order to be successful. We have shown that two related geminivirus-encoded silencing suppressors, AL2 and L2, suppress both PTGS and TGS by interacting with and inhibiting adenosine kinase (ADK), a methyl cycle co-factor. However, recent work has identified an additional suppression function of AL2 that is independent of ADK inhibition and requires AL2's ability to activate transcription of host genes. Chapter 2 of this thesis characterizes the transcription activation-dependent suppressor function of AL2 in TGS reversal and the inhibition of systemic spread of silencing. Using ADK knock-down, it was determined that methyl cycle inhibition has no impact on systemic spread and that AL2, but not L2 or AL2 lacking its transcriptional activation domain (AL2 1-114), can prevent the spread of the silencing signal. Thus inhibition of systemic silencing likely depends on the ability of AL2 to activate transcription. Subsequent studies analyzing TGS reversal in vegetative and reproductive tissues of *Nicotiana benthamiana* plants revealed that only AL2 could reverse TGS in reproductive tissue in an ADK and transcription-independent manner. Thus, a third, previously unknown suppressor function of AL2 that is independent of transcription activation and ADK
inhibition was identified. These studies show that unraveling the mechanisms that geminivirus suppressors use to inhibit various silencing pathways will lead to a better understanding of these complex host defenses.

Since TGS targets geminivirus dsDNA for repressive methylation and aspects of the RdDM pathway remain elusive, the studies described in this thesis also use geminiviruses as sensitive models to identify new pathway components and to unveil interesting roles of well-established pathway proteins. In Arabidopsis thaliana, it is well known that DRB proteins interact with specific DCL proteins to produce the canonical 21, 22, and 24 nt siRNAs for PTGS/TGS. Arabidopsis contains five DRB (DRB1-DRB5) proteins and four DCLs (DCL1-DCL4), each of which functionally partner in particular RNAi pathways. Previous work has shown that DRB1 and DCL1 functionally interact to cleave long hairpin dsRNA in the microRNA pathway, while DRB4 and DCL4 often target long dsRNA derived from replication intermediates of RNA viruses. Chapter 3 of this thesis uses the geminivirus model system to identify the DRB protein that functions with AGO4 and DCL3 in the RdDM pathway. It is shown that like dcl3 deficient plants, drb3 mutant plants are hypersusceptible to geminiviruses. Moreover, neither dcl3 or drb3 mutants are able to hypermethylate the viral genome, indicating that drb3 plants are deficient in the methylation-mediated defense pathway. DRB3 is also shown to physically interact with both AGO4 and DCL3 in the nucleus. Furthermore, analysis of geminivirus derived siRNAs found that drb3 mutants are able to produce 24 nt siRNAs similar to wild-type plants, indicating that DRB3 is not required for siRNA biogenesis. Altogether the data demonstrate that DRB3 associates with AGO4 and DCL3 in the RdDM pathway that is
crucial for defending against viral DNA. Furthermore, this work emphasizes the sensitivity of the geminivirus model system to study RdDM pathway components.

Although it is believed that Pol IV and Pol V are required to initiate methylation, a key unanswered question is how these polymerases are targeted to specific DNA sequences. Due to their ssDNA nature and amplification by rolling-circle replication, nascent geminivirus genomes evade methylation and thus provide a unique transient model system to study de novo RdDM. Chapter 4 of this thesis shows that Pol IV and V, as well as their associated chromatin remodelers, play critical roles in the methylation-mediated defense against geminiviruses. However, promoter regions of geminivirus DNA continue to be methylated in the absence of both Pol IV and V, suggesting that another polymerase may initiate DNA methylation. Subsequent experiments found that Pol IV and V initially associate with the promoter region of the viral genome and during later infection spread to a coding region. Alternatively, Pol II constitutively associated with both the promoter and coding region. DNA methylation analysis of this coding region found hypermethylation only during late stages of infection, concurrent with the appearance of Pol IV and V. This supports previous work by others indicating that Pol IV and V are important for spread of RdDM. Finally, analysis of virus-derived siRNAs in pol IV, pol V, and pol IV/V infected mutant plants showed that Pol IV and V are not required for the biogenesis of viral siRNAs. Taken together, this data indicates that Pol IV and V play important roles in amplifying, maintaining, and spreading DNA methylation, but are not required for the initiation of RdDM. This suggests that another polymerase, possibly Pol II, initially transcribes DNA leading to the production of siRNAs and ultimately resulting in the initiation of DNA methylation.
Altogether the work in this thesis furthers our understanding of geminivirus pathogenesis by unveiling new suppressor functions of AL2. Moreover, it provides novel information on the host defenses by placing DRB3 in the RdDM pathway, and re-defining the roles of Pol IV and Pol V. Together these studies illustrate the utility of the geminivirus model system for analyzing epigenetic regulation and host defense pathways.
I dedicate this to my family for all of their support and encouragement in everything I pursue.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. David Bisaro, for his encouragement and guidance throughout my graduate career. Our many long scientific discussions provided a continuous stream of experiments, while pushing the boundaries of my knowledge and imagination. I will always appreciate his enthusiasm, which inspired and motivated me throughout my graduate career. Finally, I would like to thank him for his unwavering support of my scientific fervor and unique ideas.

I truly appreciate the time, guidance, advice and support from my committee members Dr. Biao Ding and Dr. David Mackey. I would like to convey a special thanks to my committee member Dr. Keith Slotkin for the outstanding advice, knowledge and discussions that strengthened my research through our joint lab meetings. I would like to especially thank my co-advisor, Dr. Deborah S. Parris, for her extensive assistance and support for both my research and funding.

I would like to thank both my past and present lab colleagues Dr. Cody Buchmann, Dr. Gireesha Mohannath, Dr. Yan Xie, Sizhun Li, Jeffery Ostler, Udit Singhal, and Elisa Furay for their assistance, recommendations and informative conversations. I would also like to thank Tami Coursey and Jessica Storer for their continued moral support, scientific discussions and friendship. I would especially like to express my gratitude to both Dr. Priya Raja and Dr. Kenn Buckley for their technical advice, unwavering support, and friendship. All together, this group of colleagues made my tenure in the lab very
enjoyable as well as extremely productive. I would like to thank Dr. Ying Wang from Dr. Ding's lab for many enlightening and motivating discussions. From Dr. Grotewold's lab, I would like to express my gratitude to Bettina Whittler for support and great lunchtime exchanges, and Isa Casas for her insightful scientific and personal advice. I have to specially thank two great friends, Katja Machemer-Noonan and Jennifer Nauman, who provided me with some of my fondest memories of graduate school and long lasting friendships. Our lunches together are something I will always miss. Katja thank you not just for the stimulating scientific conversations, but also for all of the support and advice both scientifically as well as personally. Jenni, I cannot be appreciative enough for your ongoing encouragement, always listening and your advice that changed my outlook on life. Thank you for all of the great times inside and outside of work.

I could never fully express my gratitude for my amazing and loving husband, Matt. I would not be where I am today without your steadfast love and support in all of my endeavors. Thank you for always making me smile and forcing me to take study breaks. I cannot wait to see where this life takes us and I know you will always be by my side making me laugh. You are truly the love of my life and words could never describe my love for you.

I would also like to thank Dave, Theresa and Andy Jackel. Since the beginning, you have been there for me and supported me in everything I did. You have always made me feel like I was one of your own and I could be more excited to be part of your family. I am looking forward to all of the great times we have ahead.

I would like to convey my sincerest appreciation to my parents, Pat and Patty Wolf, without you I would not be the person I am today. You have given me the confidence
and drive to reach all of my goals. You have given me unconditional love and support throughout my life. I am so grateful for everything you have done for me and love you more than you will ever know. My genuine gratitude to all my siblings, Justin, Jason and Jaclyn, growing up with you has provided me with the best memories and I would not be the same without you. Justin, I will always look up to you and appreciate all your advice and support throughout my life. Jason, your love for life and easy going personality are something I hope to emulate everyday. Jaclyn, there is nothing in this world like a sister, you are my best friend and my life would not be the same without you.
VITA

May 2007……………………………B.S. Molecular Biology/Biotechnology,
Clarion University of Pennsylvania
June 2007-present…………………..Graduate Teaching Associate and Research Fellow
The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Molecular Genetics
TABLE OF CONTENTS

Abstract ............................................................................................................................ ii
Dedication ........................................................................................................................ vii
Acknowledgements ........................................................................................................ viii
Vita ..................................................................................................................................... xi
List of Figures .................................................................................................................. xviii

CHAPTERS:

CHAPTER 1: INTRODUCTION ......................................................................................... 1

1.1 GEMINIVIRUSES ........................................................................................................ 1

1.1.1 Geminivirus virion structure and genome organization ...................................... 2
1.1.2 Geminivirus satellite DNAs ................................................................................. 6
1.1.3 Geminivirus proteins ......................................................................................... 6
1.1.4 Geminivirus infection cycle: transmission through replication ......................... 10
1.1.5 Geminivirus Transcription ......................................................................... 13

1.2 RNA SILENCING AS A DEFENSE MECHANISM ............................................... 15

1.2.1 RNA silencing pathways in plants ..................................................................... 16
1.2.2 Post-transcriptional gene silencing (PTGS) ....................................................... 18
1.2.3 Transcriptional gene silencing (TGS) and the methylation pathway .............. 21
1.2.4 Systemic spread of the silencing signals ......................................................... 24
1.2.5 TGS as a defense against geminiviruses ......................................................... 27
1.2.6 Suppression of PTGS by geminivirus proteins ............................................. 30
1.2.7 Suppression of TGS by geminivirus proteins ............................................... 35
1.2.8 Novel suppressor activities of the geminivirus AL2 protein ............................ 39

1.3 CHROMATIN STRUCTURE AND EPIGENETICS ............................................. 40

1.3.1 Nucleosome structure and assembly .............................................................. 40
1.3.2 Histones modifications and variants .............................................................. 42
1.3.3 Chromatin remodelers .................................................................................... 46
1.3.4 Pol IV and Pol V mediate chromatin formation and spread DNA methylation in plants .......................................................... 48
1.3.5 RdDM and the potential role of RNA Polymerase II ....................................... 53
1.3.6 Geminivirus chromatin .................................................................................. 54
1.3.7 Geminivirus chromatin as a model to study the RdDM pathway ................... 55

CHAPTER 2 ............................................................................................................. 65

ANALYSIS OF GEMINIVIRUS AL2 PROTEIN IDENTIFIES NEW SUPPRESSOR ACTIVITIES AND DEVELOPMENTAL CHANGES IN THE ROLE OF ADK IN THE N. BENHAMIANA METHYL CYCLE. ................................................. 65

2.1 INTRODUCTION .................................................................................................. 65

2.2 MATERIALS AND METHODS .......................................................................... 68

2.2.1 Systemic spread suppression assay ............................................................... 68
2.2.2 Vegetative and Reproductive plants ............................................................... 68
2.2.3 Developmental PTGS reversal assay ............................................................. 69
2.2.4 Developmental TGS reversal assay ............................................................... 69
2.2.5 RNA extraction and Northern blot analysis ................................................ 70
2.2.6 DNA isolation and Southern blot analysis ................................................... 70

xiii
2.2.7 Oligonucleotides
2.2.8 ADK activity assay
2.2.9 Methylation sensitive extension assay

2.3 RESULTS

2.3.1 AL2, but not L2, prevents spread of systemic silencing in a transcription-dependent manner

2.3.2 Plant developmental stage can determine whether transcription-dependent or – independent mechanisms can reverse PTGS.

2.3.3 Plant developmental stage can determine whether geminiviruses can reverse TGS.

2.3.4 A novel mechanism of AL2-mediated TGS reversal is employed in reproductive plant

2.3.5 TGS is independent of ADK activity in reproductive plants

2.3.6 ADK inhibition inhibits global cytosine methylation in vegetative, but not reproductive plants

2.4 DISCUSSION

CHAPTER 3

ARABIDOPSIS DOUBLE-STRANDED RNA BINDING PROTEIN DRB3 PARTNERS WITH DCL3 AND AGO4 IN METHYLATION-MEDIATED DEFENSE AGAINST GEMINIVIRUSES

3.1 INTRODUCTION

3.2 MATERIALS AND METHODS

3.2.1 Arabidopsis mutants

3.2.2 Virus inoculation

3.2.3 BiFC analysis

3.2.4 Co-immunoprecipitation and immunoblot analysis
3.2.5 Bisulfite sequencing ................................................................................................. 102
3.2.6 Small RNA analysis .................................................................................................. 102
3.3 RESULTS ..................................................................................................................... 103
  3.3.1 Arabidopsis drb3 mutants are hypersusceptible to geminivirus infection .......... 103
  3.3.2 Arabidopsis dcl3 and drb3 mutants are unable to recover from geminivirus infection ................................................................. 105
  3.3.3 Arabidopsis dcl3 and drb3 mutants cannot hypermethylate the viral genome ............................................................................. 107
  3.3.4 Both active and repressed viral genomes are present in infected plants .......... 109
  3.3.5 DRB3 interacts with DCL3 and AGO4 ................................................................. 110
  3.3.6 DRB3 is not required for the biogenesis of 24 nt siRNAs by DCL3.................... 112
3.4 DISCUSSION .............................................................................................................. 113
CHAPTER 4 ..................................................................................................................... 125
4.1 INTRODUCTION ......................................................................................................... 125
4.2 MATERIALS AND METHODS .................................................................................... 128
  4.2.1 Arabidopsis Mutants ............................................................................................. 128
  4.2.2 Virus Inoculation .................................................................................................. 129
  4.2.3 Bisulfite sequencing .............................................................................................. 130
  4.2.4 Chromatin immunoprecipitation (ChIP) ............................................................ 131
  4.2.5 Small RNA analysis ............................................................................................. 132
4.3 RESULTS ..................................................................................................................... 132
  4.3.1 Methylation pathway components Pol IV, Pol V CLSY!, and DRD1 are involved in defense against geminiviruses ................... 132
  4.3.2 Pol IV and V amplify viral DNA methylation, but are not required for its establishment .............................................................. 134
4.3.3 Pol II, IV and V differentially associate with viral DNA .................... 137

4.3.4 Pol IV and V association with a coding region correlates with methylation spread................................................................. 139

4.3.5 Pol IV and Pol V are not required for biogenesis of virus-derived siRNAs .... 140

4.3.6 Pol IV and V are required for the establishment and spread of repressive histone methylation........................................................................................................ 142

4.4 DISCUSSION .................................................................................................................. 144

CHAPTER 5 ......................................................................................................................... 158

DISCUSSION..................................................................................................................... 158

5.1 Analysis of AL2 suppressor activities and developmental changes in the role of ADK in the methyl cycle .............................................................. 158

5.1.1 AL2 transcription-dependent suppressor functions ..................... 158

5.1.2 Discovery of a transcription- and ADK-independent AL2 RNA silencing suppressor activity ............................................................. 160

5.1.3 Changes in the role of ADK in the methyl cycle of reproductive plants .... 161

5.2 GEMINIVIRUSES AS MODELS TO EXAMINE HOST RDDM PATHWAY COMPONENTS ........................................................................................................ 162

5.2.1 Identification of DRB3 as an RdDM pathway component using the geminivirus model system ........................................................................ 162

5.2.2 Geminiviruses as models to study Pol IV and Pol V in RdDM .......... 164

5.2.3 Pol IV and Pol V are not required for establishment of cytosine methylation, but are needed for its spread ......................................................... 166

5.2.4 Pol IV and Pol V are required for the establishment of dimethylated H3K9 repressive chromatin marks .............................................................. 168

5.2.5 Identifying a role for Pol II in siRNA biogenesis, DNA methylation and RdDM .................................................................................... 170

5.2.6 A model for geminivirus de novo methylation ........................................ 171

xvi
5.3 GEMINIVIRUSES AS MODELS TO STUDY CHROMATIN TARGETING AND READING ................................................................................................................................. 173

5.3.1 Determining how Pol II, Pol IV, and Pol V recognize and target DNA for methylation ........................................................................................................ 173

5.3.2 Geminiviruses as models to identify chromatin readers ........................................ 175

BIBLIOGRAPHY .................................................................................................................. 178

APPENDIX A .......................................................................................................................... 205

SUMMARY OF PUBLICATION CONTRIBUTIONS ................................................................. 205

6.1 GEMINIVIRUS AL2 AND L2 PROTEINS SUPPRESS TRANSCRIPTIONAL GENE SILENCING AND CAUSE GENOME-WIDE REDUCTIONS IN CYTOSINE METHYLATION .................................................................................................................. 205

6.2 SUPPRESSION OF METHYLATION-MEDIATED TRANSCRIPTIONAL GENE SILENCING BY βC1-SAHH PROTEIN INTERACTION DRIVES GEMINIVIRUS-DNA β SATELLITE ASSOCIATION ................................................................. 207

6.3 CHARACTERIZATION OF THE RNA SILENCING SUPPRESSION ACTIVITY OF THE EBOLA VIRUS VP35 PROTEIN IN PLANTS AND MAMMALIAN CELLS .................................................................................................................. 209

6.4 A COMPLEX CONTAINING SNR1-RELATED KINASE (SNRK1) AND ADENOSINE KINASE IN ARABIDOPSIS ................................................................................................................................. 211

APPENDIX B .......................................................................................................................... 213

SUPPLEMENTAL DATA FIGURES .......................................................................................... 213
LIST OF FIGURES

Fig. 1.1 Geminivirus genomes and transcripts................................................................. 57
Fig. 1.2 Geminivirus replicative cycle .............................................................................. 58
Fig.1.3 Schematic diagram of the geminivirus intergenic region .................................... 59
Fig.1.4 Antiviral RNA silencing pathways in Arabidopsis ........................................... 60
Fig. 1.5 The methyl cycle and its inhibition by geminivirus proteins.......................... 61
Fig. 1.6 Nucleosome structure and histone tail modifications ........................................ 62
Fig. 1.7 Subunit composition of RNA Polymerases II, IV and V ............................... 63
Fig. 1.8 RNA directed DNA methylation of geminivirus genomes.............................. 64
Fig. 2.1 AL2, but not L2, blocks the systemic silencing signal ....................................... 88
Fig. 2.2 AL2 and L2 reverse established PTGS in a developmentally-dependent manner ................................................................. 89
Fig. 2.3 Plant developmental age determines whether geminiviruses can reverse TGS ........................................................................................................................................ 90
Fig. 2.4 AL2 developmental-independent TGS reversal is transcriptionally-independent ................................................................. 91
Fig. 2.5 TGS is independent of ADK activity in reproductive plants ......................... 92
Fig. 2.6 ADK knock-down cannot reduce global cytosine methylation in reproductive plants ........................................................................................................................................ 93
Fig. 2.7 Oligonucleotides used in this study ................................................................. 94
Fig. 3.1 Arabidopsis drb3 mutants show enhanced susceptibility to geminivirus infection ........................................................................................................................................ 119
Fig. 3.2 *Arabidopsis* *drb3* and *dcl3* mutants do not recover from infection with BCTV *L2* mutant virus. ................................................................. 120

Fig. 3.3 *Arabidopsis* *drb3* and *dcl3* mutants fail to hypermethylate the viral genome. Methylation of BCTV *L2* IR DNA was assessed by bisulfite sequencing......................... 121

Fig. 3.4 Hypermethylated viral genomes are associated with H3K9me2 .................. 122

Fig. 3.5 DRB3 interacts with DCL3 and AGO4 in distinct subnuclear bodies ........... 123

Fig. 3.6 DRB3 co-immunoprecipitates with DCL3 and AGO4................................. 124

Fig. 3.7 DRB3 is not required for biogenesis of geminivirus-derived siRNAs.......... 124

Fig. 4.1 Plants deficient for Pol IV, Pol V, Pol IV/V, CLSY1 and DRD1 are hypersusceptible to geminivirus infection.................................................... 150

Fig. 4.2 Pol IV and Pol V are not required for viral genome methylation.......... 151

Fig. 4.3 Pol II, Pol IV and Pol V differentially associate with the viral genome .... 153

Fig. 4.4 Spread of DNA methylation during recovery correlates with the association of Pol IV and V ................................................................. 154

Fig. 4.5 Pol IV and V are not required to generate BCTV-specific siRNAs......... 155

Fig. 4.6 Pol IV and V are required for the establishment and spread of dimethylated H3K9 histone marks................................................................. 156

Fig. 5.1 A new model for AL2 silencing suppression activities........................ 176

Fig. B.1 *Arabidopsis* *drb3* and *dcl3* mutants do not recover from infection with *BCTV L2*. ................................................................................. 213

Fig. B.2 Cytosine methylation analysis of the BCTV *L2* genome in wild-type and mutant plants ................................................................................. 214

Fig. B.3 Cytosine methylation analysis of the BCTV *L2* genome in wild-type, *drb3*, and *dcl3* mutant plants ........................................................................ 216

Fig. B.4 Hypermethylated viral genomes are associated with H3K9me2............ 217

Fig. B.5 Interactions between DRB3, DCL3 and AGO4 proteins using BiFC analysis in *N. benthamiana* epidermal cells ....................................................... 218
Fig. B.6 Sequences of oligonucleotides used in Chapter 3 ................................................. 220

Fig. B.7 Oligonucleotides used in Chapter 4 ........................................................................ 220

Fig. B.8 Statistical analysis of bisulfite clones isolated from recovered and non-recovered tissues ........................................... 222

Fig. B.9 NRPB2-, NRPD1-, and NRPE1-FLAG transgenic plants recover from geminivirus infection ........................................................................................................ 222
CHAPTER 1

INTRODUCTION


1.1 Geminiviruses

The earliest record of a plant virus disease describes the beauty of geminivirus symptoms in a poem dated back to 752 a.d. (Saunders et al. 2003). The spectacular mosaics produced by geminiviruses ultimately led to dissemination of the Abutilon mosaic virus in the late eighteen hundreds from the West Indies to Europe (Wege et al. 2000). It was not until the twentieth century that the causative agent was identified by electronmicrograph, where paired particles were observed in the phloem cells of infected plants. It was the twinned nature of the capsid that coined the name Geminiviridae.

Geminiviruses are characterized by the paired icosahedral capsid, which encapsidates one molecule of circular single-stranded DNA (ssDNA). The Geminiviridae is divided into four genera, which were named by abbreviation of the type members, Mastrevirus (from Maize streak virus), Topocuvirus (type member Tomato pseudo curly top virus), Curtovirus (represented by Beet curly top virus), and Begomovirus (Bean golden mosaic
viruses are classified into each genus by host range, genome organization, and insect vector (Fauquet et al. 2003). Geminiviruses infect a wide range of angiosperms and cause severe crop losses around the world. Infection of staple crops such as maize, cassava, sweet potato, potato, tomato, bean, and pepper and wheat has sparked widespread interest in geminivirus research (Brown et al. 1996, Moffatt 1999).

Initially, basic geminivirus biology was studied to obtain information on viral gene function, replication, and transcription. These fundamental studies ultimately led to the use of geminiviruses as models to study plant molecular biology. The small genome size and dependence on host cell machinery make geminiviruses excellent models to study the regulation of host replication and transcription, protein expression, gene silencing, and more recently epigenetics (Raja et al. 2008). This dissertation describes how studies of geminivirus pathogenicity factors can unveil interesting plant cell biology. Moreover, it explores the use of the geminivirus system for studying how plant cells recognize DNA.

1.1.1 Geminivirus virion structure and genome organization

Although geminiviruses are disseminated around the world and cause severe crop losses, it took a long time to identify them as causative agents of disease. Virus titers can reach hundreds of thousands per cell, but whole plant titers are fairly low, because most geminiviruses are mostly phloem limited. As a result, geminiviruses are difficult to purify. However, a combination of homogenization and several rounds of centrifugation eventually led to the purification of conspicuous paired virions (Larson and Duffus 1983). Despite the twinned capsid structure, virions were found to contain only one molecule of circular ssDNA. Thus, geminiviruses that have two genome components
(bipartite begomoviruses described later in this section) require two twinned particles carrying different genome to obtain a systemic infection. Moreover, since empty geminivirus shells have not been isolated, it is believed that the ssDNA and capsid protein interaction stabilizes virion structure, and that packaging occurs concomitantly with replication. Recently, a high-resolution virion structures of two geminiviruses were obtained using electron cryomicroscopy combined with image reconstruction and bioinformatics modeling. Both models agreed that the twinned particle is formed by two incomplete icosahedra composed of 11 pentameric capsomers missing one capsomer each. The two halves are joined at a 20° angle with an opening in between (Zhang et al. 2001, Bottcher et al. 2004). The single molecule of DNA forms a bow-tie configuration in the middle of the two half particles.

Although all four genera of geminiviruses have the same capsid structure, genome organization, host range, and insect vector can vary considerably. Mastreviruses are spread by leafhoppers to mostly monocotyledonous plants. This genus includes Maize streak virus and Wheat dwarf virus. There is only one known topocuvirus, Tomato pseudo curly top virus, which is spread by treehoppers. Curtoviruses and begomoviruses infect a variety of angiosperms and are transmitted by leafhoppers and a specific species of whitefly, Bemisia tabaci, respectively. The mastre- and topocu-viruses will not be discussed further because the members of the begomovirus and curtovirus genera are the main focus of this dissertation. Tomato golden mosaic virus (TGMV), African cassava mosaic virus (ACMV), Mung bean yellow mosaic virus (MYMV), and Cabbage leaf curl virus (CaLCuV) are a few well-studied bipartite begomoviruses. These bipartite begomoviruses can be sub-classified by sequence into Old World (e.g. ACMV, MYMV)
and New World (e.g. CaLCuV, TGMV) viruses. In addition, monopartite begomoviruses originated in the Old World. This group includes the well known Tomato yellow leaf curl virus (TYLCV) and Tomato leaf curl virus (TLCV). Begomoviruses have relatively narrow host ranges. However, all of these viruses infect the model plant Nicotiana bethamiana, but only CaLCuV causes symptomatic infections in the genetic model plant Arabidopsis thaliana. By contrast, curtoviruses have a very broad host range and infect both model organisms mentioned above. Examples include the type member Beet curly top virus (BCTV) and Spinach curly top virus (SCTV).

Generally, geminivirus genomes range from 2.5-3.0 kb and encode six to seven genes. Due to the limited coding compacity of the small viral genome, geminiviruses depend on host cellular machinery for both replication and transcription. During replication, the ssDNA genome (viral or plus strand) is used as template for complementary (minus) strand synthesis, forming a double stranded DNA (dsDNA) intermediate. As isolated viral DNA is infectious, this process must be accomplished by host enzymes. The dsDNA replication form (RF) is the substrate for transcription and subsequent rolling circle replication. Begomoviruses have one or two chromosomes, while only a single genome is required for curtovirus infection. Bipartite begomoviruses require both similarly sized genomes for infectivity, which are encapsidated separately. Genome A contains genes required for replication and encapsidation, while genome B encodes movement proteins important for systemic spread throughout the vasculature of the plant (Rogers et al. 1986, Sunter et al. 1987). Monopartite begomoviruses are similar to curtoviruses in genome organization. Genes transcribed from the leftward promoters are important for replication, while rightward transcription products are required for
movement and encapsidation (Fig. 1.1). There are two systems for naming geminivirus genes, both of which designate genes and proteins by numbers. One system names genes based on whether they are encoded by the viral sense (V; viral or plus strand) or in the complementary sense (C) DNA strand. The other method simply labels genes based on their leftward (L) or rightward (R) orientation on the circular genome. In both classification systems, bipartite begomovirus genes are labeled first with either an A or B for the encoding chromosome. For this thesis, the R and L system will be used. Examples of monopartite and bipartite begomovirus (Tomato golden mosaic virus, or TGMV) and a curtovirus (Beet curly top virus, or BCTV) and the gene nomenclature are shown in Figure 1.1.

Geminivirus replicative forms are organized similar to polyomaviruses (e.g. SV40). Both viruses are very economical in genome organization with viral genes encoded on both DNA strands and from overlapping open reading frames. Transcription of viral genes occurs from divergent DNA dependent RNA polymerase II (Pol II) promoters, which are located in ~250 bp intergenic region (IR). In bipartite begomoviruses, a portion of the IR is known as the common region (CR), as it is conserved between the two genome components, although it differs between viruses. Between these two oppositely orientated promoters in the IR lies the origin of replication that includes a conserved stem loop structure that must be recognized to initiate rolling circle amplification of the plus strand. The stem structure is formed by an inverted repeat sequence, and the loop contains the invariant TAATATTAC sequence (Lazarowitz 1992). The IR is considered to be the master control region of the virus and thus is a target for plant defense mechanisms.
1.1.2 Geminivirus satellite DNAs

Some Old World monopartite begomoviruses associate with small circular satellite DNAs, DNA 1 and the more common DNA β (Mansoor et al. 1999, Saunders and Stanley 1999, Briddon et al. 2003, Briddon and Stanley 2006). Only DNA β will be discussed further in this thesis. DNA β depends on the helper begomovirus for replication and encapsidation, which is critical for spread throughout the plant and insect transmission (Murrant and Mayo 1982). Although this satellite DNA has virtually no sequence similarity to the helper virus DNA, it contains the conserved hairpin and invariant nonanucleotide loop sequence required for geminivirus replication. DNA β has been found with a variety of begomoviruses and in some combinations is required for symptom development and increased infectivity (Saunders et al. 2000, Briddon et al. 2001). The DNA β satellite encodes a single gene product called βC1, which has been identified as an important pathogenicity factor and will be discussed further in section 1.2.8.

1.1.3 Geminivirus proteins

Unless otherwise noted, the bipartite begomoviruses will be used as a model to describe the functions of typical geminivirus proteins. Leftward transcription on the A chromosome produces proteins that are involved in replication and transcription. These proteins are multifunctional, and some (but most certainly not all) of their functions have been identified. The AL1 protein (L1, also known as replication initiator protein or Rep)
is the only viral protein required for replication (Elmer et al. 1988). This remarkable protein has properties reminiscent of both SV40 T-antigen and the replication initiator protein of single-stranded bacteriophages such as ΦX174. Like the phage proteins, AL1 carries out the essential cleavage and ligation events that initiate and terminate rolling circle replication (Laufs et al. 1995a). Additionally, it mediates origin recognition by binding within the origin of replication, where it recruits host replication factors including proliferating cell nuclear antigen (PCNA), replication protein A (RPA), and replication factor C (Fontes et al. 1994, Luque et al. 2002, Castillo et al. 2003, Singh et al. 2006). It also likely serves as a replicative helicase (Choudhury et al. 2006, Clerot et al. 2006). These diverse activities appear to be regulated, at least in part, by the oligomeric state of the protein. In addition, AL1 provides an environment favorable to DNA synthesis by interacting the plant retinoblastoma-related protein (pRBR) to antagonize E2F-mediated repression of cellular replication genes such as PCNA (Nagar et al. 1995, Ach et al. 1997, Kong et al. 2000, Egelkrout et al. 2001), and further conditions cells through interactions with kinases that modulate metabolism (Shen et al. 2006, Shen et al. 2009). The AL3 protein (L3, also known as replication enhance protein; REn) is needed for optimal viral DNA amplification (Sunter et al. 1990). It is unclear how AL3 stimulates replication, however, it is known to interact with AL1 and may assist in recruiting host factors to the origin. It also may condition the host cell through independent interactions with pBRB, PCNA and other proteins (Settlage et al. 2001, Castillo et al. 2003, Settlage et al. 2005). The importance of AL1/AL3 cell conditioning was underscored by a global analysis of transcription, which confirmed that geminiviruses impact the pRBR/E2F network and promote endoreduplication in infected
cells (Ascencio-Ibánez et al. 2008). The AL2 protein (also known as transcriptional activator protein; TrAP) is a transcription factor required for the expression of viral late genes (Sunter and Bisaro 1991, Sunter and Bisaro 1992). It contains a nuclear localization signal, an acidic transcription activation domain, and a Cys-His zinc finger domain, which plays an important role in dimerization. AL2 dimerization is required for optimal transcription activation and mutations in the Cys-His (e.g. AL2-C33A) region result in a reduction in that activity (Yang et al. 2007). The mechanisms of AL2 transcription activation are not completely understood. AL2 binds dsDNA weakly in a sequence non-specific manner, thus it associates with a host transcription factor to activate CP transcription (Lacatus and Sunter 2009). This transcription regulation is cell type specific, as rightward promoters must be derepressed in phloem tissue and activated in mesophyll cells (Sunter and Bisaro, 1997). AL2 also acts as a pathogenicity factor by inhibiting the cellular stress response mediated by SNF1-related kinase (SnRK1) (Sunter et al. 2001, Hao et al. 2003). AL2 also suppresses RNA silencing by multiple mechanisms that appear to involve the activation of cellular genes and the inhibition of adenosine kinase (ADK) (Bisaro 2006). (The role of AL2 in silencing suppression is discussed in detail in section 1.2.8.) The curtovirus protein L2 is a positional homolog of AL2. Although L2 does not activate transcription, it does shares some but not all pathogenicity functions (Wang et al. 2003, Wang et al. 2005, Buchmann et al. 2009). Another important pathogenicity factor, AL4, lies completely within the ORF of AL1, but in a different reading frame. AL4 is dispensable for TGMV infection, but in ACMV isolates it can act as a pathogenicity factor by suppressing RNA silencing (Vanitharani et al. 2005, Vanitharani et al. 2004). More recently, the ACMV AL4 was shown to have
post-translational modifications that are important for membrane targeting and pathogenesis (Fondong et al. 2007). It is difficult to draw solid conclusions on the role of AL4/L4 in pathogenicity and spread of the virus due to the variable sequences and functions this protein has in different geminivirus genera.

Proteins required for cell to cell movement and systemic spread differ between curtoviruses and begomoviruses. The role of AR1 (R1 in curtoviruses, also known as coat protein; CP) is similar in both genera. AR1 is necessary for encapsidation of progeny virions, insect transmission and thus plant-to-plant spread in both genera. It is also required for spread between cells and systemic spread throughout the plant in curtoviruses and monopartite begomoviruses. However, it is not required for cell to cell and systemic spread in bipartite begomoviruses (Gardiner et al. 1988). In bipartite begomoviruses, the B genome encodes only two genes, BR1, which is the nuclear shuttle protein (NSP), and BL1 (MP; movement protein). Both are required for cell-to-cell movement and spread of the virus in infected plants. BR1 is known to bind non-specifically to ssDNA and dsDNA, while BL1 associates with plasmodesmata. Moreover, BR1 and BL1 are known to interact with each other. Thus, it is hypothesized that BR1 delivers the ssDNA to BR1 at the plasmodesmata, resulting in the local spread of viral DNA (Rojas et al. 2005). Curtoviruses also encode R2 and R3 proteins. R3 mutations cause asymptomatic infections and severely reduced viral titers (Hormuzdi and Bisaro 1993). This was not due to a negative impact on replication nor to indirect effects on CP expression. Instead the effects were reminiscent of DNA A infections alone (Klinkenburg and Stanley 1990). Therefore R3 likely plays a critical role in viral movement and spread. On the other hand, R2 (also known as precoat protein; PCP) plays
an important role in the production of ssDNA, but the molecular mechanisms by which it functions are still unknown (Hormuzdi and Bisaro 1993). The R2 protein is not a homolog of the curtovirus proteins. Monopartite begomoviruses R2 acts as a pathogenicity factor by suppressing host defenses (Zrachya et al. 2007).

1.1.4 Geminivirus infection cycle: transmission through replication

Curtoviruses are transmitted by various leafhopper species, while begomoviruses are spread by a single species of whitefly, *Bemisia tabaci*. Transmission is circulative but nonpropagative, meaning that the virus does not replicate in the insect. Upon uptake the virions enter the insect gut and must pass through the midgut epithelium. They are then transported through the hemocoel and enter the salivary glands. The capsids protect the DNA through the insect gut, where they likely interact with specific midgut epithelial receptors (Rosell et al. 1999). This specific interaction allows passage into the hemocoel, where specific insect chaperones may stabilize the viral capsids. For efficient transmission, the viral capsid needs to attach specifically to receptors in the salivary gland epithelia (Morin et al. 1999, Akad et al. 2004). This process results in a latent period of 4 to 12 hrs between initial ingestion and detectable transmission. Once the geminivirus particles reach the salivary glands, the whitefly will transmit the virus throughout its lifespan. Upon feeding on a new plant, the virion particles are delivered directly to phloem cells, including sieve elements, companion cells, and phloem parenchyma cells.

After they are injected into phloem cells, viral capsids must be disassembled and the ssDNA delivered to the nucleus. Although the exact details of this process are unknown,
the three dimensional structure of the capsomers suggests that amino acids important for nuclear localization are clustered on the outer surface of the virion (Qin et al. 1998). The virion may dock onto the nuclear pore and inject the ssDNA or the entire virion may enter the nucleus for uncoating. Once in the nucleus, the viral DNA is amplified in three distinct stages (Hamilton et al. 1982, Stanley and Townsend 1985, Slomka et al. 1988, Saunders et al. 1991, Stenger et al. 1991, Saunders et al. 1992; viral replication is reviewed in Bisaro 1996, Gutierrez et al. 2004, Hanley-Bowdoin et al. 2004, Jeske 2007, Jeske 2009). Stage A converts the circular ssDNA into covalently closed circular dsDNA (ccc dsDNA), which is followed by amplification of these dsDNA intermediates by rolling-circle replication (RCR). Finally, Stage C produces mature genomic circular ssDNA that is packaged into viral particles (Fig. 1.2).

First, the circular ssDNA genome is converted into the ccc dsDNA called the replication form (RF). Minus strand synthesis might occur in dividing phloem cells, or it might be accomplished by repair synthesis in non-dividing cells. In mastreviruses, a small DNA oligonucleotide is encapsidated along with the viral genome (Donsen et al. 1984). This primer is sufficient for initiation of RF in vitro (Hayes et al. 1987). But curtoviruses and begomoviruses do not encapsidate primers, thus the exact molecular mechanisms for this stage are not fully understood. In begomoviruses, it is believed that RNA polymerase produces a primer, which can be extended by DNA polymerase resulting in the RF (Saunders et al. 1992). The ccc dsDNA is then twisted by topoisomerases and associated with nucleosomes to form supercoiled minichromosomes (more information provided in section 1.3.4). This minichromosome is the template for further replication and transcription by host machinery.
Since geminiviruses rely on host machinery and their replication occurs mostly in differentiated cells, they must induce host DNA synthesis in order to continue through the replication cycle. The second stage of replication (RF → RF) relies on RCR, which depends on viral gene expression. The critical viral proteins required for replication are AL1 and AL3, which in turn induce the expression of host replication machinery (described in more detail in 1.1.3). The nicking-closing activity initiates and terminates RCR of the plus strand. Beyond containing the conserved geminivirus stem-loop structure, the IR also contains all cis-acting sequence elements required for replication (diagramed in Fig. 1.3). The organization of these sequence motifs is highly conserved, but exhibits genera specific characteristics (Argüello-Astorga et al. 1994). In all geminiviruses, replication is initiated by a dimer of AL1, which binds specifically to iterative sequences, termed the Rep binding site. This binding site can lie between 50-100 bp upstream of the stem-loop (Fontes et al. 1992, Choi and Stenger 1995). In order to initiate RCR, AL1 must nick the viral plus strand within the conserved nonanucleotide loop sequence (TAATATT/AC, where the / indicates the nick site). Since the AL1 binding site relatively distant from the nick site, the molecular mechanisms for initiation of RCR is still unknown. One model suggests that AL3 recognizes and binds both the stem-loop and to the DNA bound AL1. This interaction could allow AL1 to nick the stem-loop (Hanley-Bowdoin et al. 1996). Alternatively, a separate AL1 protein could recognize the nick site. Once AL1 nicks the dsDNA, it stays bound to the 5' phosphate, conserving the binding energy of the phosphodiester bond and protecting the 5' end (Heyraud-Nitschke et al. 1995, Laufs et al. 1995a, Laufs et al. 1995b, Stanley et al. 1995). DNA dependent DNA polymerase, likely Pol δ or Pol ε, uses the free 3' OH end
of the DNA as a primer. Upon completing a round of RCR, another AL1 can nick the nascent ssDNA and the 5' phosphate bound AL1 can ligate the 5' and 3' ends (Heyraud-Nitschke et al. 1995, Laufs et al. 1995b). This newly synthesized ssDNA can be converted back to dsDNA, amplifying the RF form. Eventually, a switch occurs and RCR continues in much the same way except the priming event for complementary strand synthesis is repressed and only ssDNA molecules are produced (RF → SS). How this switch is regulated is unknown, but the amount of R2 and CP proteins may play an important role in this switch.

Interestingly, two-dimensional southern blot analysis identified heterogeneous linear dsDNA intermediates (Jovel et al. 2007). These results were inconsistent with RCR being the only means of geminivirus replication. It was proposed that recombination-dependent replication (RDR) can be used to repair ssDNAs resulting from early replication termination and partial digestion by host enzymes (Jeske et al. 2001). These partial genomes would contain a free 3' end of a ssDNA or a 3' overhang of dsDNA can invade homologous dsDNA. The free 3' end would be used as a primer leading to a repaired full-length genome. Thus, RDR would be evolutionarily beneficial and might explain the recombinogenic potential of geminiviruses.

1.1.5 Geminivirus transcription

Geminivirus genome organization superficially resembles that of the polyomaviruses; with an intergenic region (IR) containing oppositely oriented promoters separated by the origin of replication. The sequences that comprise these functional elements overlap extensively within this small region (Fig. 1.3). Transcription programs have been best
studied in the bipartite begomoviruses (for review see Hanley-Bowdoin et al. 2000, Jeske 2009). Viral mRNAs are polyadenylated and initiate upstream of consensus TATA box and initiator sequences typical of RNA polymerase II promoters. Most are unspliced, although transcription patterns are otherwise complex. Overlapping transcripts and transcripts containing multiple open reading frames (ORFs) are common. For example, in the TGMV A component, there are three leftward (complementary sense), 3’ co-terminal transcripts which overlap the 3’ end of the convergent rightward, viral sense transcription unit that encodes CP (Fig. 1.1). The major leftward promoter in the IR, AL-62 (the number refers to the nucleotide coordinate of the transcription start site), sponsors transcription of an mRNA that spans the AL1, AL2, and AL3 genes. This transcript also includes AL4, which is embedded within AL1 but in a different reading frame. The AL1 and AL3 proteins, and most likely AL4, appear to be expressed from this transcript (Hanley-Bowdoin et al. 1989). Two downstream transcripts driven by individual promoters, AL-1629 and AL-1935, have also been identified in TGMV infected cells (Sunter and Bisaro 1989). The shorter AL-1629 supports expression of the AL2 protein, whereas AL3 is translated from both AL-1629 and AL-1935 (Shung et al. 2006). Thus, within individual mRNAs, mechanisms that allow translation of some downstream genes (e.g. leaky scanning) and inhibit the translation of others (e.g. cryptic ORFs) likely operate (Shung and Sunter 2009). Additional study will be needed before these mechanisms are fully elucidated. However, recent evidence suggests that the activities of the downstream promoters are linked to AL-62 promoter activity. As noted earlier, AL1 acts as an origin recognition protein by binding a specific sequence. The AL1 binding sites lie within the replication origin, but immediately upstream of the AL-62 initiation
site. As a result, AL1 is auto-regulatory and represses its own transcription (Sunter et al. 1993, Eagle et al. 1994). Interestingly, activity from the AL-1629 and AL-1935 promoters require prior AL1-dependent repression of AL-62 (Shung and Sunter 2007). This suggests that bipartite begomovirus transcription occurs in three phases: gene products involved in replication (AL1 and AL3) are expressed early, AL2 expression is delayed-early, and expression of late genes (CP and NSP) from rightward promoters, which require transactivation by AL2, is a late event (Shung and Sunter 2007).

1.2 RNA silencing as a defense mechanism

Plants have evolved several defense mechanisms to fend off the variety of pathogens encountered on a daily basis. Some of these pathways are conserved from plants to animals, including innate and adaptive responses. The innate responses are more general and include (but are not limited to) responses mediated by cellular receptors that recognize microbial or viral signatures, for example flagellin or double-stranded RNA. Activation of these receptors initiates a signaling cascade that ultimately leads to metabolic and gene expression changes immediately upon infection (reviewed in Ronald and Beutler 2010). In contrast, adaptive responses occur later in infection are pathogen-specific defense mechanisms. One adaptive defense that is conserved in fungi, plants, and mammals is RNA silencing, a term that encompasses several pathways that are commonly triggered by dsRNA. RNA silencing maintains genome stability by silencing transposons, regulates gene expression with microRNAs, mediates DNA and histone methylation, and is potentially the most effective defense against viruses through production of virus derived small interfering RNAs (siRNAs) (Carrington and Ambros
1.2.1 RNA silencing pathways in plants

RNA silencing is mediated by 18-25 nucleotide (nt) RNAs of two main types; small interfering RNAs (siRNA) and micro RNAs (miRNA). MicroRNAs are encoded by the host genome to regulate gene expression. Alternatively, siRNAs that mediate the PTGS defense pathways are not encoded, instead siRNAs originate from foreign dsRNA (such as replicative forms of RNA viruses) and overlapping or aberrant transcripts. Both are processed (diced) from larger dsRNA precursors by Dicer-Like (DCL) proteins with the assistance of DRB proteins. DCL proteins contain RNase III domains, which are responsible for the catalytic activity of the proteins. DCLs also contain double stranded RNA binding motifs (dsRBMs), which bind dsRNA in a sequence non-specific manner and mediate protein-protein interactions (Saunders and Barber 2003, Chang and Ramos 2005, Hiraguri et al. 2005). In *Arabidopsis*, DCL1, DCL2, DCL3 and DCL4 generate RNAs that are mostly 21 nt, 22 nt, 24 nt, and 21 nt in length, respectively. DCLs partner with specific dsRNA binding proteins (DRB in *Arabidopsis*) in order to bind and process dsRNA and also to load the siRNAs into RISC complexes (Tabara et al. 2002, Liu et al. 2003, Baulcombe 2004, Lippman and Martienssen 2004, Voinnet 2005). These conserved pathways depend on several key players, including Dicer proteins, double stranded RNA binding proteins (DRB), RNA-dependent RNA polymerases (RDR), and Argonaute proteins (AGO). Plants have evolved highly elaborate silencing pathways that require the concerted action of four Dicer-like proteins (DCL), five DRBs, six RDRs, and ten AGOs, all of which functionally partner in distinct ways to mediate different aspects of RNA silencing (reviewed in Broderson and Voinnet 2006, Vaucheret 2006).
2003, Hiraguri et al. 2005, Jiang et al. 2005). Although there are five DRB proteins in
Arabidopsis, only two have been associated with a specific DCL. DRB1 (a.k.a. HYL1) partners with DCL1 to efficiently produce miRNAs and selects the guide strand that is loaded into RISC (Han et al. 2004, Vasquez et al. 2004, Hiraguri et al. 2005, Curtin et al. 2008, Eamens et al. 2009). DRB2 also appears to work with DCL1 in the accumulation of shoot apical meristem miRNAs (Eamens et al. 2012). DCL4 functionally and physically interacts with DRB4 in the trans-acting siRNA pathway (Adenot et al. 2006, Nakazawa et al. 2007, Curtin et al. 2008). DRB4 is not required for 21 nt siRNA biogenesis but these siRNAs are reduced in drb4 plants. While the roles of DRB1 and DRB4 have been clearly established, the DRB proteins that associate with DCL2 and DCL3 are not involved in 22 and 24 nt siRNA production and thus have not been determined (Curtin et al. 2008, Eamens et al. 2012). Expression analysis of DRB3 and DRB5 found that both are cytoplasmic proteincins expressed in the shoot apical meristem, but their functions remain elusive (Eamens et al. 2012). The small RNA duplexes that are produced by DCL and DRB proteins have two nucleotide 3' overhangs, and the terminal nucleotide is 2'-O-methylated by Hua Enhancer 1 (HEN1). HEN1-mediated methylation of siRNAs and miRNAs stabilizes their ends by blocking oligouridylation and subsequent degradation (Li et al. 2005, Yu et al. 2005). The guide strand of a siRNA or miRNA is incorporated into an RNA induced silencing complex (RISC) and targets this effector to complementary RNA. RISC complexes may function in the nucleus or the cytoplasm, and always include an AGO protein. Possible outcomes of small RNA-directed RISC activity include cleavage (slicing) of target RNAs and/or translational repression, depending on the degree of homology between the small RNA and target
RNA. Generally perfect complementarity leads to target cleavage whereas imperfect base pairing to mRNA causes translational repression by unknown mechanisms. In addition, specialized nuclear RISC directs methylation of cytosine residues and/or histone methylation, leading to the assembly of heterochromatin.

Most endogenous siRNAs produced in Arabidopsis are directed against transposons and inverted repeat sequences, and thus are part of the heterochromatic pathways that will be discussed in 1.2.3 (Kasschau et al. 2007). However, some loci produce small RNA that direct PTGS (e.g., trans-acting siRNAs and natural antisense RNAs) and it is pathways similar to these that have been adapted for defense against viral RNA in plants. There is growing evidence that the miRNA and related silencing pathways have antiviral roles in mammalian cells, and that mammalian viruses produce proteins that are able to suppress these responses (Lecellier et al. 2005, Li and Ding 2006, Triboulet et al. 2007, de Vries and Berkhout 2008, Qian et al. 2009, Wu et al. 2009, Zhu et al. 2012).

1.2.2 Post-transcriptional gene silencing (PTGS)

RNA silencing was first recognized as an antiviral defense in plants (Lindbo et al. 1993, Ratcliff et al. 1997, Ruiz et al. 1998). Its importance is supported by the observation that virus infection is invariably accompanied by the appearance of siRNAs derived from the viral genome, and even more so by the fact that all successful plant viruses examined to date encode at least one silencing suppressor. The vast majority of plant viruses have genomes composed of RNA and are therefore targeted by the cytoplasmic, antiviral PTGS pathway (reviews in Voinnet 2005, Ding and Voinnet 2007, Mlotshwa et al. 2008, Ruiz-Ferrer and Voinnet 2009). A diagram illustrating
cytoplasmic and nuclear antiviral silencing pathways, and sites of action of selected viral suppressor proteins is presented in Fig 1.4.

Analysis of siRNAs in plants infected with RNA viruses has revealed that substrates for DCL activities include inverted repeat structures in viral transcripts as well as double-stranded replicative intermediates (Molnar et al. 2005, Ho et al. 2006). Thus, the dsRNA that serves as the source of primary viral siRNAs is produced by viral RDRs. A defining feature of PTGS is the appearance of 21-22 nt siRNAs, and although functional redundancy obscured the issue from some time, recent evidence indicates that DCL4 is in many cases the primary antiviral Dicer (Blevins et al. 2006, Bouche et al. 2006, Deleris et al. 2006, Diaz-Pendon et al. 2007, Qu et al. 2008). The 21 nt DCL4-related siRNAs are typically more abundant than the 22 nt siRNAs generated by DCL2. However, DCL2 redundantly participates in antiviral silencing when DCL4 is inactivated by mutation or by a viral silencing suppressor. Significantly, enhanced host susceptibility with dramatic reductions in virus-derived siRNA are usually observed only in dcl2/dcl4 double mutants (Deleris et al. 2006, Diaz-Pendon et al. 2007, Qu et al. 2008). DRB4, which associates with DCL4, also contributes to antiviral silencing (Adenot et al. 2006, Nakazawa et al. 2007, Qu et al. 2008). DCL1, which generates miRNAs, and DCL3, which produces siRNAs associated with methylation, do not appear to be directly involved in defense against RNA viruses but may have regulatory roles (Diaz-Pendon et al. 2007, Qu et al. 2008).

Amplification of viral siRNA reinforces the antiviral silencing response in plants. Source material for secondary virus-specific siRNAs is generated primarily by the action of RDR6, and redundantly by RDR1, which can synthesize dsRNA using primed or
unprimed ssRNA templates (Diaz-Pendon et al. 2007, Qu et al. 2008, Voinnet 2008). These cellular RDRs are activated by "aberrant" RNAs, including those that are not capped or lack polyadenylation, a characteristic of many viral RNAs. Sliced viral transcripts may also share these features. Synthesis of dsRNA additionally involves multiple host factors, including Suppressor of Gene Silencing 3 (SGS3) (Mourrain et al. 2000). The resulting dsRNA is subsequently processed by DCL4 (and redundantly by DCL2) to generate secondary siRNAs (Moissard et al. 2007, Voinnet 2008). Both primary and secondary siRNAs are ultimately loaded into RISC complexes, and several lines of evidence indicate that AGO1 is the major antiviral slicer (Baumberger and Baulcombe 2005, Zhang et al. 2006, Pantoleo et al. 2007, Qu et al. 2008). However, AGO7 can redundantly serve, particularly in the absence of AGO1 function, although AGO1- and AGO7- containing RISC complexes prefer slightly different substrates (Qu et al. 2008).

Transcripts produced by viruses that package DNA genomes, including the geminiviruses and pararetroviruses such as Cauliflower mosaic virus (CaMV), are also targeted by cytoplasmic PTGS (Fig. 1.4). Transcript secondary structure and convergent transcripts provide dsRNA substrates for DCL activity (Chellappan et al. 2004, Akbergenov et al. 2006, Moissard and Voinnet 2006). Aberrant transcripts are likely also produced, both as a result of over-expression and by RISC-mediated transcript cleavage. However, unlike RNA viruses, DNA virus transcripts are generated by RNA polymerase II, and so the initial substrates for both primary and secondary siRNA production are made by a cellular enzyme. Also unlike RNA viruses, all four Dicers appear to work in concert to generate the antiviral response, and 24 nt viral siRNAs
produced by DCL3 are the most abundant species in infected tissues (Blevins et al. 2006, Moissard and Voinnet 2006). As DCL3 is a methylation pathway component, this implicates methylation in DNA virus defense, as will be elaborated later.

1.2.3 Transcriptional gene silencing (TGS) and the methylation pathway

The covalent modification of DNA and associated histones is a fundamental epigenetic mechanism that regulates gene expression during normal development, plays a major role in preventing the rearrangement of repeated sequences and in silencing transposons, and distinguishes heterochromatic and euchromatic regions of the genomes. How specific sequences are targeted for repressive methylation is not completely understood, but in some organisms, including the fission yeast *Schizosaccharomyces pombe* and plants, siRNAs guide methylation to homologous sequences in a process termed RNA-directed DNA methylation (RdDM) (Chan et al. 2005, Martienssen et al. 2005, Henderson and Jacobsen 2007, Chan 2008).

Genome-wide profiling studies have shown that cytosine methylation is highly associated with heterochromatic sequences, including centromeric repeats, transposons, and retroelements (Zhang et al. 2006, Cokus et al. 2008, Lister et al. 2008). The scale of the effort to silencing these regions is reflected by the fact that methylation-associated 24 nt siRNAs, most of which map to centromeric and pericentromeric regions, is easily the most abundant size class in wild-type *Arabidopsis* (Kasschau et al. 2007). Methylation, especially in promoter regions, can silence transcription, and it is somewhat counterintuitive that TGS can be triggered and maintained by transcription of inverted or tandemly repeated sequences. TGS can also be initiated experimentally by ectopic
expression of RNA corresponding to promoters (Mette et al. 2000, Jones et al. 2001). Key players in the RdDM pathway include DCL3 and DRB3, as well as RDR2 and AGO4 (Hamilton et al. 2002, Zilberman et al. 2003, Chan et al. 2004, Lippman et al. 2004, Xie et al. 2004, Zilberman et al. 2004, Qi et al. 2005, Tran et al. 2005, Raja unpublished). In addition, the plant-specific RNA polymerases IV and V (formerly Pol IVa and Pol IVb, respectively) are also central to methylation (Herr et al. 2005, Kanno et al. 2005, Onodera et al. 2005, Qi et al. 2005). Pol IV and Pol V are not essential for viability in Arabidopsis, but appear to be specialized forms of RNA Pol II that have evolved to function in siRNA-mediated methylation (Ream et al. 2009). Because they likely have the ability to transcribe methylated DNA, their activities may provide an answer to the paradox that transcription is required for TGS (Pikaard et al. 2008). (Pol IV and Pol V and their roles in RdDM will be discussed in detail in Section 1.3.4)

A model for siRNA biogenesis and methylation posits that Pol IV transcribes ssRNA from transposons and other methylated and repressed sequences, likely including geminivirus genomes (Fig. 1.4). These non-coding transcripts are recruited to sites of ribonucleoprotein complex assembly in the nucleolus (Cajal bodies) where they are converted to dsRNA by RDR2. The dsRNA is processed by DCL3 and incorporated into an AGO4-containing RISC complex which leaves the nucleolus and, programmed by the siRNA, directs methylation of complementary sequences, presumably by recruiting cytosine and histone methyltransferases (Li et al. 2006, Pontes et al. 2006). Downstream components include the SWI/SNF2 chromatin remodeling enzyme, Defective in RNA-directed DNA methylation 1 (DRD1) (Kanno et al. 2004, Chan et al. 2006), and Domains Rearranged Methyltransferase 1 and 2 (DRM1, DRM2), cytosine methyltransferases.

An important issue is how siRNAs guide RISC to specific sequences. Genomic DNA might unwind to allow base pairing with siRNA, or transcription might produce a nascent, non-coding scaffold RNA to which siRNAs could pair. Either mechanism provides a sequence-specific interaction for docking AGO4-containing RISC, where it could serve as platform to recruit chromatin-modifying enzymes. Recent evidence from \textit{S. pombe} and plants favors the scaffold model, with RNA Pol V likely producing the scaffold transcript in \textit{Arabidopsis} (Buhler et al. 2006, Irvine et al. 2006, Wierzbicki et al. 2008, Wierzbicki et al. 2009). In addition to recruiting methyltransferases, AGO4 also has slicer activity, and sliced scaffold transcripts could serve as a template for RDR-mediated dsRNA production that could amplify siRNAs to reinforce and maintain methylation (Irvine et al. 2006, Qi et al. 2006). Unlike mammals, plants can methylate cytosines in CG, CNG, or CHH (where H can be a C, T, or A) sequence contexts. Interestingly, the RdDM pathway results in \textit{de novo} cytosine methylation in all sequence contexts and is carried out by DRM1/DRM2 (Cao et al. 2002).

The RdDM pathway that results in \textit{de novo} DNA methylation is sustained from through generations by Cytosine Methyltransferase 3 (CMT3) and Methyltransferase 1 (MET1). Although there is some overlap and redundancy among the methyltransferases, CMT3 and MET1 primarily maintain CNG and CG methylation, respectively (Bartee et al. 2001, Lindroth et al. 2001, Jones et al. 2001, Cao et al. 2002). Thus \textit{de novo}
methylation is characterized by the presence of 24 nt siRNAs and CHH methylation, while maintenance corresponds to methylation mainly at CNG and CG sites. Once the initial trigger is removed, the SWI/SNF2 chromatin remodeler Decrease in DNA Methylation 1 (DDM1) together with Variant in Methylation 1 (VIM1) recruit MET1 to maintain methylation at hemimethylated CG sites (Jeddeloh et al. 1999, Woo et al. 2007). CMT3 works together with the H3K9 dimethyltransferase (H3K9me2) SUVH4 (Su(VAR)3-9 Homolog 4) in a positive feedback loop, where CMT3 binds methylated histone peptides \textit{(in vitro)} and SUVH4 binds methylated CNG (Jackson et al. 2002, Malagnac et al. 2002, Jackson et al. 2004, Lindroth et al. 2004). Thus, H3K9me2 is associated with CNG methylation throughout the genome, and particularly at centromeric and pericentromeric regions (Bernatavichute et al. 2008). Further, CG methylation, which is more robustly maintained, appears to coordinate non-CG methylation and other epigenetic events, allowing them to be stably inherited (Mathieu et al. 2007). Together the \textit{de novo} and maintenance RdDM pathways coordinate to preserve heterochromatin and defend the host against foreign DNA from both transposons and viruses.

1.2.4 Systemic spread of silencing signals

A remarkable feature of RNA silencing is the ability of a silencing signal to spread cell-to-cell and systemically in the host plant, a process that can immunize tissues in advance of an invading virus (Palaqui et al. 1997, Himber et al, 2003, Voinnet 2005). Once the silencing signal is produced it first begins to travel to adjacent cells (local spread, 10-15 surrounding cells). In some cases, the signal can then spread systemically throughout the plant, moving from source to sink through the phloem. Subsequently, the
signal can spread locally, resulting in efficient silencing throughout the entire plant. Unfortunately, the mechanistic details and the source of the mobile RNA are still the subject of much debate (Reviewed in Liang et al. 2011, Melnyk et al. 2011, Molnar et al. 2011). Moreover, short distance and long distance spread appear to be mechanistically distinct pathways due to the fact that viral suppressors and cadmium differentially inhibit these processes (Ueki and Citovsky 2001, Hamilton et al. 2002, Himber et al. 2003).

Similar to viruses, local silencing signals likely travel through plasmodesmata (symplastic movement) (Lough and Lucas 2006). This is supported by the observation that local spread is excluded from guard cells, which are symplastically isolated from mesophyll cells (Voinnet et al. 1998). Recently, screens were conducted to identify the genetic requirements for generating and spreading the local silencing signal. Transgenic plants expressed an inverted-repeat under the control of a phloem-specific promoter. Local spread of silencing resulted in photobleaching in the cells surrounding the veins (Himber et al. 2003, Dunoyer et al. 2005, Dunoyer et al. 2007, Smith et al. 2007). Both of these screens recovered mutations in miRNA, tasiRNA and RdDM pathway components. More specifically, DCL1, DCL4, AGO1, and HEN1 were all identified from the miRNA and tasiRNA pathways (Dunoyer et al. 2005, Yang et al. 2006). Interestingly, the RdDM pathway components Pol IV, RDR2, and CLSY1 were also found to have an important function in local spread (Himber et al. 2003, Smith et al. 2007). Grafting experiments were then used to determine the role of these genes in the production or perception of the long distance signal (Schwach et al. 2005, Brosnan et al. 2007). Pol IV, RDR2, DCL3, and AGO4 RdDM pathway components were found to be required for reception of the long distance silencing signal (Brosnan et al. 2007).
Interestingly, these experiments also identified a requirement for RDR6 in receiving and probably amplifying the silencing signal (Schwach et al. 2005, Brosnan et al. 2007). It is not yet understood how the PTGS and TGS pathway components coordinate to produce and perceive these mobile signals.

As some groups have worked to identify the genes involved in spread, others have attempted to identify the RNA species that acts as the mobile signal. It is believed that an RNA species mediates systemic silencing because its effects are nucleotide-sequence specific. Thus, the signal could be composed of long dsRNAs, long ssRNAs, the 21 to 24 nt single-stranded sRNAs, or the 21 to 24 nt dsRNAs. Since, dicer proteins were all identified in the grafting and genetic experiments described above, it seemed logical to perform grafting experiments using wild-type plants as the signal source and dcl mutants as the recipient tissue. High-throughput sequencing was then used to show that transgene-derived and endogenous 22 and 24 nt siRNAs were able to move from the wild-type shoot to the mutant root. This suggests that 22 and 24 nt siRNAs (DCL2 and DCL3 products, respectively) can systemically spread throughout the plant (Molnar et al. 2010). In addition, dcl4 mutants were deficient in production of the local silencing signal and this phenotype was complemented by a DCL4 transgene, suggesting that 21 nt siRNAs also act as the mobile element for this process. Moreover, the observation that p19, which binds to 21 nt siRNA duplexes, was able to block spread suggests that a 21 nt dsRNA spreads silencing (Dunoyer et al. 2010). Altogether, these results suggest that 21 to 24 nt siRNAs duplexes act as the mobile signal. Moreover, it appears that multiple silencing pathways are involved in the spread of silencing and these pathways may differ
for local and systemic spread, as well as for the silencing target (e.g. transgene, endogenous gene, or viruses).

Systemic silencing plays a very important role in viral defense and may lead to a phenomenon known as recovery. Initially infected tissue exhibits viral symptoms, but as the plant grows the new tissue is disease free, i.e. recovered. These tissues are resistant to secondary infections by an RNA-mediated defense pathway (Ratcliff et al. 1997). In addition, several viral silencing suppressors, including p19 encoded by *Cymbidium ringspot virus* (CymRSV) and the 2b protein of *Cucumber mosaic virus* (CMV), can prevent the systemic spread of silencing, providing further evidence for its importance in antiviral defense.

1.2.5 TGS as a defense against geminiviruses

Given that plants use methylation to suppress the activity of endogenous invasive DNAs such as transposons, it is reasonable that they would also use methylation to repress viral minichromosomes. Early support for this came from studies which demonstrated that *in vitro* methylation of geminivirus DNA greatly reduced replication in tobacco protoplasts, mostly by inhibiting transcription (Brough et al. 1992, Ermark et al. 1993). Other indirect evidence followed. Transcriptional silencing of transgenes driven by the CaMV 35S promoter was observed in CaMV infected plants (Al-Kaff et al. 1998). Similarly, it was found that transgenes driven by geminivirus promoters could be hypermethylated and transcriptionally silenced in plants inoculated with cognate, but not heterologous geminiviruses (Seemanpillai et al. 2003, Bian et al. 2006). These observations suggested that sequence-specific signals capable of eliciting TGS are
generated during infection, and that these signals can target homologous promoters in cellular chromatin. It was also reported that geminivirus-infected plants can recover from infection following transient expression of a construct designed to express dsRNA corresponding to the viral IR, the master control region that contains divergent early and late gene promoters. Thus, TGS signals could also act on promoters in replicating virus chromatin (Pooggin and Hohn 2003). Similar experiments further suggested that silencing was associated with cytosine and H3K9 methylation of the viral genome (Bian et al. 2006, Dogar 2006).

A recent study has provided direct, genetic evidence for chromatin methylation as an epigenetic defense against geminiviruses (Raja et al. 2008). This study showed that methylation-deficient *Arabidopsis* mutants exhibit hypersensitivity to the begomovirus CaLCuV and the curtovirus BCTV. All of the methylation pathway mutants examined, including *nrpd2a* (deficient for both Pol IV and Pol V), *ago4*, and *ddm1*, proved hypersusceptible to CaLCuV and BCTV. Similar enhanced susceptibility was observed in mutant plants lacking the non-CG methyltransferases *drm1/drm2* and *cmt3*, and the H3K9 methyltransferase *kyp2/suvh4*. Hypersusceptibility was also evident in adenosine kinase mutants (*adkl* and *adk2*). ADK increases the efficiency of the methyl cycle that generates the essential methyltransferase co-factor S-adenosyl methionine (SAM), and is a target of the CaLCuV AL2 and BCTV L2 silencing suppressors (discussed further in 1.2.6 and 1.2.7). Experiments with *met1* mutants (CG maintenance methyltransferases) proved difficult because a complete loss of MET1 functions results in severe growth defects. However, *met1* heterozygotes exhibit normal growth and showed moderately enhanced susceptibility to geminiviruses. Importantly, bisulfite sequencing revealed that
increased disease severity was in all cases accompanied by reduced cytosine methylation in the CaLCuV and BCTV intergenic regions. Compared to wild-type plants, total methylation was reduced approximately 10 to 25%, depending on the mutant. That larger reductions were not observed is testimony to the functional redundancy of silencing pathway components. Nevertheless, these findings suggest that CG methylation is an important aspect of host defense against geminivirus, since even relatively modest reductions are sufficient to cause a severe infection phenotype (Raja et al. 2008).

Analysis of histone methylation has so far been limited to wild-type plants. The available evidence indicates that both abundant H3K9me2 and H3 acetylated at K9 and/or K14 are associated with the CaLCuV IR. Although it is possible that viral genomes contain mixed histone modifications, the presence of both active (acetylated H3) and repressive (H3K9me2) marks associated with the viral control region suggests that populations of active and repressed genomes co-exist in infected plants (Raja et al. 2008). This would be consistent with the detection of viral chromosomes with different nucleosome densities, suggesting relatively open or closed chromatin (Pilartz and Jeske 2003).

While the sensitivity of methylation-deficient mutants to geminiviruses is quite striking, the most compelling argument for methylation as an antiviral defense comes from studies that have associated methylation with host recovery from infection. Recovery is a phenomenon where new tissues or shoots initiated after the establishment of a symptomatic infection exhibit symptom remission and contain very little virus. Recovery has been observed in several geminivirus-host combinations. For example, wild-type *N. benthamiana* and *Arabidopsis* plants can recover from infection with BCTV
L2- mutant virus, which lacks a pathogenicity factor that opposes methylation (discussed in more detail in 1.2.7) (Hormuzdi and Bisaro 1995, Wang et al. 2003).Remarkably, however, Arabidopsis ago4 plants cannot recover from infection with the mutant virus, conclusively linking recovery and L2 function with the methylation pathway. Further, the IR of BCTV L2- genomes obtained from recovered tissues of wild-type plants was found to be hypermethylated, with nearly 80% of cytosines in all sequence contexts affected (Raja et al. 2008).

Increased methylation of viral DNA, as judged by methylation sensitive restriction analysis, has also been observed following recovery of watermelon plants from infection with the begomovirus Cucurbit leaf crumple virus. In addition, symptom remission was enhanced in zucchini plants (which normally exhibit weak recovery) by transient expression of dsRNA corresponding to the viral IR (Hagan et al. 2008). Another study found that recovery of pepper plants from Pepper golden mosaic virus infection was likewise accompanied by increased viral DNA methylation. Interestingly, bisulfite sequencing revealed that cytosine methylation was denser in the IR than in the CP coding region. Bisulfite analysis of the entire Tomato yellow leaf curl-China virus (TYLCCNV) also found cytosine methylation was concentrated at promoter regions (Yang et al. 2011b). Consistent with this observation, the bulk of the 24 nt viral siRNAs in infected plants maps to the IR, while most of the 21 and 22 nt PTGS-associated viral siRNAs originate from coding regions (Rodriguez-Negrete et al. 2009, Yang et al. 2011a).
1.2.6 Suppression of PTGS by geminivirus proteins

Perhaps the strongest evidence that RNA silencing is a potent antiviral defense is the fact that all plant viruses examined to date encode at least one protein that has silencing suppressor activity. Viral suppressors are structurally diverse and can affect different and sometimes multiple points in silencing pathways. Their purpose is to promote viral invasiveness by enhancing virus replication in infected cells, and/or by inhibiting local or long distance spread of antiviral silencing. They may act by interfering with small RNA production, by binding and sequestering small RNA, or by directly or indirectly inhibiting the activity of silencing-related proteins (for review see Li and Ding 2006, Diaz-Pendon and Ding 2008). Many RNA virus suppressors have been identified and characterized, and examples are shown in Fig. 1.4.

Compared to suppressors encoded by RNA viruses, relatively little is known about the activities of DNA virus silencing suppressors, but what is understood so far indicates that they are quite unique. The Baulcombe laboratory first showed the geminivirus AL2 protein could reverse established PTGS (Voinnet et al. 1999). AL2 had initially been characterized as a transcription factor that stimulates the expression of virus late genes by both activation and derepression mechanisms (Sunter and Bisaro 1992, Sunter and Bisaro 1997, Sunter and Bisaro 2003, Lacatus and Sunter 2008). AL2 has a basic N-terminal region that contains a nuclear localization signal (NLS), a central zinc finger-like domain (CCHC motif), and a C-terminal acidic-type transcription activation domain (Hartitz et al. 1999, van Wezel et al. 2001, van Wezel et al. 2003). However, AL2 is a non-canonical transcription factor in that it only weakly binds ssDNA and dsDNA in a sequence non-specific manner, and is likely recruited to responsive promoters by
interacting with cellular proteins (Hartitz et al. 1999, Lacatus and Sunter 2008, Lacatus and Sunter 2009). AL2 does not bind siRNA or miRNA (Chellappan et al. 2005, Wang et al. 2005). Rather, several studies have demonstrated that the ability of AL2 to reverse PTGS and inhibit its systemic spread requires an intact NLS, the CCHC motif, and the transcription activation domain, suggesting that silencing suppression in these contexts requires AL2 to stimulate transcription. More specifically, it is presumed that AL2 activates the transcription of cellular genes that negatively regulate silencing pathways (van Wezel et al. 2002, Dong et al. 2003, Trinks et al. 2005). This mode of silencing suppression has been termed transcription-dependent, although interactions with silencing pathway components through the activation domain cannot formally be ruled out.

That AL2 can, in principle, alter the host transcriptome comes from studies showing that geminivirus infection can activate the expression of reporter transgenes driven by the cognate viral CP promoter in an AL2-dependent fashion (Hong et al. 1997, Sunter and Bisaro 1997). Transcriptional profiling following transient expression of AL2 in Arabidopsis protoplasts identified several genes whose expression was increased. One of these was Werner Exonuclease-like 1 (WEL1), a homologue of Werner syndrome-like Exonuclease (WEX) (Trinks et al. 2005). Although its role is unclear, WEX is required for PTGS (but not TGS) directed against transgenes (Glazov et al. 2003), and it has been proposed that WEL1 over-expression might compete for factors needed for WEX function (Trinks et al. 2005). Direct evidence to support this hypothesis, and for up-regulation of WEL1 transcription during virus infection, is needed. Nevertheless, the
transcription-dependent suppression mechanism in general is experimentally well supported.

Evidence for transcription-independent silencing suppression by AL2 and the related L2 protein from the curtovirus BCTV has also been obtained (Wang et al. 2005). AL2 and L2 share a similar genome position and overlap, in different reading frames, the AL1/L1 and AL3/L3 genes that are highly conserved between begomoviruses and curtoviruses (Fig. 1.1). However, L2 shares little direct homology with its begomovirus counterparts, except for the central CCHC motif. In addition, unlike AL2, L2 lacks an obvious activation domain and is not a transcription factor (Stanley et al. 1992, Sunter et al. 1994, Hormuzdi et al. 1995). Despite these differences, AL2 and L2 share pathogenicity functions, and both interact with and inactivate SNF1-related kinase (SnRK1) and adenosine kinase (ADK) (Sunter et al. 2001, Hao et al. 2003, Wang et al. 2003, Baliji et al. 2007). The SnRK1 interaction inhibits the cellular stress response, which appears to be a component of plant basal defense (Hao et al. 2003). However, the interaction with ADK, a nucleoside kinase that phosphorylates adenosine to 5'-AMP, offers several advantages, and of these silencing suppression is the most relevant to this thesis.

AL2 and L2 inactivate ADK in vitro and when co-expressed with ADK in E. coli and yeast, and ADK activity is reduced in an AL2/L2-dependent manner in geminivirus-infected plant tissues (Wang et al. 2003). Bimolecular fluorescence complementation studies (BiFC) in N. benthamiana cells have also demonstrated that dimer AL2 moves to the nucleus to activate transcription, whereas AL2:ADK and L2:ADK complexes accumulate mainly in the cytoplasm (Yang et al. 2007). A link between ADK activity
and silencing comes from the observation that ADK is needed to maintain the methyl cycle responsible for generating S-adenosyl methionine (SAM), a methyl donor and essential methyltransferases cofactor (Lecoq et al. 2001, Weretilnyk et al. 2001, Moffatt et al. 2002) (Fig. 1.5). It is likely for this reason that ADK-deficient mutant plants display silencing defects (Moffatt et al. 2002). A connection between methylation and PTGS was established by early observations that post-transcriptional silencing is frequently associated with methylation of the coding regions of targeted genes, and particularly reporter transgenes (Ingelbrecht et al. 1994, Smith et al. 1994, English et al. 1996, Jones et al. 1998). Moreover, HEN1 is a SAM-dependent methyltransferase that plays an important role in siRNA stability, thus inhibition of ADK could indirectly reduce HEN1 activity.

Evidence that AL2 and L2 can suppress PTGS in a transcription-independent manner by ADK inhibition was obtained from studies of transient local silencing in *N. benthamiana* leaves. These experiments showed the silencing directed against GFP could be suppressed by wild-type AL2 and L2 proteins, and also by a missense mutant (AL2 C33A) impaired for transcription activation or a truncated mutant lacking the activation domain (AL21-114) (Wang et al. 2005, Yang et al. 2007). Further, local suppression could be phenocopied by ADK knockdown and by an adenosine analogue (A-134974) that inhibits ADK activity (Wang et al. 2005). Knock-down of S-adenosyl homocysteine hydrolase (SAHH) also results in suppression of PTGS directed against GFP (R.C. Buchmann and D.M. Bisaro, unpublished). Thus both intrinsic (SAHH) and associated (ADK) methyl cycle enzymes are important for PTGS, and AL2 and L2 inhibit the methyl cycle and PTGS in a transcription-independent manner by inactivating ADK. To
summarize, begomovirus AL2 proteins appear to suppress silencing by activating the expression of cellular genes and by inhibiting ADK, whereas curtovirus L2 is limited to the latter mechanism (Bisaro 2006).

Other geminivirus proteins have been shown to suppress PTGS, including the AL4 protein encoded by begomoviruses, the V2 protein in the monopartite begomoviruses, as well as the βC1 gene from the DNA β satellite (Cui et al. 2004, Saunders et al. 2004, Vanitharani et al. 2004, Chellappan et al. 2005, Zrachya et al. 2007, Glick et al. 2008, reviewed in Bisaro, 2006). Since the majority of the work described in this thesis focuses on the suppressor activities of AL2 and L2, the other geminivirus PTGS suppressors will not be further discussed.

1.2.7 Suppression of TGS by geminivirus proteins

Given that RdDM is a critical defense and that AL2 and L2 inhibit a methyl cycle enzyme, it was hypothesized that these proteins could suppress TGS. A recent study used two different experimental strategies to address this question (Buchmann et al. 2009). The first approach used dsRNA designed to target the 35S promoter driving a GFP transgene this induced heritable, epigenetic silencing, and this line of *N. benthamiana* plants generated was called 16-TGS. Line 16-TGS plants show no GFP expression in a nuclear run-on assay. However, over-expression of TGMV AL2 and BCTV L2 proteins, geminivirus infection, or knock-down of methyl cycle enzymes (ADK, SAHH, and MET1) cause reversal of TGS, resulting in GFP expression. This work was supported by a second experimental approach that employed dexamethasone-inducible promoters to drive expression of transgenes in *Arabidopsis*. AL2, AL2_{1-114} (mutant deficient in
transcription activation), L2, and dsADK transgenes were induced by dexamethasone and expression of silenced endogenous genes and transposons was analyzed by RT-PCR. The dsADK transgene was designed to produce dsRNA that could target ADK mRNA by RNA interference. Interestingly, induction of AL2\textsubscript{1-114}, L2, and dsADK led to the expression of an F-box pseudogene, AtSN1 and Athila retrotransposons, but only AL2 with its ability to activate transcription could also cause expression of the CACTA-like DNA transposon. These results suggested the AL2 and L2 reverse TGS by inhibiting ADK in a transcription-independent manner, and that AL2 also reverses TGS by a transcription-dependent mechanism. Furthermore, induction of AL2 and L2 expression resulted in a global reduction in cytosine methylation in the transgenic Arabidospis plants (Buchmann et al. 2009).

The importance of L2 in RdDM inhibition is highlighted by the fact that wild-type plants infected with BCTV L2- mutant viruses show a novel recovery phenotype. Recovery occurs when the primary symptomatic tissue 21 days post-infection (dpi) is removed and plant is allowed to continue growth. Removal of the shoot apical meristem activates axilllary meristems and new, secondary shoots appear. Normally, BCTV infection results in highly symptomatic secondary growth, but infection with BCTV L2-results in secondary shoots with little to no symptoms, thus this phenomenon was termed recovery (Hormuzdi et al. 1994). Southern blot analysis of secondary shoot extracts demonstrated that recovery was associated with significant reductions in viral DNA levels. This recovery event was correlated with the RdDM defense pathway when it was determined that methylation mutants, such as ago4, are no longer able to recover from BCTV L2- infections. Furthermore, bisulfite analysis of BCTV IR DNA isolated from
wild-type recovered tissue contained ~70% methylated cytosines, while virus obtained from non-recovered ago4 plants was hypomethylated, containing only a 40% methylation (Raja et al. 2008). Together, these results underscore the significance of RdDM as a host defense against geminiviruses.

Strikingly, the monopartite begomovirus *Tomato yellow leaf curl China virus* (TYLCCNV) AL2 protein is deficient in methyl cycle inhibition. However, when TYLCCNV is isolated from symptomatic tomato plants in the field it is always accompanied by a betasatellite DNA (TYLCCNB, a.k.a. DNAβ). TYLCCNB depends on TYLCCNV for replication and encapsidation, but in return it provides a single protein product, βC1. βC1 induces disease symptoms, suppresses PTGS, and results in greatly increased virus titers (Saunders et al. 2004, Cui et al. 2005, Saeed et al. 2005, Yang et al. 2008). A recent study found that βC1 can also reversing TGS. TYLCCNV alone does not reverse TGS of the GFP transgene in line 16-TGS plants, but addition of DNA β or the over-expression of satellite encoded βC1 results in GFP expression in this TGS reversal assay. TYLCCNV with TYLCCNB and over-expression of βC1 alone also reduced global genomic DNA methylation similar to BCTV L2 and TGMV AL2. Interestingly, TYLCCNB was able to complement the BCTV L2- mutation in 16-TGS plants and resulted in loss of the recovery phenotype typical of BCTV L2- infection. Finally a yeast two-hybrid candidate screen found that βC1 interacted with the critical methyl cycle enzyme SAHH. The significance of this interaction was determined when *in vitro* biochemical assays showed that purified βC1 inhibits SAHH enzymatic activity (Yang et al. 2011). Thus, the betasatellite has co-evolved with its helper virus,
performing suppressor functions that may allow TYLCCNV AL2 to perform other tasks more efficiently.

Similar to TYLCCNV, the bipartite begomovirus Pepper golden mosaic virus (PepGMV) may encode an AL2 protein that is deficient for TGS suppression. PepGMV undergoes natural recovery. In this case infected pepper plants begin to show symptoms ~10 dpi, but undergo natural recovery around 15 dpi leaves at the top of the plant appear asymptomatic. Loss of symptoms is associated with reductions in viral titer, as well as increased methylation of the IR (Rodriguez-Negrete et al. 2009). Thus, the link between recovery and RdDM suggests that the PepGMV AL2 protein may not be able to efficiently inhibit methylation and is unable to reverse TGS. Another example of natural recovery is seen in ACMV infected cassava. It is known that ACMV encodes a relatively strong PTGS suppressor (AC4), however its AL2 suppressor is fairly weak, and plants recover from infection. Interestingly, a particularly severe cassava disease outbreak in Uganda was caused by mixed infections of ACMV and East African cassava mosaic virus (EACMV), which encodes a relatively strong AL2 TGS suppressor (Vanitharani et al. 2004). These observations support the idea that geminiviruses are constantly evolving and combining with satellite DNAs or other viruses to generate productive infections, and this requires suppression of TGS.

Interestingly, another curtovirus, Beet severe curly top virus (BSCTV) targets yet another methyl cycle enzyme. SAM, the methyl donor for transmethylation, can be converted to decarboxylated SAM (dcSAM) by SAM decarboxylase (SAMDC). dcSAM is an aminopropyl donor for the biosynthesis of spermidine and spermine, which promote plant growth and cellular metabolism (Mattoo et al. 2010). Due to structural similarities,
dcSAM may compete with SAM for methyltransferase interaction and therefore may competitively inhibit transmethylation (Heby et al. 1988). Consequently, cellular levels of dcSAM must be regulated and this occurs by transcriptional, translational, and posttranslational control of SAMDC (Shantz et al. 1992, Stanley et al. 1994, Hanfrey et al. 2003, Yerlikaya and Stanley 2004). A yeast two-hybrid screen with BSCTV C2 (homologous to L2) identified SAMDC as an interaction partner, and this interaction was shown to inhibit the 26S proteasome-mediated degradation of SAMDC. Thus, C2 stabilizes SAMDC, resulting in an increase in the dcSAM/SAM ratio, and competitively inhibiting SAM mediated transmethylation (Zhang et al. 2011) (Fig. 1.5). Whether BCTV L2, BCTV C2/L2, and begomovirus AL2 interact with both ADK and SAMDC remains to be determined. Altogether, recent work has shown that geminiviruses have evolved multiple strategies to combat TGS defense mechanisms, highlighting the importance of this host defense pathway. In addition, targeting methyl cycle enzymes appears to be a particularly expedient means to inhibit methylation-mediated TGS.

1.2.8 Novel suppressor activities of the geminivirus AL2 protein

Although AL2 and L2 share suppressor activities, AL2 is known to suppress PTGS and TGS using multiple mechanisms (Bisaro 2006). Inhibition of ADK, a methyl cycle cofactor, is a well-described mechanism for suppression of both TGS and local silencing by PTGS (Wang et al. 2005, Yang et al. 2007, Raja et al. 2008, Buchmann et al. 2009). This mechanism is shared by AL2 and L2 and is independent of transcription activation. On the other hand, AL2-mediated transcription activation of endogenous negative regulators of RNA silencing has also been proposed to suppress both PTGS and TGS
(van Wezel et al. 2003, Trinks et al. 2005, Buchmann et al. 2009). Thus, AL2 can suppress RNA silencing using transcription -independent and -dependent mechanisms, which have yet to be thoroughly investigated. The goal of Chapter 2 of this thesis was to determine whether, and in general terms how, AL2 and L2 inhibit the systemic spread of silencing or cause the reversal of both PTGS and TGS. It is anticipated that analysis of AL2 and L2 suppression activities will lead to a better understanding of host silencing and methylation machineries.

1.3 Chromatin structure and epigenetics

Chromatin compaction is essential for packaging DNA into the nucleus, but it also plays a critical role in all nuclear processes including replication, transcription, recombination, repair, genome stability and chromosome segregation. A wide variety of chemical modifications to both DNA and histones modulate the chromatin state and result in a highly dynamic structure. The collection of these chemical modifications is called the epigenome, which translates to 'above the genome' and refers to any change that results in a heritable phenotype without directly changing the nucleotide sequence (reviewed in Berger 2007, Kouzarides 2007, Pluger and Wagner 2007, Roudier et al. 2009).

1.3.1 Nucleosome structure and assembly

Chromatin compaction begins with the folding of 147 bp of DNA around a single nucleosome formed by an octamer of H3, H4, H2A, and H2B canonical histones (two
copies of each) (Luger et al. 1997). Histones consist of two domains. The majority of the amino acid sequence contributes to the histone fold domain, while the N- and C-terminals of histones are relatively unstructured tail domains. The fold domain constitutes the nucleosome core, with the H3-H4 histones forming a dimer-of-dimers, which in turn interacts with 2 H2A-H2B heterodimers (Arents et al. 1991, Davey et al. 2002). The minor groove of the DNA wraps around this core nucleosome structure via non-sequence specific hydrogen bonds (Luger and Richmond 1998). The unstructured tail domains of the histones protrude out from the protein-DNA core and their function will be discussed in detail in 1.3.2. Since these canonical nucleosomes package, protect, and supercoil DNA, it is not surprising that nucleosome assembly is linked to DNA replication. Canonical histone (H3, H4, H2A, and H2B) expression occurs during S phase of the cell cycle. As helicases separate the two strands of DNA ahead of the replication fork, nucleosomes are displaced. Behind the replication fork the daughter DNA strands will associate with displaced nucleosomes, as well as newly synthesized nucleosomes. First, DNA wraps around the H3-H4 tetramer and this recruits two H2A-H2B heterodimers, which bind to each side of the H3-H4 (Widom 1999, Andrews et al. 2010). This disassembly and assembly process is assisted by histone chaperones, which ensure proper histone-DNA interactions (Laskey et al. 1978, Fyodorove and Kadonaga 2003, Andrews and Lager 2011a, reviewed in Andrews and Lager 2011b).

The repetitive structure of DNA wrapped around nucleosomes is the first layer of chromatin compaction and appear as beads on a string by electromicrograph. Chromatin can then be condensed further by the linker histone H1, which binds to the short region of DNA between the reiterating nucleosomes and allows the chromatin to fold into a helical
filament. This converts the 'beads on a string' into a 30 nm fiber. Subsequently, a complex and poorly understood succession of folding and looping events is mediated by scaffolding proteins completes the chromatin architecture. Although chromatin can be heavily condensed, it must be a dynamic structure in order for proteins to recognize, transcribe, and replicate DNA. Thus, DNA is organized into regions of heavily condensed chromatin, heterochromatin, which is not conducive for transcription. Alternatively, euchromatin is relatively open and promotes gene expression, while other regions containing genes primed for transcription are maintained in a state between hetero- and eu-chromatin. All of these regions are associated with specific histone post-translational modifications, DNA methylation patterns, and histone variants (Berger 2007, Kouzarides 2007). Each of these marks modulate chromatin compaction and are acted on by chromatin readers and remodelers, effector protein complexes that recognize and reinforce specific modifications and/or allow access to DNA for transcription. Therefore, regulating chromatin structure with specific epigenetic marks is critical for controlling gene expression, cell cycle regulation, and ultimately development.

1.3.2 Histones modifications and variants

Specific alterations to the histone tails and exchange of the core histones for variants can modulate nucleosome association with DNA (reviewed in Pfluger and Wagner 2007, Berger 2007, Roudier et al. 2009, Law and Jacobsen 2010, Deal and Henikoff 2011). Histone post-translational modifications (PTMs) can occur prior to the formation of a nucleosome or after the incorporation into DNA. Greater than 60 different amino acid residues of the four core histones have been shown to have PTMs and the unstructured
tails of H3 and H4 are especially rich. In particular, lysine residues can be acetylated, ubiquitinylated, sumoylated, or methylated. Arginine residues can be methylated and serine and threonines can be phosphorylated (Fig 1.6). Additional complexity is added since arginine and lysine can be mono-, di-, and trimethylated. Moreover, the location of the modification, in the promoter, 5' or 3' end of the gene, can determine the transcriptional response (Bernstein et al. 2007).

Recent genome-wide mapping has revealed that certain modifications cluster together and are associated with either silenced genes and transposons or actively transcribed genes (Berger 2007, Li et al. 2007, Kouzarides 2007). The role of some marks seems to be more defined. Phosphorylation of serine and threonine and methylation of arginine are associated with active genes, while sumoylated lysine marks repressed genes. Acetylation and sumoylation may be mutually antagonistic. Acetylation of histone tails (e.g. H3K9/14 or H4K8/16) loosens the association with DNA by altering the charge of the nucleosomes, thus acetylation is associated with active genes and generally occurs in the promoter regions (Shogren-Knaak et al. 2006). Interestingly, methylated lysine residues are recognized by effector proteins either to mediate transcriptional repression or promote transcription initiation and elongation (Francis et al. 2004, Pray-Grant et al. 2005, Sims et al. 2007, Eskeland et al. 2010). Thus, methylation of lysine residues has diverse effects on transcription and has been extensively studied.

Dimethylated H3K9 (H3K9me2) is a highly studied modification that is associated with cytosine methylation (5mC) and H4K20me1, all of which are enriched in heterochromatin, as well as repeat elements throughout the genome (Lippman et al. 2004, Zhang, et al. 2006, Turck et al. 2007, Vaughn et al. 2007, Bernatavichute et al. 2008,
Zhang et al. 2009, Roudier et al. 2011). Interestingly, H3K9me2 and 5mC usually occurs independently from the repressive H3K27me1, which also plays a role in silencing repeat loci (Jacob et al. 2009). Interestingly, 5mC in promoters is associated with repression, while moderately expressed genes contain 5mC throughout the gene body (Zhang et al. 2006, Zilberman et al. 2006). In contrast to the heterochromatic PTMs, H3K27me3 is associated with promoters and transcribed regions of genes (Zhang et al. 2007, Deal and Henikoff 2010). H3K27me3 is antagonistic to gene body DNA methylation and is associated with repression of tissue-specific and developmentally regulated genes (Turck et al. 2007, Zhang et al. 2007, Oh et al. 2008, Jacob et al. 2010). Methylation can also be associated with highly expressed genes, which can be marked with H3K4me3 and H3K36me3 in the promoter and 5' end (Oh et al. 2008, Zhang et al. 2009, Roudier et al. 2011). Actively transcribed genes are also associated with H3K36me2 at the 3' end (Oh et al. 2008). This suggests that these PTMs are important for transcription initiation and elongation, respectively. Interestingly, H3K4me1 and 2 are deposited at genes that suggest a role in transcription activation, but are not associated with highly expressed loci (Oh et al. 2008, Tanurdzic et al. 2008, Zhang et al. 2009, Roudier et al. 2011). That Histone PTMs are reversible allows active chromatin to be repressed and vice versa. For example, the formation of heterochromatin requires methyltransferases for PTMs described above, but also involves deacetylases, ubiquitin proteases, and histone demethylases. The combined efforts of these enzymes together with histone and DNA methyltransferases promotes the formation of heterochromatin (Aufsatz et al. 2002, Sridhar et al. 2007, Deleris et al. 2010, Earley et al. 2010, Law and Jacobsen 2010, Searle et al. 2010, Lu et al. 2010).
Combined with PTMs, histone variants add an extra layer of chromatin complexity. Histone variants, unlike the canonical histones, are typically encoded by a single locus and are expressed constitutively. Beyond genome packaging, histone variants play integral roles in transcription initiation and termination, DNA repair, chromosome segregation and sperm chromatin packaging (reviewed in Ingouff and Berger 2010, Talbert and Henikoff 2010, Deal and Henikoff 2011). Histone variants differ from the canonical histones in amino acid sequence, and some appear to be conserved throughout eukaryotes (Malik and Henikoff 2003). These universal variants are well studied and include CenH3, H3.3, H2A.Z, and H2A.X.

Although some of these variants differ from the core histones by just a few amino acids, their deposition and function are quite different. For example, H3.3 varies from H3 at just 3 amino acids in plants (Malik and Henikoff 2003). H3.3 is mostly deposited independently of replication at promoters, gene regulatory elements and in the transcribed regions of expressed genes (Mito et al. 2005, Deal et al. 2010, Goldberg et al. 2010). Thus, H3.3 plays a role in regulating transcription. Unlike H3.3, H2A.Z differs from H2A at many amino acid residues, but H2A.Z is also deposited independently from replication preferentially into nucleosomes flanking transcription start sites (Suto et al. 2000, Raisner et al. 2005, Mavrich et al. 2008, Zilberman et al. 2008). H2A.Z additionally prevents DNA methylation, thus playing a role in transcription regulation and formation of heterochromatin boundaries (Raisner et al. 2006, Zilberman et al. 2008, Conerly et al. 2010). H3.3 and H2A.Z double-variant nucleosomes are unstable in vivo and have been mapped to the promoter and 5' ends of highly transcribed genes where they likely promote nucleosome turnover (Jin and Felsenfeld 2007, Jin et al. 2009). Even
more specialized than H3.3 and H2A.A is CenH3. CenH3 is incorporated into the centromeres where it is required to form the kinetochores and plays a crucial role in chromosome segregation (Santaguida and Musacchio 2009). Finally, H2A.X, although not well studied in plants, differs from H2A in the C-terminal motif (Malik and Henikoff 2003). In response to double stranded breaks, the H2A.X C-terminus is phosphorylated, which in turn recruits DNA repair machinery (van Attikum and Gasser 2009).

Altogether, PTMs and histone variants have diverse roles in regulating gene expression in particular cell types and in response to various environmental cues, maintaining genome integrity, cell division, and development. But future work will be required to understand the crosstalk between the numerous PTMs and the histone variants. Interestingly, chromatin modifications have been shown to regulate innate immune response genes during pathogen attack (Berr et al. 2012). The role of the various chromatin marks has not been heavily studied in geminivirus defense. Only H3K9me2 and acetylated H3K9/K14 have been extensively analyzed at the geminivirus IR, thus it will be interesting to map the PTMs and histone variants that are associated throughout the viral genome. Moreover, the virus could serve as a model to study how the histone code regulates gene expression and replication.

1.3.3 Chromatin remodelers

Most histone modifications must be recognized by effector proteins in order to regulate the dynamic balance between genome packaging and DNA access. Chromatin remodelers are one example of effector proteins that hydrolyze ATP to move, eject and restructure nucleosomes. Thus, chromatin remodelers together with histone
modifications play important roles in transcription initiation and elongation, DNA replication, recombination and repair (Vignali et al. 2000, Geiman and Robertson 2002). Five basic assets are shared by all remodeler complexes: (1) an affinity for the nucleosome, as well as DNA; (2) specialized domains that recognize specific histone modifications; (3) a DNA-dependent ATPase domain that hydrolyzes ATP in order to break histone-DNA contacts; (4) a domain that regulates the ATPase activity either by itself or through protein-protein interactions; (5) a domain that mediates protein-protein interactions with transcription factors or other chromatin modifying effectors (reviewed in Becker and Hörz 2002, Saha et al. 2006, Clapier and Cairns 2009). Beyond these basic properties, specialized domains allow remodelers to be classified into four distinct families, each of which can be separated into subfamilies (Flaus et al. 2006). The family most relevant to this thesis is the switching defective/sucrose nonfermenting (SWI/SNF) remodelers, which includes the plant specific SNF2-like chromatin remodeling proteins.

SNF2 family members can modulate transcriptional activation or repression depending on the factor and possibly the proteins with which it interacts. Three particular remodelers have been shown to play critical roles in establishing and maintaining heterochromatin by helping direct histone and DNA methylation. Decrease in DNA methylation 1 (DDM1), defective in RNA-directed DNA methylation 1 (DRD1), and classy1 (CLSY1) were all identified independently by forward genetic screens searching for mutations that reduced DNA methylation (Jeddeloh et al. 1999, Brzeski, and Jerzmanowski 2003, Chan et al. 2004, Kanno et al. 2004, Smith et al. 2007, Greenberg et al. 2011). DDM1 mutations result in rapid reductions in cytosine methylation at repeat loci. Gradual reductions of 5mC are also observed at low copy
sequences in subsequent generations (Jeddeloh et al. 1999). These results are consistent with the finding that DDM1 plays a crucial role in recruiting the MET1 CG methyltransferase to hemimethylated CG sites (Woo et al. 2007). DRD1 and CLSY1 are much less studied, but both appear to work in the RdDM pathway with Pol IV and Pol V, respectively. In clsy1 mutants Pol IV is mislocalized, suggesting that CLSY1 possibly works upstream of Pol IV. CLSY1 has also been shown to play a role in the systemic spread of silencing (Smith et al. 2007, Greenberg et al. 2011). DRD1 co-purifies with Pol V subunits and is required for the generation of Pol V-dependent transcripts (Law et al. 2010). Moreover, drd1 mutants are no longer able to recruit Pol V to specific genomic loci (Weirzbicki et al. 2008). Thus, although DRD1 chromatin remodeling activity has not been shown in vitro, it likely plays an important role in recruitment of transcriptional gene silencing machinery to target loci. Considered together, DDM1 is important for maintenance of heterochromatic states, while CLSY1 and DRD1 are believed to be involved in the establishment of TGS. Future work is required to address how DDM1, DRD1, and CLSY1 remodelers target DNA and restructure nucleosomes.

1.3.4 Pol IV and Pol V mediate chromatin formation and spread DNA methylation in plants

Chromatin remodelers allow DNA to be accessible to the transcriptional machinery, but if the DNA is already repressed through histone and DNA modifications why open it for transcription? Paradoxically, establishment and maintenance of heterochromatin requires transcription. In S. pombe, RNA Pol II transcription produces a scaffolding transcript that recruits histone methylation, leading to the formation of heterochromatin
Motamedi et al. 2004, Djupedal et al 2005, Kato et al. 2005). In plants, Pol IV and Pol V are believed to transcribe methylated DNA and mediate transcriptional gene silencing at target loci (Herr et al. 2005, Onodera et al. 2005, Pontier et al. 2005). Pol IV and V are specialized polymerases that evolved from Pol II through gene duplications and sub-functionalization (Ream et al. 2009, Tucker et al. 2011). Pol II, IV and V have 12 core subunits in Arabidopsis thaliana (Ream et al. 2009). The largest subunits of these enzymes are encoded by unique genes designated as NRPB1, NRPD1, and NRPE1, respectively ("NRP" stands for nuclear RNA polymerase and B, D, and E are the second, fourth, and fifth letters in the alphabet; the number 1 indicates the largest subunit. Pol I and Pol III are polymerases A and C, respectively.). Pol IV and Pol V share a second largest subunit, NRP(D/E)2, while Pol II utilizes a distinct NRPB2 gene (Herr et al. 2005; Kanno et al. 2005; Onodera et al. 2005, Pontier et al. 2005). Interaction between the two largest subunits forms the catalytic site for RNA synthesis. The other 10 subunits can be unique or shared among the polymerases, allowing functional diversity (Fig. 1.7) (Ream et al. 2009, Tucker et al. 2011, Tan et al. 2012). These non-catalytic subunits play important structural and regulatory roles in initiation, elongation, termination or RNA processing (Cramer et al. 2008; Werner and Grohmann 2011). Unlike Pol II, Pol IV and V are not essential for plant viability, but Pol IV and V are important for genome stability, development and antiviral and antibacterial defenses through their distinct roles at the beginning and end of RdDM, respectively (reviewed in Matzke et al. 2009, Meyer 2010, Haag and Pikaard 2011, Simon and Meyers 2011) (Fig. 1.8).

Pol IV is believed to initiate the RdDM pathway by transcribing methylated DNA and the chromatin remodeler CLSY1 is possibly required for recruitment (Smith et al. 2007).
Pol IV physically colocalizes with RdDM loci and in pol iv mutants other proteins of the RdDM pathway are mislocalized, suggesting that Pol IV recruits the downstream RdDM components to target DNA (Pontes et al. 2006). Pol IV transcribes DNA to produce non-coding ssRNA, providing a substrate for RDR2. Subsequent action by DCL3 creates 24 nt siRNAs. Although Pol IV transcripts have not been identified and characterized, Pol IV, CLSY1, and RDR2 are required for production of 24 nt siRNAs at ~95% of RdDM target loci (Zhang et al. 2007). The 24 nt siRNAs then associate with an AGO4 containing RISC complex, which can target homologous sequences in DNA for methylation (Fig. 1.8).

Pol V is required for the AGO4-siRNA guided silencing system. Thus, Pol V is not required for the production of siRNAs at most RdDM loci, but plays a crucial role in targeting RdDM (Kanno et al. 2005, Pontier et al. 2005, Mosher et al. 2008). Recently, Pol V-dependent non-coding transcripts were characterized by 5' RACE (rapid amplification of cDNA ends) at multiple RdDM loci (Weirzbicki et al. 2008). These transcripts are either triphosphorylated or capped at the 5' end. Like Pol II the first nucleotide was a purine, but unlike Pol II, the transcripts did not initiate at a specific site. These non-coding RNAs (ncRNAs) spanned promoter regions of hypermethylated loci. Mutations in the active site of Pol V eliminated these transcripts, and chemical cross-linking of the transcripts to Pol V verified that the ncRNAs were Pol V-dependent (Weirzbicki et al. 2008, Haag et al. 2009). Moreover, transcripts were independent of 24 nt siRNA biogenesis, indicating that Pol IV and Pol V work autonomously in parallel pathways that ultimately result in RdDM (Weirzbicki et al. 2008).
Pol V works in concert with several effectors, which can assist in recruitment, elongation, targeting or amplifying RdDM. A complex of proteins termed DDR is believed to recruit or stabilize Pol V association with chromatin (Law et al. 2010). This complex contains the chromatin remodeler DRD1 (discussed in section 1.3.3), as well as Defective in Meristem Silencing 3 (DMS3) and Required for DNA Methylation 1 (RDM1). DMS3 is a plant-specific hinge domain protein similar to structural maintenance of chromosomes (SMC) proteins such as cohesions and condensins. RDM1 is a ssDNA binding protein with a strong preference for methylated DNA (Kanno et al. 2004, Kanno et al. 2008, Ausin et al. 2009, Gao et al. 2010). Gel filtration chromatography and immunoprecipitation followed by mass spectrometry found that Pol V subunits and the DDR complex co-purify with each other (Law et al. 2010). Furthermore, Pol V no longer associated or transcribed target loci in drd1 and dms3 mutants (Weirzbicki et al. 2008, Weirzbicki et al. 2009). More recently, chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) found that Pol V association throughout the genome required the DDR complex (Zhong et al. 2012). Although not identified in interactions with the DDR complex, defective in meristem silencing 11 (DMS11) interacts with DMS3, which stimulates DMS11 ATPase activity. Thus, DMS11 may facilitate chromatin remodeling and Pol V recruitment independent of or together with the DDR complex (Lorkovic et al. 2012).

Once Pol V associates with chromatin, AGO4-siRNA RISC complexes are recruited. Beyond the base-pairing that occurs between the siRNA and the Pol V transcript, AGO4 is also recruited to target loci through interaction with the WG/GW repeats in the C-terminal domain of NRPE1 (Li et al. 2006, El-Shami et al. 2007, Weirzbicki et al. 2009).
Two proteins facilitate the recruitment of the RISC complex to Pol V-scaffold transcripts, involved in De Novo 2 (IDN2, a.k.a. required for DNA methylation 12, RDM12) and Kow domain-containing Transcription Factor 1 (KTF1, a.k.a. RDM3 and SPT5-like). IDN2 binds dsRNAs and may stabilize the interaction between the siRNA and the target transcript (Aeusin et al. 2009, Wierzbicki et al 2009, Law et al. 2010). KTF1 binds RNA in vitro and interacts with AGO4 likely through its WG/GW domain (Bies-Etheve et al. 2009, He et al. 2009, Huang et al. 2009). Together these proteins may recruit and stabilize the association of AGO4-siRNA and Pol V transcripts. This stabilization is important for two reasons. First, the stable interaction allows the RISC complex to recruit DNA and histone methylation machinery. AGO4 interacts with the RDM1 component of the DDR complex, which in turn associates with DRM2, the de novo cytosine methyltransferase, thus linking Pol V transcription and AGO4 association with the DNA methylation machinery (Gao et al. 2010). Second, AGO4 binds small RNAs through its MID and PAZ domains, while the PIWI domain contains an Asp-Asp-His triad responsible for AGO4 slicing activity (Song et al. 2004, Qi et al. 2006). This slicing activity occurs on the long RNAs that are base-paired with the AGO-associated siRNAs (Song et al. 2004, Rivas et al. 2005). Thus, AGO4 mediated slicing produces a substrate for RDR2, ultimately resulting in the biogenesis of secondary siRNAs. This catalytic activity could explain why some RdDM loci have reduced siRNAs in pol v mutants.

Secondary siRNAs and/or further rounds of Pol IV transcription and RdDM lead to the spread of DNA methylation into flanking DNA sequences (Kanno et al. 2008, Voinnet 2008, Ahmed et al. 2011). A recent forward genetic screen was conducted to identify genes involved in the 'transitivity' of RdDM. Using hairpin-induced silencing of
a GFP transgene, Pol IV, RDR2, DCL3, and Pol V were required for the spread of DNA methylation. This work proposed a model where Pol IV transcribed the methylated DNA and RDR2 and DCL3 played an important role in secondary siRNA biogenesis, while Pol V was needed to target RdDM (Daxinger et al. 2009). Presumably, a Pol V nascent transcript cleaved by AGO4 could be amplified by RDR2 and diced by DCL3 to produce secondary siRNAs. These siRNAs could again target a Pol V transcript and recruit DRM2 (Kanno et al. 2004, Chan et al. 2005, Kanno et al. 2005, Kanno et al. 2008, Voinnet 2008).

Although a lot of work has been done to elucidate the mechanisms of RdDM, many questions remain. In particular, how do Pol IV and Pol V target DNA for methylation? Do they recognize specific sequences and/or modifications? How is DNA methylation established, since the current model suggests that Pol IV and Pol V transcribe methylated DNA? How do the 10 structural subunits contribute to the sub-functionalization of Pol II, Pol IV and Pol V? Is there crosstalk between Pol II and Pol IV/Pol V to mediate and target RdDM to specific genes or promoters throughout development?

1.3.5 RdDM and the potential role of RNA Polymerase II

Although it is believed that Pol IV and Pol V have evolved to take over the role of Pol II in RdDM, recent work suggests that Pol II may retain some ancestral role. It was discovered that Pol II associated and produced nc-transcripts at well-defined RdDM target loci. Recruitment of Pol II to some of these loci requires DRD1, the chromatin-remodeler coupled with Pol V nc-transcription. Moreover, Pol II can interact with AGO4 through its WG/GW repeats in its C-terminal domain (Li et al. 2006, El-Shami et al.
The role of Pol II in RdDM was assessed further by analyzing RdDM in weak *nrpb2* alleles (the second largest subunit of Pol II). In these mutants, RdDM loci had less Pol IV, Pol V, and AGO4 association, suggesting that Pol II may recruit RdDM machinery to target loci. Decreased siRNA accumulation and silencing correlated with the reduction in Pol IV and Pol V association, respectively (Zheng et al., 2009). Another study utilized mutations in Mediator, a multisubunit complex known to recruit Pol II to protein coding genes, to analyze the role of mediator and Pol II in RdDM. Similar to *nrpb2*, mediator mutants reduced Pol II nc-transcript production and Pol V recruitment (Kim et al. 2011, Kim and Chen 2011). The link between Pol II and RdDM was strengthened further when RDM1 was found to colocalize with Pol II, AGO4, and DRM2. As described in the previous section, RDM prefers to bind methylated ssDNA and also works with Pol V (Gao et al. 2010). Taken together, this recent work suggests that Pol II plays a critical role in RdDM, but its exact function is not well understood.

### 1.3.6 Geminivirus chromatin

The structure of geminivirus chromatin has not been extensively studied, but what is known is consistent with the viral transcription program, and further suggests that viral nucleosome organization is similar to that of host chromatin. A circular DNA molecule 2.6 kb in size could accommodate 10-15 nucleosomes, depending on the length of linker DNA. Assuming an average of 200 bp associated with each, one would expect the viral genome to be complexed with a maximum of 13 nucleosomes. An early electron microscope study showed geminivirus nucleoprotein complexes with a "beads-on-a-string" appearance typical of open chromatin, with an average of 12 and a maximum of
13 per string (Abouzid et al. 1988). Later studies employing nuclease digestion and topoisomer analysis confirmed this general structure (Pilartz and Jeske 1992, Pilartz and Jeske 2003). In addition, nuclease hypersensitive sites were detected at regions corresponding to the IR and at least one of the downstream leftward promoters, consistent with the notion that these regions can exist in a more open and active conformation. Closer inspection of nuclease digestion patterns also provided evidence for populations of genomes with alternative nucleosome phasing, suggesting that nucleosome positions might not be fixed (Pilartz and Jeske 2003). The overall picture that emerges from these studies is one of dynamic viral chromosomes, possibly varying between a repressed state (13 nucleosomes) and active states with one or more relatively open sites corresponding to promoter regions (11-12 nucleosomes). Recent studies found that the CaLCuV IR is associated with both active (acetylated H3) and repressed (dimethylated H3K9) histone tail modifications (Raja et al. 2008), also suggesting that different populations of active and repressed viral genomes co-exist during viral infection. It will be interesting to determine whether and how viral proteins might influence chromatin structure. It also remains to be seen whether histone variants that could stabilize nucleosome arrangements characteristic of active or repressed states are incorporated into viral chromatin. To date, little work has been done to assess the histone modifications associated with the viral chromatin on both the IR and downstream promoter regions.

1.3.7 Geminivirus chromatin as a model to study the RdDM pathway

The geminivirus dsDNA RF associates with histones to form minichromosomes and histones can be modified by host enzymes similar to endogenous chromatin. Plants
deficient for methyl cycle enzymes, RdDM pathway components, as well as histone and DNA methyltransferases are hypersusceptible to geminivirus infection. Thus, the defense pathways that target endogenous chromatin and transposons also target geminiviruses (Henderson 2007, Raja et al. 2008, Buchmann et al. 2009, Lisch 2009). Moreover, plant recovery from BCTV L2- infection is associated with hypermethylation of the viral DNA and this requires the RdDM pathway component AGO4 (Raja et al. 2008). Thus, deficiencies in RdDM pathways produce clear viral phenotypes providing us with an extremely sensitive model system to study the plant RdDM pathway.

It was discussed in section 1.2.1 that the DRB protein that partners with DCL3 in the RdDM pathway remains to be determined. Due to the highly elaborated siRNA pathways in plants, functional redundancy between the DRBs, DCLs, RDRs, and AGOs has been observed and has blurred the lines between silencing pathways. Thus, sensitive assays must be used to break through the redundant roles of silencing proteins in order to identify the key RdDM pathway components. In Chapter 3, the novel geminivirus model system was used to show that DRB3 partners with DCL3 and AGO4 in the RdDM defense against geminiviruses. This validated the sensitivity of the geminivirus model system to study the RdDM pathway. In sections 1.3.4 it was established that it is unclear how Pol IV and Pol V are recruited and target DNA for de novo methylation and spread. Moreover, the exact role of Pol II in RdDM is undefined (described in section 1.3.5). Thus in Chapter 4 the geminivirus model system was employed to determine the roles of Pol IV and Pol V in the RdDM defense against geminiviruses.
**Fig. 1.1 Geminivirus genome and transcripts.** The diagrams depict the dsDNA replicative forms of a typical bipartite begomovirus (e.g. TGMV, CaLCuV), a monopartite begomovirus (e.g. TYLCV), a curtovirus (e.g. BCTV), and satellite DNA (DNA β). Red arrows indicate that the protein has been directly demonstrated to have silencing suppressor activity in at least one virus. Not all potential suppressors have as yet been tested for activity. There are currently two gene nomenclature systems in use. One refers to genes as leftward (L) or rightward (R) with respect to direction of transcription relative to the conventional map. The other refers to genes as complementary sense (C) or viral sense (V). Some viral protein have also been named according to core function, including replication initiator protein (Rep), transcriptional activity protein (TrAP), replication enhancer (REn), capsid protein (CP), movement protein (MP), and nuclear shuttle protein (NSP). Major transcripts (indicated by thin
black arrows) are shown using data from TGMV as a model for bipartite begomoviruses. Note that the bidirectional transcription units overlap at their 3' ends. In some cases, transcript names (AL-61, AL-1935, Al-1629) are provided (see text). The position of a conserved hairpin that marks the site of replication initiation is indicated by an asterisk within the intergenic region (IR). The common region (CR), a sequence of ~230 base pairs that is nearly identical between the A and B genome components of bipartite viruses is shown as a blue box.

Fig. 1.2 Geminivirus replicative cycle. Geminivirus virions enter cells and uncoat to release the positive-sense ssDNA in the nucleus. Initiation of the negative strand (red line) is not completely understood and results in the production of a covalently closed dsDNA replicative form (dsRF, a.k.a. RF 1). The dsRF is a platform for viral gene expression, which ultimately results in the initiation of rolling-circle replication (RCR).
RCR creates a new dsRF that can re-enter the replicative cycle, as well as ssDNA that can spread cell-to-cell or be packaged into virions.

Adapted from Bisaro, 1996

**Fig. 1.3 Schematic diagram of the geminivirus intergenic region.** Schematic diagram of the TGMV replication origin is depicted. The relative positions of Rep-binding sites, the stem loop and the invariant sequence (TAATATTAC), and the site where plus-strand synthesis initiates are shown. The iterative Rep binding sites (small orange boxes) play an integral role in origin recognition, specificity and transcription regulation (Fontes et al. 1994b, Sunter et al. 1993, Eagle et al. 1994). The intergenic region also contains sequence elements required for recruitment of Pol II transcription machinery, including TATAA boxes, a G-box family transcription factor binding site, and a TrAP response element (conserved late element)(Lazarowitz 1992, Argüello-Astorga et al. 1994). These divergent Pol II promoters guide initiation of Rep and CP transcription (indicated by the black arrows).
Fig. 1.4 Antiviral RNA silencing pathways in *Arabidopsis*. Cytoplasmic events (blue background) illustrated include the generation of primary viral siRNAs by DCL4 in collaboration with DRB4. DCL2 can redundantly function if DCL4 is inactivated. The sources of primary dsRNA include the products of viral RDRs (vRDR) for RNA viruses or RNA polymerase II (Pol II) for geminiviruses and other DNA viruses. The siRNAs enter RISC complexes that employ AGO1 as the primary slicer activity (indicated by scissors), with redundant activity provided by AGO7. RDR6-mediated secondary siRNA production, in conjunction with SGS3 and other proteins, is also shown. Transcripts normally possess a cap (black circle) and a polyA tail (An). Aberrant or sliced transcripts which lack these features may directly serve as substrates for dsRNA synthesis by RDR6 (or redundantly by RDR1, not shown). Cell-to-cell and systemic spread of silencing is presumably mediated by small RNAs and enhanced by amplification. Nuclear events (brown circle) depicted include Pol II mediated transcription of viral mRNAs from active geminivirus minichromosomes (nucleosomes illustrated as blue circles) and non-coding Pol IV transcription from...
methylated, repressed minichromosomes. Transcripts produced by either enzyme are subject to DCL3-mediated processing either directly (if they possess sufficient double-stranded character) or following conversion to dsRNA by RDR2. The 24 nt siRNA products of DCL3 guide docking of AGO4-containing RISC onto nascent transcripts generated by POL V. Subsequent methylation of viral DNA and associated histones involves recruiting cytosine (DRM1, DRM2, CMT3, MET1) and histone (e.g. SUVH4) methyltransferases, and requires chromatin remodelers, including DRM1 and DDM1. Pol V transcripts may also be sliced by AGO4 (scissors) to act as additional RDR2 substrates.Suppressor proteins encoded by RNA viruses (red stop signs) or DNA viruses (orange stop signs) interfere with silencing pathways either by binding small RNAs or interacting with pathway components, as indicated. Note that the geminivirus AL2 and L2 proteins suppress silencing in a transcription-independent manner by inhibiting the methyl cycle that generates S-adenosyl methionine (SAM). Methyl cycle inhibition is accomplished by inactivation of adenosine kinase (ADK) (see Fig. 1.5). A transcription-dependent mechanism unique to AL2 that presumably involves transcriptional up-regulation of a cellular gene (WEL1) is also illustrated.

Yang et al. 2011

**Fig. 1.5 The methyl cycle and its inhibition by geminivirus proteins.** S-adenosyl methionine (SAM) is the methyl donor for most transmethylation reactions. The product, S-adenosyl homocysteine (SAH), inhibits transmethylation by competing with SAM for methyltransferases (MTases). SAH is converted to homocysteine (Hcy) and adenosine by S-adenosyl homocysteine hydrolase (SAHH). Phosphorylation of adenosine by adenosine kinase (ADK) is critical because the SAHH-catalyzed reaction is reversible and the equilibrium lies in the direction of SAH synthesis. By removing adenosine, ADK
promotes flux through the cycle and minimizes competitive inhibition of methyltransferase reactions by SAH. Thus, ADK inactivation by geminivirus AC2/AL2 and C2/L2 proteins globally interferes with methylation. TYLCCNV AC2/AL2 may be deficient for ADK inhibition. The C2/L2 protein has also been shown to stabilize SAM decarboxylase (SAMDC), which causes decarboxylated SAM (dcSAM) levels to rise. dcSAM is a competitive inhibitor of SAM. Whether AC2/AL2 proteins also promote increases in dcSAM has yet to be tested. The betasatellite encoded βC1 protein directly antagonizes the methyl cycle by inhibiting SAHH. Note that this diagram lists only DNA and histones as methyltransferase substrates, although any type of transmethylation reaction may require SAM as a cofactor. THF: tetrahydrofolate, PPi: pyrophosphate; Pi: inorganic phosphate.

**Fig. 1.6 Nucleosome structure and histone tail modifications.** Nucleosomes are formed by approximately 147 bp of DNA wrapped around an octamer of H3, H4, H2A, and H2B canonical histones. Various post-translational modifications (PTMs) are added to the unstructured N- and C-terminal tails of the histones that protrude from this protein-DNA structure. Greater than 60 amino acid residues have been shown to be modified and these regulate transcription and replication in a variety of ways. As shown in this diagram, lysine (K) residues can be acetylated (Ac), ubiquitinylated (Ub), or methylated (Me). Arginine (R) residues can be methylated, and serine (S) and threonine (T) can be phosphorylated (Ph). Additional complexity is added since arginine and lysine can be mono-, di-, and trimethylated.
Fig. 1.7 Subunit composition of RNA Polymerases II, IV and V. RNA polymerase II, IV and V are composed of 12 core subunits that can be unique or shared among each of these polymerases. Although Pol IV and Pol V evolved from Pol II, each of these enzymes contains a unique largest subunit. Pol IV and Pol V share a second largest subunit and together these 2 subunits confer catalytic activity. Pol II, IV and V share various other small subunits, but the functional consequences of subunit diversity are unknown.
Fig. 1.8 RNA directed DNA methylation of geminivirus genomes. Viral or cellular target sequences are transcribed by Pol IV and possibly by Pol II. The ssRNA nontranscript is converted to dsRNA by RDR2. DCL3 processes 24 nt siRNA from the dsRNA, which is loaded into an AGO4-RISC complex with the aid of DRB3. AGO4-RISC recruits the cytosine methyltransferases DRM2. CMT3 and MET1 maintain cytosine methylation at CNG and CG sites, respectively. KYP2 carries out H3K9 methylation. Chromatin remodelers involved include CLSY1, DRD1, and DDM1. All proteins represented in color have been shown to play a role in the RdDM defense against geminiviruses and loss of these proteins results in hypersusceptibility to geminivirus infection. White ovals designate proteins that have yet to be tested.
CHAPTER 2

Analysis of geminivirus AL2 protein identifies new suppressor activities and developmental changes in the role of ADK in the *N. benthamiana* methyl cycle.

The following chapter will be submitted for publication as follows: **Jamie N. Jackel**, R. Cody Buchmann*, Udit Singhal, and David M. Bisaro (* These authors contributed equally to this work). Analysis of geminivirus AL2 protein identifies new suppressor activities and developmental changes in the role of ADK in the *N. benthamiana* methyl cycle. Figures 2.1 and 2.2 were done by R. Cody Buchmann and Figures 2.3 and 2.4 I performed in collaboration with Udit Singhal.

2.1 Introduction

Viruses belonging to the *Geminiviridae* package small, circular single-stranded DNA (ssDNA) genomes in unique double icosahedral particles. The ssDNA genome, which may be monopartite or bipartite, is replicated by rolling circle and recombination mediated mechanisms using double-stranded DNA (dsDNA) intermediates that associate with cellular histones to form minichromosomes. Replication and transcription from viral templates is accomplished by host polymerases and accessory proteins. Viral gene products are typically multifunctional proteins that collectively direct host machinery to viral templates, initiate specific steps in replication and/or transcription, enable virus spread within and between hosts, provide a cellular environment favorable to replication, and suppress host defenses (Hanley-Bowdoin et al. 2004, Rojas et al. 2005, Bisaro 2006, Jeske 2009).
RNA silencing is a general term that refers to a set of mechanistically related and evolutionarily conserved processes, including post-transcriptional gene silencing (PTGS) and transcriptional gene silencing (TGS). In plants, PTGS typically leads to translational inhibition or cytoplasmic degradation of mRNAs. By contrast, TGS is a nuclear phenomenon that results in RNA-directed DNA methylation (RdDM) and epigenetic silencing. An antiviral role for both PTGS and TGS is well-established, and their importance as a defense is clear from the fact that virtually all plant viruses encode proteins that act as silencing suppressors (Ding and Voinnet 2004, Ruiz-Ferrer and Voinnet 2009, Raja et al. 2010, Wu et al. 2010, Burgyan and Halveda 2011, Ding and Lu 2011). These viral counter-defensive proteins are essential pathogenicity factors that act by a variety of mechanisms to block different aspects of silencing. A remarkable feature of silencing that greatly enhances its defensive value is its ability to spread cell-to-cell and systemically throughout the plant (Brosnan and Voinnet 2011). However, the molecular basis for systemic spread, and the mechanisms by which some viral suppressors are able to prevent it, are not completely understood. Likewise, some viral suppressors are known to reverse established silencing, suggesting they target silencing maintenance pathways that are not well characterized.

An example of a geminivirus silencing suppressor is AL2 encoded by members of the genus *Begomovirus*, which includes *Tomato golden mosaic virus* (TGMV) and *Cabbage leaf curl virus* (CaLCuV). The 15-kDa AL2 (also known as AC2, C2, or TrAP, transcriptional activator protein) was originally characterized as a transcription factor required for the expression of viral late genes (Sunter and Bisaro 1992, Sunter and Bisaro 2003). The related L2 protein (also known as C2) found in the *Curtovirus* genus,
Beet curly top virus (BCTV), is not required for late viral gene expression and lacks transcription activation activity (Hormuzdi and Bisaro 1995, Yang et al. 2007). However, AL2 and L2 share pathogenicity functions (Sunter et al. 2001, Hao et al. 2003, Wang et al. 2003). AL2 has been shown to suppress PTGS by multiple mechanisms. One involves AL2-mediated transactivation of host genes that may act as endogenous negative regulators of RNA silencing (transcription-dependent suppression) (Trinks et al. 2005). Another is based on the ability of both AL2 and L2 to interact with and inactivate adenosine kinase (ADK), a methyl cycle-associated enzyme required for the efficient production of S-adenosyl methionine (SAM), an essential methyltransferase cofactor (Moffatt et al. 2002, Wang et al. 2003). We found that AL2 and L2 can suppress local silencing (PTGS) by a mechanism that involves ADK inactivation (transcription-independent suppression) (Wang et al. 2005, Yang et al. 2007).

More recently, we demonstrated that small RNA-directed methylation leading to TGS acts as an epigenetic defense against geminiviruses, and that AL2 can suppress methylation and TGS by both transcription-dependent and -independent means, while L2 is limited to the latter mechanism (Burgyan and Halveda 2011, Raja et al. 2008). However, the roles of the each mechanism in suppressing different aspects of PTGS and TGS remained to be described. The goal of the present study was to determine whether, and in general terms how, AL2 and L2 inhibit systemic spread of silencing and cause the reversal of PTGS and TGS. In addition to furthering our knowledge of viral pathogenesis, an analysis of AL2 and L2 suppression activities is expected to lead to a greater understanding of host silencing and methylation pathways.
2.2 Methods

2.2.1 Systemic spread suppression assay

The assay to analyze suppression of systemic silencing has been described (Silhavy et al. 2002). Briefly, transgenic *Nicotiana benthamiana* plants containing an active 35S-GFP transgene, line 16c (Ruiz et al. 1998), were agroinfiltrated with a mix of cultures containing binary plasmids to express double-stranded GFP RNA (dsGFP) and a test protein or dsRNA construct. β-glucuronidase (GUS) was a negative control and p19 was a positive control. Constructs expressing the controls and geminivirus proteins AL2, AL21-114, AL2-C33A, and L2 have been described (Wang et al. 2005, Yang et al. 2007). Constructs expressing dsRNA to silence ADK (dsADK) and S-adenosyl homocysteine hydrolase (dsSAHH) were prepared from previously described cDNAs (Buchmann et al. 2009). Plants were screened for GFP expression with a hand-held long wave UV lamp (Blak-Ray Model B 100 YP) and photographed with a Nikon D40 equipped with 52 mm UV lens and a yellow filter.

2.2.2 Vegetative and reproductive plants

Seeds of *N. benthamiana* line 16c or 16-TGS were sown in 6 cm pots. For PTGS assays, after 4-5 weeks post-dsGFP infiltration, plants were screened under UV light for GFP expression. Plants with GFP expression were removed from the population. Then, plants were repotted in new 6 cm pots for reproductive growth or in 16.5 cm pots for vegetative growth. For TGS assays, 2 week old seedlings were screened under UV light.
for GFP expression. Again, plants expressing GFP were removed from the population and repotted in either 6 or 16.5 cm pots. Vegetative plants were shorter with larger leaves, had more lateral growth, and did not have flowers. Reproductive plants were tall and flowered.

2.2.3 Developmental PTGS reversal assay

16c *N. benthamiana* plants were grown for 2 weeks then agroinfiltrated with dsGFP to trigger systemic silencing. 4-5 weeks after planting, they either were transferred to larger pots to keep them at vegetative stage of growth or allowed to bolt, and checked from 6-8 weeks for complete silencing (indicated by the entire plant showing red under UV light). The plants were then infiltrated with agrobacterium containing a recombinant PVX cDNA vector. The PVX either was an empty vector, contained HC-Pro (a protein known to reverse silencing), AL2, the mutant AL2\textsubscript{1,14}, or L2 (Buchmann et al. 2009). PTGS reversal was analyzed using water-mounted tissue sections with a laser scanning confocal microscope.

2.2.4 Developmental TGS reversal assay

TGS reversal assay was performed as described previously (Buchmann et al. 2009). 16-TGS *N. benthamiana* seedlings were split into vegetative and reproductive populations. At three weeks of age, plants were rescreened for silencing of the GFP transgene under UV light. Only plants with no GFP expression were infected with Agrobacterium containing geminiviruses, PVX, or TRV vectors. As a negative control either BCTV \textit{L2'}, PVX empty vector or TRV empty vector were used. Geminiviruses
and TRV/PVX recombinant vectors were used previously to show TGS reversal in vegetative plants, thus serving as a positive control (Buchmann et al. 2009). TGS reversal was analyzed under UV light and photographed as described previously.

2.2.5 RNA extraction and Northern blot analysis

RNA extraction and Northern blots were performed as described previously (Buchmann et al. 2009). Trizol (Invitrogen) was used to isolate RNA. Northern blots were run with 4 µg of RNA and stained with EtBr to visualize the 28S rRNA as a loading control. GFP blots were probed with an antisense GFP riboprobe in vitro synthesized with [α-32P] UTP (3000 Ci/mmol, Perkin-Elmer) using T7 MAXIscript kit (Ambion). Blots analyzed for TRV or PVX viral titers were probed with a mix of antisense 5' labeled 20 µM oligonucleotides using [γ-32P] ATP (3000 Ci/mmol, Perkin-Elmer) and T4 polynucleotide kinase (Fermentas).

2.2.6 DNA isolation and Southern blot analysis

DNA extraction and Southern blots techniques were described previously (Raja et al. 2008). In short, DNA was extracted from individual plant samples using DNeasy columns (Qiagen). 1 µg of DNA was restricted over-night with either Scal for BCTV or XcmI for CaLCuV and 1% gel electrophoresis was performed. Membranes were cross-linked and probed with antisense oligonucleotides for either 18S rDNA for a loading control or geminivirus DNA. A mixture of five antisense oligonucleotides were 5'
labeled using [$\gamma$-$^{32}$P] ATP (3000 Ci/mmol, Perkin-Elmer) and T4 polynucleotide kinase (Fermentas).

2.2.7 Oligonucleotides

Oligonucleotides used for probing Northern and Southern blots are listed in Fig 2.7.

2.2.8 ADK activity assay

ADK activity was determined as described previously (Wang et al. 2003). Crude protein was extracted from individual mock, geminivirus, or TRV infected vegetative and reproductive plants. Protein extracts (250 ng) were incubated with 15 µM adenosine and [$\gamma$-$^{32}$P] ATP (3000 Ci/mmol, Perkin-Elmer) at 30°C for 20 minutes. The reaction was stopped with 2 µL 500 mM MgCl₂ and the products were fractionated using thin-layer chromatography plates (Sigma) in 1 M glacial acetic acid. Radiolabeled AMP was quantitated with phosphorimaging.

2.2.9 Methylation sensitive extension assay

The extension assay was conducted as described previously (Pogribny et al. 1999, Boyko et al. 2007, Buchmann et al. 2009). *N. benthamiana* DNA was extracted from mock or viral infected vegetative and reproductive plants. 1 µg of the DNA was restricted over-night with 10-fold excess of MspI (New England Biolabs). Single-nucleotide extensions were performed with 1X PCR Buffer, 1.0 mM MgCl₂, 0.25 units of *Taq* polymerase (New England Biolabs), and 0.5 µL [$\alpha$-$^{32}$P] dCTP (800 Ci/mmol,
Perkin-Elmer) in a 30 µL reaction for 1 hour at 56°C. 12.5 µL was loaded onto two DE-81 (Whatman) filter paper discs and washed as previously described. Radioactivity was determined using a scintillation counter. Samples were normalized to the respective mock control and students t test was performed.

2.3 Results

2.3.1 AL2, but not L2, prevents spread of systemic silencing in a transcription-dependent manner.

Using agroinfiltration assays in Nicotiana benthamiana (Johansen and Carrington 2001), we showed that the TGMV AL2 and AL2_{1-100} proteins (the latter lacking the transcription activation domain), BCTV L2 protein, and double-stranded RNA corresponding to ADK (dsADK) are able to suppress local silencing directed against the coding region of a transiently expressed green fluorescent protein (GFP) mRNA (Wang et al 2005). Thus the transcription-independent mechanism shared by AL2 and L2, which depends on methyl cycle inhibition due to ADK inactivation, is sufficient to suppress local PTGS.

A related silencing assay using N. benthamiana line 16c plants, which contain a GFP transgene driven by the constitutive 35S promoter (35S-GFP) (Ruiz et al. 1998), was adopted to determine whether the viral proteins, and methyl cycle inhibition, could also interfere with systemic spread of PTGS. Line 16c plants were co-agroinfiltrated with a construct to express inverted repeat RNA corresponding to the GFP coding region (dsGFP) to induce systemic PTGS, along with test or control constructs. The p19 protein...
of *Cymbidium ringspot virus* is a strong suppressor of systemic silencing and served as a positive control (24). β-glucuronidase (GUS) was a negative control. Test constructs expressed L2, AL2, AL2_{1-114} or AL2-C33A. AL2_{1-114} lacks the minimal transcription activation domain (Hartitz et al. 1999). AL2-C33A is defective in self-interaction and has reduced transcription activation activity. However, neither the C33A substitution nor the absence of the activation domain affects the AL2:ADK interaction (Yang et al. 2007). Constructs expressing inverted repeat RNAs corresponding to *N. benthamiana* ADK (dsADK) and S-adenosyl homocysteine hydrolase (dsSAHH) were also tested. ADK promotes flux through the methyl cycle, while SAHH is an essential methyl cycle enzyme (Moffatt et al. 2002).

Following co-infiltration of lower leaves of 16c plants with dsGFP and GUS (negative control), systemic silencing of the GFP transgene was visible under UV light after ~10 days on the crown leaves as dark red areas against the salmon color of GFP-expressing tissue (Fig. 4.1). Dark red areas, caused by chlorophyll auto-fluorescence, were apparent in all but ~20% of the plants tested. Remarkably, despite its relatively weak activity in local assays (Wang et al. 2005), AL2 proved nearly as effective as p19 at inhibiting systemic silencing, with an average of ~90% non-silenced plants over three independent experiments. This is consistent with a previous study, which used a different assay to show that AL2 (AC2) protein from *Mungbean yellow mosaic virus* interferes with silencing spread (Trinks et al. 2005). By contrast, AL2_{1-114} and L2, which lack transcription activation activity, did not prevent systemic silencing and were essentially equivalent to the GUS negative control. Both the dsADK and dsSAHH constructs also proved ineffective (Fig. 4.1). These observations indicate that the transcription-
independent suppression mechanism involving methyl cycle inhibition is insufficient to prevent the initiation, subsequent spread, or perception of the systemic silencing signal. Supporting the necessity for transcription activation, AL2-C33A, with about ~30% the activation activity of wild type AL2, displayed intermediate suppression in this assay (Fig. 4.1). We concluded that AL2, but not L2, inhibits systemic spread of silencing by a transcription-dependent mechanism.

2.3.2 Plant developmental stage can determine whether transcription-dependent or -independent mechanisms can reverse PTGS.

Only a relatively few silencing suppressors, including HC-Pro and AL2, have been shown to reverse established PTGS, which suggests they are able to interfere with a maintenance step (Anandalakshmi et al. 1998, Brigneti et al. 1998, Kasschau et al. 1998, Voinnet et al. 1999, Llave et al. 2000).

To examine AL2 and L2 PTGS suppression activities, these were studied in plants exhibiting systemic GFP silencing both before and after the vegetative-to-reproductive transition. To do this, N. benthamiana line 16c plants were grown in small (6 cm) pots and lower leaves were infiltrated two weeks post-germination with the dsGFP construct to trigger systemic PTGS. (Plants that later failed to show systemic silencing were removed from the study.) Two to three weeks following administration of the silencing trigger, the plants were divided into two groups. "Reproductive" plants were transferred to fresh small pots (6 cm), and quickly flowered. "Vegetative" plants were transplanted into larger (16.5 cm) pots, which allowed them to continue vegetative growth. Thus the
GFP-silenced test plants differed with respect to the vegetative-to-reproductive transition. Silenced plants were then inoculated with recombinant *Potato virus X* (PVX) vectors expressing AL2, AL2\(_{1-114}\), and L2 (Buchmann et al. 2009). PVX (empty vector) served as negative control, and PVX::HC-Pro as a positive control. Four weeks post-inoculation, plants were observed under UV light and samples collected for GFP mRNA analysis. Samples consisted of symptomatic leaf and stem tissue pooled from 3 to 4 plants.

GFP fluorescence indicating silencing reversal was not reliably detected by visual inspection of whole plants, even with the PVX::HC-Pro positive control. Thus we examined water-mounted tissue segments using a confocal microscope, and under these conditions yellow-green GFP-expressing cells could easily be seen against a background of red cells in which the GFP transgene remained silenced (Fig. 4.2A). Curiously, the PVX vector alone reproducibly caused silencing reversal specific to stomatal guard cells in both reproductive and vegetative plants, suggesting that PVX affects silencing maintenance in this but no other cell type examined. A total of three experiments were carried out with vegetative and reproductive plants, and one additional experiment was performed only with reproductive plants. Each experiment included four plants per treatment. Leaves showing obvious PVX symptoms and 3-4 leaves below the crown were harvested and one or two tissue segments were taken from each. A segment was considered positive for reversal if a field of view from the section showed GFP expression in cells other than guard cells.

Infection of vegetative plants with PVX::HC-Pro, PVX::AL2, PVX::AL2\(_{1-114}\), and PVX::L2 resulted in reversal of PTGS directed against GFP (Fig. 4.2A). PVX::HC-Pro
elicited yellow-green GFP fluorescence in isolated patches of cells, whereas PVX vectors expressing the geminivirus proteins induced expression in a more evenly dispersed pattern. Silencing reversal was apparent in >75% to 100% of the tissue segments examined in all experiments with all recombinant viruses, with the exception of one experiment where PVX::L2 was positive in ~50% of the segments. Northern blot analysis of GFP mRNA levels confirmed these observations (Fig. 4.2B).

Surprisingly, in reproductive plants, only PVX::HC-Pro and PVX::AL2 caused significant GFP expression (Fig. 4.2A). GFP expression was observed in >75% to 100% of the tissue segments examined in each of four experiments. GFP expression outside of guard cells was never observed with PVX::AL2-1-114, and in only one of four experiments with PVX::L2-inoculated plants, and in this case fewer than 50% of sections were positive. These observations were again confirmed by Northern blot analysis (Fig. 4.2B). We concluded that AL2 is capable of reversing GFP-directed silencing in leaves from both vegetative and reproductive plants, whereas L2 and transcriptionally inactive AL2 can reverse silencing only vegetative leaves. This indicates that reversal in vegetative *N. benthamiana* plants can occur by the transcription-independent mechanism, while suppression in reproductive tissue is transcription-dependent. This in turn suggests that different mechanisms of PTGS maintenance are active during different stages of development.
2.3.3 Plant developmental stage can determine whether geminiviruses can reverse TGS.

In a previous study, we created an *N. benthamiana* line (16-TGS) containing a transcriptionally silenced 35S-GFP transgene and used it to show that CaLCuV and BCTV, and viral AL2 and L2 proteins, can reverse established TGS (Buchmann et al. 2009). These assays were carried out using vegetative plants, and thus we repeated the experiments to compare suppression activities in reproductive plants. BCTV *L2*\(^{-}\) mutant virus was used, which contains a stop codon that truncates the protein after 72 (of 173) amino acids (*L2*-2), was used as a negative control (Hormuzdi and Bisaro 1995). BCTV *L2*\(^{-}\) replicates and generates systemic symptoms similar to wild type BCTV. AL2-deficient begomoviruses (including CaLCuV and TGMV) are not systemically infectious because AL2 is required to activate the expression of the BR1 movement protein (Sunter and Bisaro 1992).

For these experiments, two-week old seedlings of *N. benthamiana* line 16-TGS were screened for GFP silencing under UV light, and only plants showing uniform red chlorophyll autofluorescence due to heritable TGS were used. Seedlings were then divided to generate vegetative and reproductive populations as described above. Three independent experiments were conducted with 8 vegetative and reproductive plants each. At three weeks of age, plants were infected with CaLCuV, BCTV, or BCTV *L2*\(^{-}\). A hand-held UV light was used to screen for TGS reversal three-weeks post-infection, and plants were considered positive when GFP expression was observed in symptomatic tissue. Tissue was harvested for RNA, DNA and ADK activity analysis, with samples consisting of symptomatic leaf and stem tissue pooled from 3 to 4 plants.
Consistent with previous results, >80% of vegetative plants infected with wild type BCTV and CaLCuV displayed GFP expression indicative of TGS reversal, whereas less than 20% of the BCTV L2− infected plants exhibited GFP expression (Fig. 4.3A and 4.3B) (Buchmann et al. 2009, Raja et al. 2010). Although reproductive plants infected with wild type BCTV showed disease symptoms similar to the vegetative plants, fewer than 20% of them showed evidence of GFP expression, similar to BCTV L2−. By contrast, comparable GFP expression was observed in leaves of vegetative and reproductive plants infected with CaLCuV, indicating that this virus is able to reverse TGS at both developmental stages (Fig. 4.3A and 4.3B). Similar results were obtained in two experiments with TGMV (data not shown).

To rule out the possibility that differences in virus replication might account for developmentally sensitive TGS reversal observed with BCTV, viral DNA levels were measured. DNA was isolated from infected vegetative and reproductive tissue, and viral dsDNA was linearized with Scal (BCTV) or XcmI (CaLCuV). Southern blots were probed with oligonucleotides specific for 18S rDNA (loading control) and the BCTV or CaLCuV intergenic regions. Not surprisingly, viral DNA levels were lowest in plants infected with BCTV L2− (Fig. 4.3C). However, for each virus, DNA levels in vegetative and reproductive tissues were similar, and thus the inability of BCTV to suppress silencing in reproductive tissue cannot be attributed to reduced replication.

We previously showed that ADK inhibition in infected plants is L2-dependent (Wang et al. 2003), and considered the possibility that L2 expression levels might be reduced in reproductive plants infected with wild type BCTV. To indirectly test this, ADK activity levels were measured in crude protein extracts obtained from vegetative and reproductive
tissues infected with BCTV or BCTV \textit{L2}'. The results of these experiments confirmed that ADK inhibition is L2-dependent, and that showed that wild type BCTV is equally capable of suppressing ADK activity in vegetative and reproductive tissue (Fig. 4.3D). Thus virus replication and L2 expression levels appear to be similar in both tissues. Taken together, these experiments indicate that, similar to PTGS, distinct mechanisms of TGS maintenance may be active during different stages of development.

2.3.4 A novel mechanism of AL2-mediated TGS reversal is employed in reproductive plants.

The experiments described in the previous section indicate that begomoviruses and curtoviruses differ in their ability to suppress TGS in reproductive tissue. To confirm this reflects the activities of their respective AL2 and L2 proteins, these were individually expressed from PVX vectors in line 16-TGS plants. We previously used this system to show that over-expression of AL2, AL2\textsubscript{1-114}, AL2 C33A, or L2 reverses TGS in vegetative plants (Buchmann et al. 2009), and here extended the experiments to include reproductive plants. Experiments were performed in triplicate with eight vegetative and reproductive plants in each experiment.

In support of results obtained with the geminiviruses, PVX::AL2 reversed TGS in both vegetative and reproductive tissue in >80% of the plants, while PVX::L2 TGS reversal was limited to vegetative plants (Fig. 4.4A and 4.4B). Surprisingly, the transcription activation-deficient AL2\textsubscript{1-114} and activation-impaired AL2-C33A also reversed silencing in both vegetative and reproductive plants. However, GFP expression
induced AL2-C33A was somewhat weaker than wild type AL2 and AL2<sub>1-114</sub>, especially in reproductive tissue. This suggests that AL2 dimerization or the non-canonical zinc finger motif, both of which are affected by the C33A substitution, might have some role in TGS reversal independent of transcription activation. PVX vector accumulation was similar in both types of tissue regardless of the geminivirus protein expressed, suggesting comparable protein expression levels in vegetative and reproductive plants (Fig. 4.4C). We concluded that TGS reversal can occur in a transcription-independent manner in vegetative plants by a mechanism that is shared by AL2 and L2. However, only AL2 can reverse TGS in reproductive <i>N. benthamiana</i> plants, and in this case by another transcription-independent mechanism that is not shared by L2. This analysis thus uncovered a novel, third mechanism for AL2-mediated silencing suppression that does not correlate with ADK inhibition or transcription activation.

2.3.5 TGS is independent of ADK activity in reproductive plants.

We showed, using a <i>Tobacco rattle virus</i> (TRV)-based VIGS vector, that knockdown of methyl cycle enzymes ADK and SAHH phenocopies geminivirus TGS reversal of a GFP transgene in vegetative plants (Buchamnn et al. 2009). However, studies described here indicate that BCTV and L2 cannot reverse TGS in reproductive tissue, yet the virus continues to efficiently inhibit ADK activity even after <i>N. benthamiana</i> plants have begun to flower (Fig. 4.3D). Therefore, TGS in reproductive tissue may be independent of either the methyl cycle or ADK. To explore the possibility that methyl
cycle enzyme requirements change through development, VIGS knock-down was repeated in reproductive plants.

Experiments using line 16-TGS were carried out as described above, except plants were infected with TRV empty vector as a negative control, or TRV containing fragments of ADK, SAHH coding regions. METHYLTRANSFERASE 1 (MET1) was a positive control. A handheld UV light was used to identify plants with GFP expression, and this tissue was harvested for RNA and ADK activity analysis. At least eight plants from both the vegetative and reproductive populations were infected and examined for GFP expression in three independent experiments.

Approximately 10% of the vegetative and reproductive plants infected with TRV empty vector showed GFP expression. On the other hand, TRV::ADK, TRV::SAHH, and TRV::MET1 infections resulted in TGS reversal in nearly all of the vegetative plants. By contrast, only TRV::SAHH and TRV::MET1 continued to show strong yellow-green fluorescence at the apex of approximately 100% of the reproductive plants. TRV::ADK infected reproductive plants displayed only background levels of GFP expression, indicating that ADK knock-down is no longer sufficient to reverse TGS in reproductive plants (Fig. 4.5A). Analysis of GFP mRNA accumulation by Northern blot hybridization supported visual observations (Fig. 4.5B). Additionally, oligonucleotide probes complimentary to the TRV CP showed similar vector replication levels (Ratcliff et al. 2001) (Fig. 4.5C).

Finally, to ensure that the loss of TGS reversal in reproductive tissue was not due to insufficient ADK knock-down, an ADK activity assay was employed. ADK assays were conducted with the total protein extracts obtained from mock inoculated, TRV, and
TRV::ADK infected vegetative and reproductive plants, and activity was normalized to the mock samples. We found that ADK activity was similar in both vegetative and reproductive tissue in mock extracts, and that activity was stimulated by TRV infection (Fig. 4.5D). A similar increase in ADK activity was previously observed following infection with PVX and Cucumber mosaic virus (CMV), and to a lesser extent with BCTV L2' (Fig. 4.3D) (Wang et al. 2003). Nevertheless, TRV::ADK reduced ADK activity to between 30% and 60% of levels seen in mock inoculated reproductive and vegetative tissues, respectively (Fig. 4.5D). These results allow us to conclude that TGS can be restored in N. benthamiana reproductive tissue despite low ADK activity levels, suggesting that the methyl cycle becomes independent of ADK.

2.3.6 ADK inhibition inhibits global cytosine methylation in vegetative, but not reproductive plants.

As a more direct measurement of methyl cycle activity, we used a methylation-sensitive cytosine extension assay to interrogate the status of N. benthamiana DNA (Buchmann et al. 2009, Yang et al. 2011). These experiments rely on the inability of MspI to cleave its target site (C/CGG) when the external cytosine is methylated. Following digestion of total genomic DNA, a single-nucleotide extension assay is performed using $^{32}$P-dCTP and Taq DNA polymerase. Under these conditions, nucleotide incorporation depends on the number of sites cleaved, which in turn is negatively correlated with methylation at CNG sites.
DNA was isolated from mock inoculated and BCTV, BCTV \(L^2\), and CaLCuV infected vegetative and reproductive plants. DNA was also obtained from plants inoculated with TRV, TRV::ADK, and TRV::SAHH. At least three individual plants were sampled, and duplicate extension reactions were performed with each sample.

In vegetative tissue, BCTV \(L^2\) infection had no significant impact on cellular DNA methylation levels when compared to mock inoculated plants, whereas increased incorporation was observed in BCTV (>1.5-fold) and CaLCuV (>2-fold) infected plants, as expected. However, in keeping with its ability to suppress TGS, only CaLCuV caused increased incorporation in reproductive tissue (>3-fold) (Fig. 4.6A).

Cytosine extension assays also showed no apparent difference between mock samples and samples from plants inoculated with the TRV vector, while increases in incorporation were noted in vegetative plants infected with TRV::ADK (>1.2-fold) and TRV::SAHH (>1.5-fold). However, consistent with its ability to suppress TGS, only TRV::SAHH treatment was able to increase incorporation in reproductive plants (>1.5-fold) (Fig. 4.6B). These results provide further evidence that suppression mechanisms involving ADK inhibition are effective in vegetative, but not reproductive, \(N.\) benthamiana plants, and further indicate that the role of ADK in sustaining the methyl cycle is developmental stage-dependent.

2.4 Discussion

The studies reported here extend our knowledge of AL2 and L2 silencing suppression and shed new light on the role of ADK in the methyl cycle. First, we found that
inhibiting the methyl cycle, either by ADK or SAHH knock-down, had no impact on systemic spread of silencing and that AL2, but not L2, can prevent silencing spread. At this time we do not know whether AL2 inhibits the generation, spread, or perception of the signal, although we do know that suppression requires the AL2 transcription activation domain. We also demonstrated that while AL2 and L2 can reverse established PTGS in vegetative plants, only AL2 is able to do this in reproductive plants after the onset of flowering. The ability of AL2 to reverse PTGS in reproductive plants also requires the transcription activation domain. Thus, these studies confirmed two transcription-dependent suppressor functions of geminivirus AL2. Both of these activities are likely due to transcriptional activation of genes encoding endogenous silencing regulatory proteins (Trinks et al. 2005).

Subsequently, vegetative and reproductive plants were infected with geminiviruses and PVX to assess the transcription-dependent suppressor functions of AL2 in TGS reversal. Once more AL2, but not L2, was able to reverse TGS in reproductive plants. These analyses were repeated and verified by allowing vegetative plants to bolt and flower and the ability of L2 to reverse TGS was always lost at the vegetative to reproductive transition. However, transcription activation AL2 mutants, AL2-1-114 and AL2-C33A, continued to reverse established TGS. Therefore, this activity appears to be unrelated to AL2s ability to reverse PTGS, block the spread of silencing, and inhibit ADK activity. These data identify a third uncharacterized suppressor function of AL2, beyond transcription activation and ADK inhibition, and is not related to the transcription-dependent TGS reversal that this study set out to identify.
Analyzing L2 and AL2 suppressor functions not only revealed previously unidentified activities, but also uncovered interesting plant physiology. Due to the fact that AL2 and L2 are known to reverse TGS through inhibition of ADK, it was important to evaluate the role of ADK in vegetative and reproductive plants. Analysis of TGS reversal and ADK activity in BCTV and TRV::ADK infected plants, indicate that although ADK activity is significantly reduced, TGS is no longer reversed in reproductive plants. The methylation sensitive nucleotide fill-in assay results found that BCTV and TRV::ADK infections cause reductions in the global DNA methylation in vegetative plants. However, BCTV and TRV::ADK infected reproductive plants have similar levels of methylation to BCTV L2- and TRV infections, respectively. Thus, the methyl cycle is able to function independent of ADK in reproductive N. benthamiana, suggesting that alternative enzymes salvage adenosine at different stages of development.

If the methyl cycle in reproductive tissue becomes independent of ADK activity, then adenosine must be salvaged through an alternative pathway in order for the methyl cycle to produce SAM properly. Although three enzymes can potentially contribute to adenosine salvage, only ADK and adenosine nucleosidases (AN) have been detected in plant extracts (reviewed in Schoor and Moffatt 2004). ADK is highly expressed, active normally throughout plant development, and a complete loss of ADK activity is lethal in Arabidopsis (Shomberg and Stephan 1997, Moffatt 2000, Moffatt et al. 2002). Therefore, ADK has been considered the predominant adenosine salvage pathway. Interestingly, adenosine can be converted to AMP through the sequential activities of AN and adenine phosphoribosyltransferase (APT). Work in Avicennia marina has shown that AN activity is higher than ADK activity, indicating that there are two functional routes
for adenosine salvage in this species (Suzki et al. 2003). Thus it is reasonable to suspect that AN could also play a role in adenosine salvage in reproductive *N. benthamiana*.

Studies of the *Arabidopsis* genome have identified five nucleoside hyrolase (NSH), named NSH 1 through 5. Nucleoside hydrolases can be split into three categories based on their preference for inosine-uridine, inosine-adenosine-guanosine, or non-specific activity (Versees et al. 2001). Depending on the substrate preference of these five genes, they act as potential candidates for AN activity. To date, little work has been done on these five genes, but NSH1 is fairly well characterized (Jung et al. 2009). More recent work has furthered our understanding of NSH1, NSH2 and NSH3 and has shown that the genes have various expression patterns during development and work to salvage both intracellular and extracellular purines (Jung et al. 2011). The results show that both NSH1 and NSH3 have the ability to salvage adenosine, thus providing two candidates to compensate for ADK activity in reproductive plants.

The studies presented here have opened the door for several pathways to explore the interesting plant physiology and viral suppressor functions. In order to improve our understanding of the *N. benthamiana* methyl cycle, the role of the five identified NSH genes in adenosine salvage must be explored. Furthermore, these same studies identified an uncharacterized, transcription and ADK-independent TGS suppressor activity of AL2, while the currently described L2 mechanisms are insufficient to inhibit the pathway that is active in reproductive plants. Interestingly, the AL2-C33A mutant sustained GFP expression, but visual and Northern blot data suggests that reproductive plants had reduced TGS reversal. The AL2-C33A mutation lies within a conserved CCHC zinc-finger domain and these regions are well-known to be involved in protein-protein
interactions. As a matter of a fact, the AL2-C33A mutant not only has reduced transcription activation, but BiFC data indicated that the AL2-C33A mutant abolishes dimerization (Yang et al. 2007). Therefore, the CCHC domain could be involved in protein-protein interactions that are required for the newly identified transcription-independent TGS suppression mechanism. Thus, the AL2-C33A and other CCHC mutants may be useful tools for characterizing this suppressor activity.

To conclude, studies of AL2 reveal it to be a remarkably versatile PTGS and TGS suppressor that uses both transcription-dependent and –independent mechanisms to inhibit local silencing and subsequent systemic spread, as well as reversing established silencing in vegetative and reproductive plants (Voinnet 1999; van Wezel et al. 2002, Vanitharani et al. 2004, Trinks et al. 2005, Wang et al. 2005, Yang et al. 2007). L2 is limited to transcription-independent mechanisms and shares many of these activities. In addition, both AL2 and L2 have recently been shown to reverse TGS (Buchmann et al. 2009). The multiple aspects of RNA silencing suppressed by AL2 and L2, and the discovery of additional silencing suppressors encoded by geminiviruses and their satellites, emphasizes the critical role that silencing pathways play in modulating geminivirus pathogenicity (Vanitharani et al. 2005, Bisaro, 2006). Much has been and will be learned about these pathways as we further unravel the mechanisms by which they are suppressed by geminivirus proteins.
Fig. 2.1 AL2, but not L2, blocks the systemic silencing signal. Top: AL2, but not L2, blocks the spread of the systemic silencing signal in a transcription-dependent manner. After 4-6 weeks, evidence of systemic silencing was observed at the crown of the plant, and plants not showing evidence of systemic silencing were counted. Over two or more experiments, total “silencing escape” plants were tallied and the percentage graphed. Error bars represent the range between individual experiments. Bottom: An illustration of systemic silencing progression, from pre-infiltration to a fully silenced leaf.
Fig. 2.2 AL2 and L2 reverse established PTGS in a developmentally-dependent manner. (A) AL2, but not L2, reverses established PTGS in reproductive plants in a transcription-dependent manner. 2-4 weeks after infiltration, tissue sections from older, symptomatic leaves were observed under a confocal microscope at 20x magnification. Deep red indicates chlorophyll autofluorescence and green indicates GFP expression. (B) Analysis of GFP mRNA from systemic tissue. As before, total RNA was extracted from symptomatic tissue and probed with $^{32}$P-GFP antisense probe in a Northern Blot. Loading controls were EtBR stained 18S rRNA bands, with a picture taken prior to transfer to the membrane.
Fig. 2.3 Plant developmental age determines whether geminiviruses can reverse TGS. (A) CaLCuV continues to reverse TGS in reproductive plants, while BCTV cannot. 3 weeks after geminivirus infection vegetative and reproductive plants were screened under UV light for TGS reversal. Deep red represent TGS and is due to chlorophyll autofluorescence, while the yellow-green color indicates GFP expression and TGS reversal. (B) Northern blot analysis of GFP mRNA supports the visual data. Northern blot shown is representative of results observed from at least three individual experiments. The 28S rRNA stained with EtBr is used as a loading control, while GFP RNA was detected with a GFP-riboprobe. (C) Southern blot analysis shows that geminivirus titers are similar in vegetative and reproductive plants. Southern blot data represents three individual experiments. 18S rDNA is used as a loading control and all DNA was detected with radiolabeled oligonucleotide probes. (D) BCTV inhibits ADK
activity in vegetative and reproductive plants in an L2 dependent manner. Crude protein extracts were obtained from Mock, BCTV, and BCTV L2- infected plants and analyzed for ADK activity with thin layer chromatography. The graph represents three individual experiments. The radioactively labeled AMP was quantitated and mock extracts were normalized to 1. Error bars represent SEM. On the right, is an image of a representative TLC plate of ADK activity.

Fig. 2.4 AL2 developmental-independent TGS reversal is transcriptionally-independent. (A) PVX::AL2, PVX::AL21-114, and PVX::AL2 C33A, but not PVX::L2, can reverse TGS in reproductive tissue. These data support geminivirus infection results and indicate that AL2 TGS reversal in reproductive plants is independent of transcription activation. 2-3 weeks after PVX infection vegetative and reproductive plants were screened under UV light for TGS reversal. (B) Northern blot analysis of GFP mRNA supports the visual data. Northern blot shown is representative of results observed from at least three individual experiments. (C) Northern blot analysis shows that PVX titers are similar in vegetative and reproductive plants. Northern blot data represents three individual experiments. RNA was isolated from individual plants, 28S rRNA is used as a loading control and all RNA was detected with radiolabeled oligonucleotide probes.
Fig. 2.5  TGS is independent of ADK activity in reproductive plants. (A) TRV::SAHH and TRV::MET1 reverse TGS in reproductive tissue, while TRV::ADK cannot. These data support geminivirus and PVX infection results and indicate that AL2 TGS reversal in reproductive plants is independent of ADK inhibition. 3 weeks after TRV infection vegetative and reproductive plants were screened under UV light for TGS reversal. (B) Northern blot analysis of GFP mRNA supports the visual data. Northern blot shown is representative of results observed from at least three individual experiments. (C) Northern blot analysis shows that TRV titers are similar in vegetative and reproductive plants. Northern blot data represents three individual experiments. RNA was isolated from individual plants, 28S rRNA is used as a loading control and all
RNA was detected with radiolabeled oligonucleotide probes. (D) TRV::ADK inhibits ADK activity in vegetative and reproductive plants. Crude protein extracts were obtained from Mock, TRV, and TRV::ADK infected plants and analyzed for ADK activity with thin layer chromatography. The graph represents internal duplicates of three individual experiments. The radioactively labeled AMP was quantitated and mock extracts were normalized to 1. Error bars represent SEM. On the right, is an image of a representative TLC plate of ADK activity.

**Fig. 2.6 ADK knock-down cannot reduce global cytosine methylation in reproductive plants.** (A) CaLCuV can significantly reduce global CNG methylation in both vegetative and reproductive plants, while BCTV can only do so in vegetative plants. (B) TRV::SAHH can significantly reduce global DNA methylation in both vegetative and reproductive plants, while TRV::ADK cannot. These results support the TGS reversal data obtained from the GFP transgene. DNA was isolated from individual plants and digested with MspI. Radiolabeled dCTP was used to fill-in the resulting G over-hang and increased incorporation represents reduced CNG methylation. BCTV L2- and TRV served as a negative control. Graphs represent an internal duplicate of three individual experiments from three individual plants. All data points are normalized to the associated mock control. Error bars represent SEM and 99% (**) or 95% (*) confidence levels are indicated as determined by the student's t test.
<table>
<thead>
<tr>
<th></th>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>aCalCuVIRA1</td>
<td>5' TTACAAAATGCCACGAAACCCTAAATCG 3'</td>
</tr>
<tr>
<td>2</td>
<td>sCaLCuVIRA2</td>
<td>5' CCTATTGTTTTTGCTTTGGTGTCCA 3'</td>
</tr>
<tr>
<td>3</td>
<td>sCaLCuVIRA3</td>
<td>5' ATGGCCCGCGATTTTTTGTTGTC 3'</td>
</tr>
<tr>
<td>4</td>
<td>aCaLCuVIRA4</td>
<td>5' ATAGCTAACTTGTCAAGGCGAT 3'</td>
</tr>
<tr>
<td>5</td>
<td>BCTVIRsR_1</td>
<td>5' GTTTTGATTTGTTGCCCATTATCTCG 3'</td>
</tr>
<tr>
<td>6</td>
<td>BCTVIRsR_2</td>
<td>5' TCTTCAGGAAGTTTCCCCGCTCA 3'</td>
</tr>
<tr>
<td>7</td>
<td>BCTVIRsF_1</td>
<td>5' GACACGTTGGAAGGGTACTGATAAGAA 3'</td>
</tr>
<tr>
<td>8</td>
<td>BCTVIRsF_2</td>
<td>5' TAAGATTGTTGACCTGTTCAATAGGAAAGG 3'</td>
</tr>
<tr>
<td>9</td>
<td>q18SrDNA-2F</td>
<td>5' AAACGGCTACCACATCCAAG 3'</td>
</tr>
<tr>
<td>10</td>
<td>q18SrDNA-2R</td>
<td>5' ACTCGAAAAAGCCCGTTATT 3'</td>
</tr>
<tr>
<td>11</td>
<td>q18SrDNA-3F</td>
<td>5' AACGCGAAAGCATTGCCCAAG 3'</td>
</tr>
<tr>
<td>12</td>
<td>q18SrDNA-3R</td>
<td>5' TGCTCGGCGATCTTTATGGT 3'</td>
</tr>
<tr>
<td>13</td>
<td>PVXgp12682R</td>
<td>5' AAACCTCGGCGCTCATGCTATTTGCTCCACATCCAAG 3'</td>
</tr>
<tr>
<td>14</td>
<td>PVXgp12763R</td>
<td>5' TGTTAATGTGTCATTTCCCTCAGGACTACA 3'</td>
</tr>
<tr>
<td>15</td>
<td>PVXgp12300R</td>
<td>5' AGTTGATTGCDGCGAGGACCTAAGAACAACC 3'</td>
</tr>
<tr>
<td>16</td>
<td>PVXgp11213R</td>
<td>5' TCGCGTAGTCCTCAAGGACCAACAAC 3'</td>
</tr>
<tr>
<td>17</td>
<td>PVXgp11917R</td>
<td>5' ATCCCTCTGGGATTAAACAGCGCCTT 3'</td>
</tr>
<tr>
<td>18</td>
<td>PVXgp1328R</td>
<td>5' TCCGCGGGTTTCTTCTCATGTAGT 3'</td>
</tr>
<tr>
<td>19</td>
<td>PVXgp13961R</td>
<td>5' CGGCTGGAGTGTAACACTTTTGCTGT 3'</td>
</tr>
<tr>
<td>20</td>
<td>TRVCP50</td>
<td>5' CCACCCCAAGAATCGTCCATCAATG 3'</td>
</tr>
<tr>
<td>21</td>
<td>TRVCP174</td>
<td>5' ATTCAGTGCCGCTAGTAACCAGTGTCT 3'</td>
</tr>
<tr>
<td>22</td>
<td>TRVCP355</td>
<td>5' CCCTTTGTGTTGGATTCGCAGGGAA 3'</td>
</tr>
<tr>
<td>23</td>
<td>TRVCP302</td>
<td>5' TCATCGTCTCGTCTTTGAACAGCCT 3'</td>
</tr>
<tr>
<td>24</td>
<td>TRVCP208</td>
<td>5' AGCTGCTCCACCACAACCTAATCC 3'</td>
</tr>
</tbody>
</table>

**Figure 2.7. Oligonucleotides used in this study.** Oligonucleotides used for probing Northern and Southern blots for viral titers or mRNA expression.
CHAPTER 3

_Arabidopsis_ Double-Stranded RNA Binding Protein DRB3 Partners with DCL3 and AGO4 in Methylation-Mediated Defense against Geminiviruses

The following chapter will be submitted for publication as follows: Priya Raja, Jamie N. Jackel, Sizhun Li, Isaac M. Heard, and David M. Bisaro _Arabidopsis_ Double-Stranded RNA Binding Protein DRB3 Partners with DCL3 and AGO4 in Methylation-Mediated Defense against Geminiviruses. I produced Figure 3.4 and Figure B.4, as well as Figures 3.2 and 3.7 in collaboration with Priya Raja.

3.1 Introduction

Plants use a sophisticated RNA silencing machinery to control the expression of selected genes and repress resident transposons. Silencing can occur post-transcriptionally through several pathways mediated by microRNA (miRNA) or small interfering RNA (siRNA), or transcriptionally by small RNA-directed DNA methylation (RdDM) (Baulcombe 2004, Broderson and Voinnet 2006, Vaucheret 2006, Henderson and Jacobsen 2007, Matzke et al. 2009, Law and Jacobsen 2010). In addition, RNA silencing acts as an effective defense against invading nucleic acids, including those of RNA and DNA viruses (Ding and Voinnet 2007, Mlotshwa et al. 2008, Ruiz-Ferrer and Voinnet 2009). The small RNAs characteristic of individual silencing pathways are generated by distinct DICER-LIKE (DCL) ribonucleases that process larger double-stranded RNA (dsRNA) and hairpin precursors. In _Arabidopsis_, DCL1 generates 21
nucleotide (nt) miRNAs, whereas DCL2, DCL3, and DCL4 produce siRNAs that are typically 22, 24, and 21 nt in length, respectively. These small RNAs program ARGONAUTE (AGO)-containing effectors known as RISC (RNA-induced silencing complex) that mediate sequence-specific translational inhibition, transcript cleavage, DNA methylation, or some combination of these, depending on the small RNAs and AGO proteins they contain.

Dicer proteins possess RNase III domains and dsRNA binding motifs (dsRBMs), the latter mediating sequence non-specific dsRNA binding as well as protein-protein interactions (Saunders and Barber 2003, Chang and Ramos 2005, Hiraguri et al. 2005). An important group of Dicer-interacting factors includes proteins related to *C. elegans* RDE-4 (Tabara et al. 2002). These proteins also have dsRBMs but lack an obvious catalytic domain. Some, for example *Drosophila* R2D2 and R3D1/Loquacious, cooperate with specific Dicers in small RNA biogenesis and/or RISC loading (Liu et al. 2003). Related proteins include TRBP and PACT in mammals, and in plants comprise a family known simply as dsRNA binding proteins (DRBs). The DRB family in *Arabidopsis* has five members. The best studied is DRB1/HYL1, which interacts with DCL1, is required for efficient miRNA accumulation, and is involved in selecting the guide strand loaded into RISC (Hiraguri et al. 2005, Han et al. 2004, Vasquez et al. 2004, Curtin et al. 2008, Eamens et al. 2009). DRB2 processes specific miRNAs subsets in the shoot apical meristem, suggesting considerable complexity in this developmentally critical tissue (Eamens et al. 2012a). DRB3 and DRB5 are dispensable for miRNA processing, but assist in silencing transcripts targeted by DRB2-associated miRNAs (Eamens et al. 2012b). DRB4 protein physically and functionally interacts with DCL4 in
trans-acting siRNA (ta-siRNA) biogenesis and in siRNA-mediated defense against RNA viruses (Curtin et al. 2008, Adenot et al. 2006, Nakazawa et al. 2007, Qu et al. 2008). In addition, the *Cauliflower mosaic virus* (CaMV) P6 silencing suppressor interacts with DRB4, suggesting that a DRB4:DCL4 complex also targets dsRNA derived from DNA virus transcripts (Haas et al. 2008). Although it has been suspected that at least one of the DRB proteins also plays a role in the nuclear RdDM pathway, studies have so far proved inconclusive, and mutational analysis indicates that none of the DRB proteins are essential for maintaining repressive transposon methylation (Curtin et al. 2008, Eamens et al. 2011). However, DRB2 and DRB4 have been shown to have antagonistic effects on the accumulation of small RNAs produced by RNA polymerase IV, a component of the RdDM pathway (Pelissier et al. 2011). Thus multiple DRB proteins may have direct or indirect roles in the establishment or maintenance of RdDM.

Geminiviruses have small (2.5 to 3.0 kb) genomes of circular, single-stranded DNA that replicate in infected cell nuclei by a rolling circle mechanism that utilizes double-stranded DNA replicative form (dsDNA RF) intermediates (Hanley-Bowdoin et al. 2004, Rojas et al. 2005, Jeske 2009). Viral genomes specify four to seven proteins, none of which have polymerase activity. Instead, geminiviruses rely on host machinery for replication and transcription, both of which occur on viral chromatin templates composed of dsDNA RF and cellular histones organized as typical nucleosomes. Viral mRNAs are subject to post-transcriptional gene silencing (PTGS), and as a counter-defensive geminivirus proteins suppress this cytoplasmic aspect of silencing (Vanitharani et al. 2005). In addition, we have shown that RNA-directed methylation of viral chromatin leading to transcriptional gene silencing (TGS) is a potent antiviral defense which is also
suppressed by geminivirus proteins (Raja et al. 2008, Buchmann et al. 2009, Raja et al. 2010, Yang et al. 2011). Genetic analysis has revealed that similar pathways are used to methylate geminivirus chromatin and resident transposons, suggesting that geminiviruses could serve as sensitive probes for the identification and analysis of methylation pathway components and effectors (Henderson and Jacobsen 2007, Raja et al. 2008, Lisch 2009). Here, by analyzing the responses of drb mutant plants to geminivirus infection, we present evidence indicating that DRB3, in addition to its role in the miRNA pathway, also participates with DCL3 and AGO4 in methylation-mediated defense against DNA viruses.

3.2 Materials and Methods

3.2.1 Arabidopsis mutants

DCL mutants were obtained from Dr. J.C. Carrington (Xie et al. 2004), and all others from the Arabidopsis Biological Research Center at The Ohio State University. The following seed stocks were used: wild-type Col-0 (CS60000) Columbia ecotype, drb2 (SALK_012017 and CS849395/At2g28380), drb3 (SALK_022644/At3g26932), drb4 (SALK_000736/At3g62800) (Adenot et al. 2006), drb5 (SALK_031307C/At5g41071), dcl2-1 (SALK_064627/At3g03300), dcl3-1 (SALK_005512/At3g43920), dcl4-2 (GABI_160G05/At5g20320), wild-type Ler-0 (CS20) Landsberg erecta ecotype, ago 4-1 gl1-1 (CS6364/At2g27040) (Zilberman et al. 2003). All mutants were homozygous, and when necessary this was verified by genotyping. Plants were reared in growth rooms at 22°C with 12 hour light/dark cycles.
3.2.2 Virus inoculation

Agroinoculation of *Arabidopsis* plants with CaLCuV or BCTV was carried out as previously described (Raja et al. 2008). Plants were inoculated within 5 days of bolting. Bolts were cut where they emerged from the rosette and inoculum was applied to the freshly cut stem, which was then punctured with an insect pin multiple times. CaLCuV symptoms were observed and plants harvested 14-21 days post-inoculation; BCTV symptoms were observed and plants harvested 21-30 days post-inoculation. Inflorescence tissue showing visible symptoms was harvested from *Arabidopsis*. For each sample, tissue was pooled from four infected plants. For BCTV recovery experiments, plants were agroinoculated with BCTV or BCTV L2 and after the primary harvest, plants were allowed to continue growing under the same conditions. The BCTV L2-I null mutant has been previously described (Hormuzdi and Bisaro 1995). Observations were made about symptom development in secondary inflorescence tissue, which was then harvested as secondary tissue 14-21 days after the harvesting of primary tissue.

3.2.3 BiFC analysis

Interactions between DCL and DRB proteins in plants were tested using BiFC (Hu et al. 2002). The construction of BiFC expression vectors using enhanced YFP has been described (Yang et al. 2007). cDNA clones for DRB3 (U66272, 1080 bp) and DRB4 (C104842, 1099 bp) were obtained from the Arabidopsis Biological Resource Center. DCL3 and DCL4 cDNAs were obtained from Dr. J.C. Carrington. PCR primers were used to amplify these with the introduction of a *Paci* site at the 5' end and an *AseI* site at
the 3' end. The PCR products were subsequently digested with PacI and AsclI and ligated into similarly digested BiFC vectors p2YN, p2YC, pYC and pYN.

Cloned plasmids were transformed into *A. tumefaciens* strain C58C1, and cultures were used to infiltrate 3-4 week old *N. benthamiana* plants as previously described (Wang et al. 2005). Briefly, agrobacterium cells grown in liquid culture (OD$_{600}$ = 1) were sedimented and incubated for 3 hours in infiltration medium containing MES, acetosyringone and MgCl$_2$. Cultures containing p2YN- and p2YC-based plasmids were mixed 1:1 immediately prior to infiltration of young leaves using a blunt syringe. RFP-histone 2B and RFP-fibrillarin were used as markers for the nucleus and nucleolus, respectively (Chakrabarty et al. 2007). Leaf tissue was analyzed by microscopy approximately 36 hours post-infiltration using a Nikon PCM 2000 confocal laser scanning microscope equipped with argon and green helium neon lasers with excitation wavelengths of 488 nm and 544 nm, respectively. To record YFP fluorescence, a band-pass emission filter (EM515/30HQ) with a 450- to 490 nm excitation wavelength and 515 nm emission wavelength was used. To record RFP fluorescence, a 565 nm long-pass filter (E565LP) was employed. Images were captured using Simple PCI Software and compiled with Adobe Photoshop.

### 3.2.4 Co-immunoprecipitation and immunoblot analysis

Immunoprecipitation (IP) experiments were performed to verify that DRB3 can physically associate with DCL3 and AGO4. Adenosine kinase (ADK) antibody was a negative control (Wang et al. 2003). Total protein extracts were prepared from leaves of transgenic, dcl3 mutant *Arabidopsis* seedlings expressing a DCL3-FLAG fusion protein.
(gift from Dr. Craig Pikaard) (Pontes et al. 2006), *N. benthamiana* plants transiently expressing DRB3 fused with double hemagglutinin, 6 histidine tags (HA<sub>2</sub>HIS<sub>6</sub>-DRB3), or *N. benthamiana* co-expressing HA<sub>2</sub>HIS<sub>6</sub>-DRB3 and FLAG-AGO4. The DRB3 and AGO4 proteins were expressed from the *Tobacco mosaic virus*-based TRBO vector as described (Lindbo 2007). Immunoprecipitation was performed with monoclonal FLAG antibody (Sigma-Aldrich F3165). To assess DRB3-DCL3 interaction, total protein extract (2.0 g) from leaves of *Arabidopsis* plants expressing DCL3-FLAG was mixed with 0.2 g extract from *N. benthamiana* leaves expressing HA<sub>2</sub>HIS<sub>6</sub>-DRB3. To verify DRB3-AGO4 interaction, total protein extract from 2 g of *N. benthamiana* leaves co-expressing HA<sub>2</sub>HIS<sub>6</sub>-DRB3 and FLAG-AGO4 was used. FLAG or ADK antibodies (10 µg) were incubated with protein G agarose (Millipore) and protein extracts for 2 hours at 4°C.

Proteins in immune complexes were fractionated by 10% SDS–PAGE (acrylamide:bisacrylamide, 49:1) and analyzed by immunoblot using 1:1000 dilutions of FLAG antibody, ADK antibody, or monoclonal HA-peroxidase antibody (Sigma-Aldrich H6533). Secondary antibodies included 1:1000 dilutions of horseradish peroxidase (HRP)-linked anti-mouse IgG (Sigma-Aldrich, A5906) for anti-FLAG, or HRP-linked anti-rabbit IgG (Sigma-Aldrich, A1949) for anti-ADK. Enhanced detection was performed using Supersignal West Pico (for DRB3 and AGO4) or Femto (for DCL3) chemiluminescent substrates (Pierce).
3.2.5 Bisulfite sequencing

Bisulfite sequencing was performed as described (Raja et al. 2008, Frommer et al. 1992). DNA isolated from infected plant tissue was linearized overnight using appropriate restriction enzymes (SacI for BCTV, XmnI for CaLCuV). Proteinase K digestion was subsequently carried out overnight, followed by bisulfite conversion using CT conversion reagent (EZ-DNA Methylation Gold; Zymo Research). Primers were designed against converted template, and the intergenic region of the viral genome was amplified by PCR. The PCR product was purified using Promega Wizard columns and TA cloned, and individual clones were sequenced at The Ohio State University Plant Microbe Genomics Facility. Conversion controls are discussed in Results. The forward and reverse primers used to amplify CaLCuV and BCTV intergenic regions (IR) following bisulfite conversion are listed in Fig. B.6 of supplemental material. Bisulfite reactions were carried out in a thermocycler. Data were analyzed and dot plots prepared using Kismeth (Gruntman et al. 2008).

3.2.6 Small RNA analysis

RNA was isolated from 0.2 g of symptomatic floral tissue using Ribozol (Amresco) or Trizol (Invitrogen) reagent. Each sample was pooled from 3-4 plants. Small RNAs were analyzed as described (Qu et al. 2008). Total RNA (5-8 µg) was loaded on a 16% denaturing polyacrylamide gel containing 8 M urea in 0.1X TBE. The gel was run until 10 minutes after the bromophenol blue dye ran off the gel. The gel was electroblotted onto Hybond nylon membrane in 0.5X TBE, 4°C, at 40 V, for 1 hour. RNA was
crosslinked to the membrane using a Stratagene UV crosslinker at 1700 kJ. The membrane was prehybridized in Ambion Ultrahyb oligo for 1 hour at 40-42°C. A mixture of oligonucleotides (20 µM) was used as probe. Oligonucleotides were labeled with Fermentas T4 polynucleotide kinase at 37°C for 30-60 minutes. Labeled probes were denatured and hybridization was carried out overnight at 40-42°C. The membrane was washed 3 times for 20 minutes each with 2xSSC, 0.5% SDS at the same temperature. Ethidium bromide stained rRNA was used as a loading control. Oligonucleotides used as size markers are listed in supplemental material Fig. B.6, as are sense and antisense intergenic region and coat protein oligonucleotides that were used as hybridization probes.

3.3 Results

3.3.1 Arabidopsis drb3 mutants are hypersusceptible to geminivirus infection

Methylation-deficient Arabidopsis mutants exhibit hypersensitivity to geminiviruses that correlates with reduced cytosine methylation of the viral genome (Raja et al. 2008). To assess whether Arabidopsis DRB proteins might be involved in methylation-mediated defense, drb2, drb3, drb4, and drb5 mutant plants were challenged by geminivirus infection. DRB1 was not included in this study because it is known to exclusively participate in the miRNA pathway. Initial experiments were performed with Cabbage leaf curl virus (CaLCuV, genus Begomovirus) and Beet curly top virus (BCTV, genus Curtovirus) to evaluate responses to distinct types of geminiviruses. To observe a range
of symptom enhancement, plants were inoculated under conditions that normally elicit mild symptoms.

Geminivirus-infected \( drb2 \), \( drb4 \), and \( drb5 \) mutant plants showed mild to moderate disease enhancement relative to wild-type plants. Disease ratings with these mutants ranged from 1 to 4, where 1 indicates leaf curling and floral deformation typically seen with wild type plants, and 4 indicates very severe curling and deformation with extensive stunting (Raja et al. 2008). By contrast, disease symptoms in \( drb3 \) plants inoculated with either CaLCuV or BCTV were greatly enhanced and accompanied by considerable stunting (disease rating 3-4) (Fig. 3.1A and 3.1B). Bisulfite sequencing was carried out to assess the methylation status of the viral intergenic region (IR). The IR (~300 bp) contains the origin of replication flanked by divergent promoters, and in CaLCuV includes 79 cytosines in different contexts, with 19 CG, 13 CNG, and 47 CHH sites (where H is any residue other than G). Following bisulfite treatment, the viral strand was amplified by PCR and 12 clones were sequenced per treatment.

Bisulfite sequencing revealed that enhanced disease was accompanied by a substantial reduction in methylation. The proportion of methylated cytosine residues in the IR of \( drb3 \) mutants (~20%) was considerably less than was observed in viral genomes obtained from wild type plants (>40%, Col-0 ecotype) (Fig. 3.1C). We concluded that the phenotype of geminivirus-infected \( drb3 \) mutant plants was consistent with a defect in the methylation pathway.
3.3.2 *Arabidopsis dcl3* and *drb3* mutants are unable to recover from geminivirus infection

Perhaps the most compelling argument for methylation as an antiviral defense comes from studies that have associated the methylation pathway with host recovery from geminivirus disease. Recovery occurs when tissues arising after the establishment of a systemic infection exhibit symptom remission and contain very little virus. *Nicotiana benthamiana* and *Arabidopsis* plants rarely recover from infection with wild type BCTV, but nearly always recover from infection with BCTV L2′ mutant virus, which lacks the L2 (also known as C2) pathogenicity factor that suppresses both PTGS and TGS and non-specifically inhibits methylation (Hormuzdi and Bisaro 1995, Wang et al. 2003, Wang et al. 2005, Raja et al. 2008, Buchmann et al. 2009). The small amount of geminivirus DNA found in recovered tissue is hypermethylated in the IR. In addition, we previously showed that *ago4* plants are unable to recover from BCTV L2′ infection and cannot hypermethylate the IR, confirming that the RdDM pathway is required for recovery (Raja et al. 2008, Hagen et al. 2008, Rodriguez-Negrete et al. 2009). Based on these observations, we reasoned that an analysis of mutant plants for recovery might provide a sensitive and definitive means of identifying methylation pathway components.

To test this hypothesis, *dcl2*, *dcl3*, and *dcl4* plants were examined. Considerable functional redundancy exists among *Arabidopsis* DCL enzymes, although DCL3 and the 24 nt siRNAs it generates are clearly associated with chromatin methylation (Xie et al. 2004, Gasciolli et al. 2005, Henderson et al. 2006). However, mild to moderate symptom enhancement was observed following BCTV or CaLCuV infection of *Arabidopsis dcl2, dcl3*, and *dcl4* plants in earlier studies, precluding unequivocal identification of an
individual DCL protein as the key player in methylation-mediated defense (Raja et al. 2008, Blevins et al. 2006). Similar results were again obtained following inoculation of these \textit{dcl} mutants with BCTV \textit{L2}. However, after the establishment of systemic infections and removal of primary infected tissue, only \textit{dcl3} mutant plants proved unable to recover from infection. Severe symptoms appeared in all new axillary secondary shoots of all 32 \textit{dcl3} plants tested (Fig. 3.2). By contrast, symptom remission was observed in new shoots of wild type, \textit{dcl2}, and \textit{dcl4} plants (32 each, see Fig. B.1). Thus, a host recovery assay could discriminate between functionally redundant DCL activities and identify DCL3 as the enzyme most associated with methylation-mediated defense.

To confirm a role for DRB3 in methylation, a panel of \textit{drb} mutant plants was inoculated with BCTV \textit{L2} and, as observed with wild type BCTV, symptoms were much more severe in \textit{drb3} mutants. As expected, new secondary shoots of wild type plants infected with BCTV \textit{L2} recovered from infection and showed little evidence of disease, as did nearly all new shoots of the \textit{drb2}, \textit{drb4}, and \textit{drb5} mutants (32 plants each, see Fig. S1 in supplemental material). However, while new tissues of \textit{drb4} and \textit{drb5} plants did not display floral deformation, they were occasionally somewhat stunted. By contrast, \textit{drb3} mutants did not recover, and severe symptoms appeared in all new shoots of the 32 \textit{drb3} plants inoculated with BCTV \textit{L2} (Fig. 3.2). Together with previous studies of \textit{ago4} mutants (Raja et al. 2008), these experiments confirm that RdDM conditioned by DCL3 and AGO4 is required for host recovery from geminivirus infection, and provide strong genetic evidence linking DRB3 to this pathway.
3.3.3 Arabidopsis dcl3 and drb3 mutants cannot hypermethylate the viral genome

We previously showed that the relatively small amount of BCTV L2 DNA in recovered wild type plants is hypermethylated, while the same virus from comparable but non-recovered tissue of ago4 mutants is not (Raja et al. 2008). Thus, we compared the methylation status of BCTV L2 DNA obtained from recovered tissue with that from non-recovered tissue of drb3, dcl3, and ago4 mutants. Mutant drb4 and dcl4 plants were included in this analysis because a DRB4:DCL4 complex is involved in PTGS-mediated defense against RNA viruses and DNA virus transcripts (Blevins et al. 2006, Bouche et al. 2006, Deleris et al. 2006, Qu et al. 2008, Haas et al. 2008). Bisulfite sequencing was carried out as before, except that primers designed to amplify the viral strand of the BCTV IR were employed. The BCTV IR contains 42 cytosines in the following contexts: 10 CG, 7 CNG, and 25 CHH. PCR products were cloned and sequenced. In two independent experiments the sequences of 12 and 9 clones, respectively, were determined per treatment. It should be noted that, to preclude common artifacts of bisulfite sequencing, rigorous conversion controls are routinely included in all bisulfite experiments. Here, as an internal control, a bacterial plasmid containing CaLCuV DNA was added to BCTV-infected sample extracts prior to treatment with bisulfite reagent. Alternatively, as a parallel control, a plasmid containing BCTV DNA was added to an uninfected plant DNA extract. Data were accepted only when control DNA conversion was complete (12 clones analyzed per treatment). In some cases, plasmid DNA was methylated in vitro with CG methylase before being added to plant extracts. In these cases, data were accepted only when conversion was complete at non-CG sites, and all
CG sites remained unconverted. Bisulfite sequencing data for all plants examined are illustrated graphically in Fig. 3.3A. Cytosine methylation profiles representing individual BCTV $L2^-$ IR clones obtained from wild type (Col-0), $drb3$, and $dcl3$ plants are also presented (Fig. 3.3B, 3.3C, and 3.3D). Profiles for all analyzed plants are shown in appendix B, Fig. B.2.

The majority of genomes from recovered tissues of wild type plants were hypermethylated in all sequence contexts, with ~63% (Col-0) to ~77% (Ler-0) of total cytosines methylated (Fig. 3.3A). By contrast, methylation levels were substantially reduced in viral genomes obtained from $drb3$, $dcl3$, and $ago4$ mutants. Specifically, considerably less methylation was observed in non-recovered $drb3$ plants compared to recovered $drb4$ plants (34% and 54%, respectively), and in non-recovered $dcl3$ plants relative to recovered $dcl4$ plants (48% and 70%, respectively). The greatest reduction relative to comparable wild-type (Ler-0) was noted in genomes from $ago4$ mutant plants (38% compared to 77%). In a second, independent experiment involving wild type, $drb3$, and $dcl3$ plants, similar results were obtained. In this case, the proportion of cytosines methylated were 62%, 35%, and 43%, respectively (see appendix B, Fig. B.3). Given that similar outcomes were observed with $drb3$, $dcl3$, and $ago4$ in independent experiments reported here and in previous studies (Raja et al. 2008), it is highly unlikely that our results are significantly impacted by selection of sibling clones for bisulfite sequencing. These results confirm that geminivirus genomes obtained from recovered tissues (wild type, $drb4$, and $dcl4$) are more densely methylated than those from non-recovered $drb3$, $dcl3$, $ago4$ plants, and further associate DRB3 with the antiviral RdDM pathway that involves DCL3 and AGO4.
3.3.4 Both active and repressed viral genomes are present in infected plants

Histone H3 associated with geminivirus DNA carries modifications characteristic of both active (acetylated H3; H3Ac) and repressed (H3K9me2) chromatin (Raja et al. 2008), and cytosine methylation profiles reveal mixtures of mostly hypermethylated (likely repressed) or mostly hypomethylated (likely active) genomes (Fig. 3.3). Together, these observations suggest that populations of active and repressed viral genomes co-exist in infected plants, and that the major effect of specific host or viral mutations that lead to hypersusceptibility or recovery is to alter their relative proportions.

Chromatin immunoprecipitation followed by bisulfite sequencing (ChIP-BS) was employed to address this issue. Beginning with non-recovered tissue from a BCTV L2- infected mutant plant (input), ChIP-BS was performed with antibody specific for H3Ac (acetylated at lysine 9 or 14) or H3K9me2, and precipitated DNA was treated with bisulfite reagent. Following PCR amplification with IR primers, clones were obtained and sequenced. We observed that H3Ac precipitation led to a moderate enrichment in hypomethylated genomes compared to input DNA (~40% compared to ~50% of cytosines methylated). By contrast, all of the genomes associated with H3K9me2, a hallmark of repressed chromatin, were hypermethylated (~90% of cytosines methylated) (Fig. 3.4). Similar results were obtained in a second, independent experiment carried out using recovered tissue from a plant infected with BCTV L2+. Precipitation with H3Ac antibody again led to a moderate enrichment of hypomethylated genomes relative to input DNA (~55% compared to ~70% of cytosines methylated), and all genomes associated with H3K9me2 were hypermethylated (~90% of cytosines methylated) (see appendix B,
These findings support the view that populations of active and repressed genomes are present in infected plants, and indicate that even during a productive, symptomatic infection, a significant fraction of viral genomes are silenced by cytosine and H3K9 methylation.

3.3.5 DRB3 interacts with DCL3 and AGO4

The similar recovery phenotypes observed with *ago4, dcl3,* and *drb3* mutant plants prompted us to ask whether these proteins physically interact. We first employed bimolecular fluorescence complementation (BiFC) to examine *in vivo* interactions between these proteins, using DCL4 and DRB4 as a positive interaction control. BiFC involves fusion of potentially interacting proteins with the N- or C-terminal portions of yellow fluorescent protein (YFP), and subsequent co-expression (Hu et al. 2002). Protein association reconstitutes YFP, producing a fluorescent signal that at once reveals the interaction and where it occurs in the cell. In these studies, *N. benthamiana* leaves were co-infiltrated with mixtures of *Agrobacterium tumefaciens* cells harboring constructs that express the fusion proteins (Yang et al. 2007). Test proteins were expressed as both N- and C-terminal fusions with both the N- and C-terminal portions of YFP. Proteins were judged to interact when signal was observed in multiple combinations, and not to interact when signal was absent in all combinations.

In agreement with genetic evidence linking DRB3 to the methylation pathway, we found that DRB3 and DCL3 interact in the nucleus (Fig. 3.5A; additional images in appendix B, Fig. B.5). The DRB3:DCL3 signal was observed throughout and was especially intense in sub-nuclear bodies, some but not all of which were associated with
nucleoli (Fig. 3.5B). This is consistent with previous localization of DCL3, AGO4, and siRNA in Cajal bodies, ribonucleoprotein processing centers that sometimes associate with nucleoli (Pontes et al. 2006). By contrast, no evidence of complex formation was observed when DRB3 and DCL4 fusion proteins were co-expressed. Thus, our data show that DRB3 specifically interacts with the RdDM-associated DCL3 and not with PTGS-associated DCL4. As a positive interaction control, we confirmed that DRB4 interacts with DCL4 (Hiraguri et al. 2005, Nakazawa et al. 2007). DRB4:DCL4 complexes were similarly observed in nucleolus-associated bodies, although they were also present in the cytoplasm (Fig. 3.5A and 3.5B).

AGO4 has been reported to localize in two distinct sub-nuclear bodies, namely Cajal bodies and smaller AB-bodies, which also contain RNA polymerase V and DRM2 (Li et al. 2006, Pontes et al. 2006, Li et al. 2008). We observed that DRB3 interacts with AGO4 throughout the nucleus, although some signal was also apparent in the cytoplasm (Fig. 3.5A). Nuclear complexes were concentrated in small punctate spots that were distinct in appearance from the larger bodies containing DRB3:DCL3 and did not associate with nucleoli (Fig. 3.5B).

Co-immunoprecipitation (co-IP) experiments were carried out to confirm the interactions. Unfortunately, despite considerable effort, we were unable to express DCL3 in yeast, plant, or E. coli cells using standard methods. Thus dcl3 mutant, transgenic Arabidopsis plants expressing a complementing DCL3-FLAG transgene were employed (Pontes et al. 2006). Arabidopsis DRB3 was expressed as a double hemagglutinin peptide, 6 histidine fusion (HA$_2$His$_6$-DRB3), and AGO4 as a FLAG fusion (FLAG-AGO4), in N. benthamiana plants using a Tobacco mosaic virus-based vector (Lindbo
2007). When protein extracts containing FLAG-DCL3 and HA$_2$His$_6$-DRB3 were mixed and incubated with FLAG antibody, both DCL3 and DRB3 were immunoprecipitated (Fig. 6A). Neither protein was precipitated using a control antibody raised against adenosine kinase (ADK). Similarly, when extracts from *N. benthamiana* plants co-expressing FLAG-AGO4 and HA$_2$His$_6$-DRB3 were incubated with FLAG antibody, both AGO4 and DRB3 were immunoprecipitated (Fig. 3.6B). Together, these BiFC and co-IP studies demonstrate physical associations between DRB3:DCL3 and DRB3:AGO4 that are consistent with functional cooperation in methylation-mediated antiviral defense.

3.3.6 DRB3 is not required for the biogenesis of 24 nt siRNAs by DCL3

While some dsRBM proteins do not significantly impact small RNA levels, others facilitate the biogenesis of small RNA species by the Dicers with which they associate. We asked whether DRB3 is necessary for the accumulation of 24 nt siRNA generated by DCL3 in response to BCTV. While it was previously shown that levels of 24 nt siRNA derived from an RNA virus were not diminished in *drb3* plants (Curtin et al. 2008), virus-specific siRNAs of this size class are considerably more abundant in geminivirus infected plants (Blevins et al. 2006, Akbergenov et al. 2006).

RNA was isolated from wild type, *drb3*, and *dcl3* plants inoculated with BCTV, and infected *drb4*, *dcl4*, and *ago4* plants were included in this study for comparison. RNA samples were fractionated on polyacrylamide gels and probed with a mixture of $^{32}$P-labeled oligonucleotides specific for the BCTV IR and the coat protein (CP)-coding region. Previous experiments using separate IR or CP oligonucleotide probes have
shown that both regions spawn abundant BCTV-derived siRNAs of all size classes (data not shown).

As expected, dcl3 mutants lacked 24 nt siRNAs corresponding to the BCTV genome, and dcl4 mutants showed much reduced 21 nt siRNA levels (Fig. 3.7). However, little change was evident in the type and abundance of small RNA species present in the drb3, drb4, and ago4 mutant plants compared to corresponding wild type plants (Col-0 or Ler-0). We concluded that DRB3, DRB4, and AGO4 are not required for the synthesis or accumulation of BCTV-derived siRNAs.

3.4 Discussion

Nuclear RdDM that involves DCL3 and AGO4 provides a robust defense against geminiviruses. The goal of this study was to determine whether an Arabidopsis DRB protein participates in this epigenetic defense. By taking advantage of the extreme sensitivity of methylation-deficient plants to geminivirus infection, we generated genetic evidence that strongly links DRB3 with this pathway. First, drb3 mutant plants show enhanced susceptibility to geminivirus infection that is accompanied by reduced levels of cytosine methylation in the viral IR. In addition, like dcl3 and ago4 mutants, drb3 plants fail to recover from infection with a virus lacking a suppressor protein that inhibits methylation (BCTV L2'), and are unable to carry out the IR hypermethylation that is typically observed in recovered, wild-type plants. Significantly less methylation was observed in non-recovered drb3 compared to recovered wild type plants (~34-35% and 62-63% methylated cytosines, respectively).
Physical interaction is another indication of functional cooperation between DRB3, DCL3, and AGO4. We found that DRB3:DCL3 complexes preferentially accumulate in relatively large bodies that are occasionally associated with nucleoli, similar to Cajal bodies known to contain DCL3, AGO4, and siRNA (Pontes et al. 2006). Surprisingly, DRB3:AGO4 complexes were not seen in these structures, but were enriched in smaller, punctate spots reminiscent of AB bodies that contain AGO4, RNA polymerase V, and DRM2 (Li et al. 2006, Pontes et al. 2006, Li et al. 2008). Why sites of DRB3:DCL3 and DRB3:AGO4 complex enrichment do not overlap is unclear, but raises the intriguing possibility that DRB3 choreographs nuclear events by interacting with upstream and downstream pathway components in different locations. DRB3:DCL3 interaction was confirmed in co-immunoprecipitation experiments in which, for technical reasons, interacting proteins were separately expressed in different plants. DRB3-DCL3 complexes were formed after the plant extracts were mixed. DRB3:AGO4 interaction was verified by co-IP following co-expression of these proteins in *N. benthamiana* plants. An earlier study found that a DRB3-YFP fusion protein was present in the cytoplasm following transient expression in *N. benthamiana* cells and, while the existence of a nuclear pool of DRB3 was not addressed, DRB3 did not co-localize with DRB1 in sub-nuclear D-bodies, which contain DCL1 and miRNA precursors (Eamens et al. 2012a). Cytoplasmic localization is consistent with a downstream role for DRB3 in the miRNA pathway (Eamens et al. 2012b). Nuclear localization of DRB3:DCL3 and DRB3:AGO4 complexes is consistent with an additional role for DRB3 in the methylation pathway and, since BiFC will only detect proteins when they are in complex, a cytoplasmic pool of DRB3 would have escaped detection in our analysis.
While our experiments clearly show that *drb3* mutant plants are methylation-deficient from the perspective of geminiviruses, they differ from an earlier study, which found that methylation levels of several cellular transposons were unaltered in *drb3* plants, or in any of the *drb* mutants, including multiple mutants (Curtin et al. 2008). We suggest that this may be ascribed in part to functional redundancy, and further that the enormous demand imposed on the methylation pathway by geminivirus infection allowed us to detect reduced viral DNA methylation in *drb3* plants which could not be observed at endogenous loci. In support of this view, the results of our ChIP-BS experiments indicate that a substantial fraction of the viral genomes in infected plants (which can approach 0.5 x 10^6 copies per cell) are repressed by dense cytosine methylation and by association with H3K9me2 (Fig. 3.4). Thus, we argue that DRB3 is the primary DRB protein involved in the methylation-mediated antiviral defense. While paralogues of DRB3, DCL3, and AGO4 appear to be capable of directing viral genome methylation in the absence of these enzymes, the functionally redundant activities are clearly not sufficient to generate the very high methylation levels necessary to support recovery. It should also be mentioned that while DRB3 expression can be detected in all tissues, levels are particularly high in apical and axillary shoots, which are the tissues where recovery is initiated (Eamens et al. 2012b, Hruz et al. 2008). However, we cannot formally rule out the possibility that, rather than redundant DRB, DCL, or AGO activities, as yet uncharacterized, non-canonical pathways are responsible for geminivirus genome methylation observed in non-recovered *drb3*, *dcl3*, and *ago4* mutant plants.

Maintenance of CG and CNG methylation by the methyltransferases MET1 and CMT3, respectively, is another factor that might obscure a role for DRB3 at transposons
and other endogenous sequences. Non-symmetrical CHH methylation is not recognized by maintenance methyltransferases, and as a result genomic CHH methylation in wild type \textit{Arabidopsis} is at most 20\% at specific loci, and averages less than 2\% genome wide (Cokus et al. 2008, Lister et al. 2008). Geminivirus dsDNA replicative forms, on the other hand, must be methylated \textit{de novo} in infected cells, and maintenance thorough mitosis is not required. Evidence of this is the large proportion of CHH methylation that we routinely observe on the BCTV IR, which can reach 60 to 80\% in wild type Col-0 and Ler-0 infected with a mutant virus that cannot suppress methylation (BCTV \textit{L2}). Thus, it is possible that DRB3 is primarily involved in the establishment of methylation, and if so the effects of its absence would be obvious on viral genomes but difficult to detect at resident transposons.

It bears repeating that rigorous conversion controls are employed in all bisulfite experiments, and we are confident that the pattern of mostly methylated or mostly unmethylated clones that we routinely observe is accurate. However, it is worth considering that geminivirus DNAs exist in multiple forms in infected cells: 1) circular ssDNA, which cannot be methylated, 2) circular dsDNA replicative forms, which are templates for replication and transcription and may or may not be methylated, and 3) linear dsDNA forms of heterogeneous length, which are believed to be replication by products or non-functional genome fragments that also may or may not be methylated. By Southern blot analysis, we have not observed reproducible alterations in the ratio of circular ssDNA to dsDNA in any of the methylation-deficient mutants we have examined. Thus the reductions in viral genome methylation we have observed in known methylation-deficient mutants cannot be ascribed to alterations in relative viral ssDNA to
dsDNA ratios. In addition, previous studies of *Tomato yellow leaf curl China virus* and BCTV have shown that during primary infections most viral genome methylation is found at sequences encompassing or adjacent to viral promoters, including the IR ((Yang et al. 2011) and unpublished results). Similar results have been obtained with other geminiviruses (Rodriguez-Negrete et al. 2009). This methylation could exist either on circular dsDNA forms or on heterogenous linear DNA, or both. It has been reported that the bulk of methylated geminivirus DNA exists as heterogeneous linear dsDNA forms which presumably correspond to all regions of the viral genome (Paprotka et al. 2011). Thus it is possible that fragmented viral DNA is preferentially methylated, or alternatively that heavily methylated viral genomes accumulate double-strand breaks.

Our bisulfite sequencing protocol does not discriminate between viral DNAs of different conformation and, of course, the level of methylation observed over any particular region or with any particular virus is to some extent biased by the primers used to amplify converted DNA. Nevertheless, we observe reproducible differences in the proportion of methylated to unmethylated viral IR clones obtained from wild type plants, or non-methylation mutants (e.g. *dcl4*), compared with mutant plants that are known to be methylation deficient (e.g. *dcl3* and *ago4*). Further, we see the highest methylation levels in recovered wild type plants using a virus (BCTV *L2*) that lacks a suppressor protein known to inhibit methylation. Thus our results accurately reflect relative methylation levels of aggregate populations of viral DNA in infected cells.

Small RNA gel blot experiments showed that the absence of DRB3 does not significantly impact the biosynthesis or accumulation of geminivirus-derived siRNAs, including the 24 nt size class generated by DCL3. That methylation of viral genomes
clearly occurs in \textit{dcl3} mutant plants suggests that small RNAs processed by other DCL enzymes can redundantly participate in the methylation pathway, or that non-canonical mechanisms contribute to geminivirus genome methylation. In any case, the absence of an effect of DRB3 on small RNA profiles is not unprecedented, as some DRB proteins appear to have little or no role in small RNA biogenesis. For example, only a small decrease in 21 nt virus-specific siRNA levels was observed in \textit{drb4} plants infected with \textit{Turnip crinkle virus}, an RNA virus (Qu et al. 2008). Further, although DRB3 was recently shown to participate in miRNA-mediated silencing of specific transcripts, it appears to have no role in miRNA biogenesis (Eamens et al. 2012b). Thus, DRB3 appears to act downstream of both DCL1/DRB2-mediated miRNA processing and DCL3-mediated 24 nt siRNA processing.

Similar to \textit{drb3}, the \textit{ago4} mutation also did not cause a loss or reduction of BCTV-specific siRNAs. This is at odds with a previous study that reported a reduction in 24 nt siRNAs originating from some endogenous loci in the absence of AGO4, and it was proposed that AGO4 binding might stabilize siRNAs (Zilberman et al. 2004). We suggest that our disparate results may be due to the extreme abundance of geminivirus-derived siRNAs in infected plants, which could saturate RISC complexes and obscure a stabilizing effect of siRNA binding.

In conclusion, we set out to determine if a particular DRB protein or proteins contributes in a significant way to geminivirus genome methylation, and our results strongly implicate DRB3. Beginning with an unbiased screen, we showed that: 1) \textit{drb3} mutants uniquely show enhanced susceptibility to geminiviruses; 2) \textit{drb3} mutants cannot recover from infection; 3) \textit{drb3} mutants fail to hypermethylate the viral genome, which is
essential for recovery; 4) the response of *drb3* mutants to geminivirus infection is essentially identical to *dcl3* and *ago4* mutant plants, which are defective for well-characterized components of the RdDM pathway; and 5) DRB3 physically associates with DRB3 and AGO4 in the nucleus. Thus, although participation of DRB3 in cellular DNA methylation was not directly addressed and remains open to debate, these studies provide unequivocal evidence that DRB3 plays an important role in the RdDM pathway that conditions an epigenetic defense against geminiviruses, and lay the ground-work for mechanistic studies of DRB3 function.

**Fig. 3.1** *Arabidopsis drb3* mutants show enhanced susceptibility to geminivirus infection.  (A) Photographs illustrate CaLCuV disease symptoms in wild-type (Col-0 ecotype) and *drb3* mutant plants, shown 14 days post-inoculation.  (B) BCTV symptoms shown 21 days post-inoculation. BCTV has an inherently longer latent period than CaLCuV.  (C) Histograms show the percentage of cytosines methylated in CaLCuV IR DNA isolated from wild-type and *drb3* plants, as determined by bisulfite sequencing.
Fig. 3.2 Arabidopsis *drb3* and *dcl3* mutants do not recover from infection with BCTV L2\(^-\) mutant virus. (A) The photographs show secondary tissue of wild-type (Col-0), *drb3*, *drb4*, *dcl3*, and *dcl4* plants infected with BCTV L2\(^-\) virus. Note recovery (near absence of symptoms) in wild-type, *drb4*, and *dcl4* mutants, and severe disease symptoms in the *drb3* and *dcl3* mutants. Although deformation of floral tissue was not observed in secondary tissue of infected *drb4* mutants, plants were occasionally stunted. An example of a stunted *drb4* plant is shown in this figure. (B) Close-up views of shoots from the same infected plants.
Fig. 3.3 *Arabidopsis drb3 and dcl3 mutants fail to hypermethylate the viral genome.* Methylation of BCTV $L^2$ IR DNA was assessed by bisulfite sequencing. (A) Histograms indicate the percentages of cytosine residues methylated in different sequence contexts in viral DNA obtained from secondary tissues of wild-type (Col-0 or Ler-0) or mutant plants. The *drb* and *dcl* mutants are in Col-0, whereas *ago4* is in the Ler-0 ecotype background. Cytosine methylation profiles of BCTV $L^2$ DNA obtained from (B) recovered wild-type Col-0, (C) non-recovered *drb3*, and (D) non-recovered *dcl3* plants are shown. The dots represent all cytosines in the IR and are color coded according to sequence context (red: CG, blue: CNG, green: CHH). Filled circles indicate methylation, and each line represents the sequence of an individual clone, arranged from most to least methylated. When combined with data from a second, independent
experiment (Figure B.3), *drb3* and *dcl3* treatments were significantly different from wild-type at the 99% (***) and 95% (*) confidence levels, respectively, as determined by Student's *t* test (values indicate the mean +/- SE).

**Fig. 3.4** Hypermethylated viral genomes are associated with H3K9me2. ChIP-BS was performed with extracts from BCTV L2+ infected, non-recovered tissue using antibodies to H3Ac or H3K9me2. Input (extract without ChIP) and ChIP DNAs were bisulfite treated and PCR primers spanning the viral IR were used to amplify associated DNA. Products were cloned and sequenced. The difference between input and H3K9me2 fractionated samples was significant at the 99% (***) confidence level, as determined by Student's *t* test.
Fig. 3.5 **DRB3 interacts with DCL3 and AGO4 in distinct subnuclear bodies.** BiFC analysis of the indicated DCL, DRB and AGO proteins in *N. benthamiana* epidermal cells was performed. Constructs expressing proteins fused to the N- or C-terminal portion of YFP were delivered by agroinfiltration to *N. benthamiana* leaves. Cells were photographed 36 hr post-infiltration using a confocal laser scanning microscope. RFP-histone 2B (RFP-H2B) and RFP-fibrillarin were used as markers for the nucleus and nucleolus, respectively. Protein combinations are indicated above each photograph. The photographs in (A) show lower-magnification views (20X) while (B) shows high magnification views (100X).
Fig. 3.6 DRB3 co-immunoprecipitates with DCL3 and AGO4. Immunoprecipitation (IP) was performed with FLAG antibody (α-FLAG) or ADK antibody (α-ADK, negative control), and proteins in immune complexes were detected by Western blot analysis using HA antibody (α-HA), α-FLAG, or α-ADK. (A) Extracts from plants expressing DCL3-FLAG or HA$_2$His$_6$-DRB3 were mixed ~10:1 to compensate for extremely low levels of DCL3-FLAG expression (below the level of detection in input samples). (B) Extracts were obtained from plants co-expressing FLAG-AGO4 and HA$_2$His$_6$-DRB3.

Fig. 3.7 DRB3 is not required for biogenesis of geminivirus-derived siRNAs. RNA blot hybridization shows the accumulation of virus-specific siRNAs in BCTV-infected wild-type (Col-0 and Ler-0) and mutant (drb3, drb4, dcl3, dcl4, and ago4) plants. Ethidium bromide stained rRNA served as a loading control. The positions of oligonucleotide size markers are indicated at the left.
CHAPTER 4

Arabidopsis RNA polymerases IV and V are required to establish repressive histone methylation, but not cytosine methylation, on geminivirus genomes

The following chapter will be submitted for publication as follows: Jamie N. Jackel and David M. Bisaro. Roles of RNA Polymerases II, IV, and V in the establishment of de novo RNA-Directed DNA Methylation of Geminivirus Genomes.

4.1 INTRODUCTION

Epigenetic modification is a major force shaping the organization and expression of eukaryotic genomes. Repressive chromatin methylation, for example, controls the expression of selected genes, transposable elements, and repeated sequences, and leads to the establishment of transcriptional gene silencing (TGS) and heterochromatin. Many methylation pathway components and effectors have been identified. However, there remains much to learn about how cytosine and histone methylation, such as dimethylation at histone H3 Lysine 9 (H3K9me2), is initiated and maintained. Plants employ a distinct, nuclear branch of RNA silencing, termed RNA-directed DNA methylation (RdDM), to target methylation of cytosine residues (Henderson and Jacobsen 2007, Matzke et al. 2009, Law and Jacobsen 2010, Greenberg et al. 2011). In Arabidopsis, current pathway models include two specialized, plant-specific DNA-dependent RNA polymerases, Pol IV and Pol V (Herr et al. 2005, Kanno et al. 2005,
Onodera et al. 2005, Pontier et al. 2005). Pol IV and Pol V are multisubunit enzymes related to Pol II that are believed to be capable of transcribing methylated DNA templates (Pikaard et al. 2008, Haag and Pikaard 2011). Non-coding Pol IV transcripts are used as template by RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) to produce dsRNA that is processed into 24 nt siRNAs by DICER-LIKE 3 (DCL3). The siRNA is incorporated into an ARGONAUTE 4 (AGO4) or AGO6-containing RNA induced silencing complex (RISC), and the guide strand is targeted by homology to a Pol V scaffold transcript, which remains associated with template DNA. The tethered RISC directly or indirectly recruits chromatin remodeling enzymes as well as cytosine and possibly histone methyltransferases, resulting in the methylation and silencing of chromatin. Two putative SNF2 family chromatin remodelers have been implicated in this pathway. CLASSY 1 (CLSY1) is thought to act with Pol IV (Smith et al. 2007), and DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1) is known to recruit Pol V to at least one genomic locus (Wierzbicki et al. 2008).

It is remarkable that a similar nuclear pathway that depends on non-coding transcription by Pol II directs histone modifications in Schizosaccharomyces pombe (Djupedal et al. 2005, Kato et al. 2005, Buhler et al. 2006). In Arabidopsis, Pol II, Pol IV, and Pol V each contain unique large catalytic subunits (NRPB1, NRPD1 and NRPE1, respectively) but share several subunits (Huang et al. 2009, Ream et al. 2009, Tucker et al. 2011), raising the possibility that the more complex TGS pathway in plants might also involve transcription by Pol II, at least at some loci. Consistent with this notion, a recent study determined that nrpb2-3 plants (containing an amino acid substitution that impairs function of the second largest Pol II subunit) show reduced association of Pol IV and Pol
V at transposons, suggesting that Pol II transcription recruits the silencing machinery to maintain DNA methylation. Moreover, Pol II was found to interact with AGO4, implying that Pol II might directly mediate methylation (Zheng et al. 2009). A more recent analysis of the Arabidopsis methylome in nrpb2-3 plants confirmed that Pol II participates in DNA methylation at some intergenic sites, and also supports a role for Pol II that is independent of Pol IV and Pol V (Stroud et al. 2013). While it is generally accepted that Pol IV and Pol V participate in plant RdDM, the extent of Pol II involvement remains an open question, in part because the nrpb2-3 mutation that was used in these studies is a weak allele (Pol II null mutants are lethal). Conversely, while Pol II-dependent, RNA-directed histone modification is known to occur in S. pombe, the roles of Pol IV and Pol V in targeting repressive histone methylation in plants have not been determined.

Geminiviruses package small (2.5-3.0 kb) circular ssDNAs that replicate in the nucleus through dsDNA intermediates that associate with histones to form minichromosomes. These dsDNA replicative forms serve as template for replication and transcription and, because geminiviruses do not encode polymerases, both processes are dependent on host machinery (Hanley-Bowdoin et al. 2004, Jeske 2009). Geminiviruses are targeted by at least two distinct branches of RNA silencing. Viral mRNAs are subject to post-transcriptional gene silencing (PTGS) (Muangsan et al. 2004, Blevins et al. 2006), and viral genome methylation leading to TGS is additionally employed as an epigenetic defense (Raja et al. 2008m Rodriguez-Negrete et al. 2009). As a counter defense, geminiviruses encode proteins that suppress PTGS and/or inhibit methylation and suppress TGS (Trinks et al. 2005, Wang et al. 2005, Buchmann et al. 2009, Yang et al. 2009).
2011), reviewed in (Bisaro 2006; Raja et al. 2010). The sensitivity of methylation-deficient mutant plants to geminivirus infection, and the existence of effective viral counter-measures, together highlight the importance of methylation as a key defense against DNA viruses.

Because host machinery is responsible for replication and transcription from minichromosome templates, and genetic analysis has shown that similar pathways are used to silence viral DNA and endogenous transposons (Raja et al. 2008), geminiviruses can serve as models for these fundamental processes and their epigenetic regulation. Further, viral chromatin must be methylated *de novo* in infected cells, and the transient nature of the virus replication cycle obviates the need to maintain epigenetic marks through cell division. Thus geminiviruses offer a unique opportunity to examine processes that initiate and target DNA methylation and histone modification. Here, the geminivirus *Beet curly top virus* (BCTV) is used to analyze the roles of Pol IV and Pol V, and chromatin remodeling enzymes, in the establishment of viral DNA and H3K9 methylation. Our results indicate that, while Pol IV and Pol V are important for amplification and spread of viral DNA methylation, they are not required for its establishment or for the production of virus-derived siRNAs. However, both Pol IV and Pol V are essential for deposition of H3K9me2 on viral chromatin.
4.2 MATERIALS AND METHODS

4.2.1 Arabidopsis Mutants

Mutants used in this study were obtained from the Arabidopsis Biological Resource Center at The Ohio State University. The following seed stocks were used: wild-type Col-0 (CS60000) Columbia ecotype, clsy1 (SALK_018319/At3g42670), nrpd1 (SALK_128428/CS66150 and CS66151/At1g63020) (Kanno et al. 2005), drd1 (SALK_132061C/At2g16390), nrpe1 (nrpd1b-11, SALK_022919/At2g40030), nrpd/e2 (nrpd2a, SALK_109513C, SALK_090385C, and SALK_095689/At3g23780). The nrpd1/e1 double mutant was generated by crossing nrpd1 (CS66151) and nrpe1 (SALK_022919) plants. Similar phenotypes were observed and representative results are shown when multiple mutant lines were tested. All mutants were homozygous, and when necessary this was verified by genotyping. Plants were reared in growth rooms at 22°C with 12 hour light/dark cycles.

4.2.2 Virus Inoculation

Arabidopsis plants were infected with BCTV using agroinoculation as previously described (Sunter et al. 2001). At least 32 plants for each experiment were inoculated within 5 days of bolting. Bolts were cut where they emerged from the rosette and inoculum was applied to the freshly cut stem, which was then punctured with an insect pin multiple times. The agrobacterium harboring the BCTV clone was grown to an optical density (600 nm) of 1.0. BCTV symptoms were observed and the inflorescence tissue was harvested 21 days post-inoculation. For each sample, tissue was pooled from
at least three infected plants. For BCTV recovery experiments, at least 32 plants were agroinoculated with BCTV, BCTV L2-1, or BCTV L2-2 and after the primary harvest, plants were allowed to continue growing under the same conditions. The BCTV L2-1 and L2-2 null mutants have been described (Hormuzdi and Bisaro 1995). Secondary inflorescence tissue was harvested 14-21 days following the primary harvest. Similar results were obtained both BCTV L2-1 and BCTV L2-2, thus BCTV L2′ is used throughout to indicate either mutant.

4.2.3 Bisulfite sequencing

Bisulfite sequencing was performed as described (Frommer et al. 1992, Raja et al. 2008). DNA isolated from infected plant tissue was linearized overnight using appropriate restriction enzymes (SacI for BCTV). Proteinase K digestion was subsequently carried out overnight, followed by bisulfite conversion using CT conversion reagent (EZ-DNA Methylation Gold, Zymo Research). Primers were designed against converted template. The intergenic region (IR) or coat protein (CP) coding region of the viral genome was amplified by PCR. The PCR product was purified using GeneJet PCR purification kit (Fermentas) and TA cloned using the pDrive cloning system (Qiagen), and individual clones were sequenced at The Ohio State University Plant Microbe Genomics Facility. As an internal control, a bacterial plasmid containing Cabbage leaf curl virus (CaLCuV) DNA was added to BCTV-infected sample extracts prior to treatment with bisulfite reagent. Alternatively, as a parallel control, a plasmid containing BCTV DNA was added to an uninfected plant DNA extract. Data were accepted only when control DNA conversion is complete (12 clones analyzed per treatment). In some
cases, plasmid DNA was methylated in vitro with CG methylase before being added to plant extracts, and data were accepted only when conversion was complete at non-CG sites and all CG sites remain unconverted. The forward and reverse primers used to amplify BCTV IR and CP coding regions following bisulfite conversion are listed in Supplemental Table 1. Bisulfite reactions were carried out in a thermocycler. Data were analyzed and dot plots prepared using Kismeth (Gruntman et al. 2008).

4.2.4 Chromatin Immunoprecipitation (ChIP)

ChIP was performed as described previously (Johnson et al. 2002, Raja et al. 2008). Symptomatic Arabidopsis inflorescence tissue (0.3 g) was harvested from a pool of at least 3 plants and finely sliced using a razor blade. Tissue was cross-linked under vacuum with formaldehyde for 20 min, then quenched for 10 min with glycine. Tissue was ground in liquid nitrogen and sonicated in lysis buffer under conditions that fragment DNA to an average size of ~500 bp (range 250-800 bp). Protein A agarose beads were used to pre-clear extracts. Immunoprecipitation was carried out with overnight at 4°C using the commercially available antibodies αFLAG (Sigma, F1804), αH3K9me2 (Abcam, ab1220), or αH3Ac (acetylated at lysine 14 and 9) (Millipore, 06-599). Cross-links were reversed at 65°C overnight. DNA was extracted with phenol-chloroform and further purified using mini-prep columns (Promega). Purified DNA (1 µl) was used as PCR template with primers specific for viral DNA or a control sequence (Fig. B.7 of appendix B). Semi-quantitative PCR was performed with an annealing temperature of 55°C for all primers sets. Different cycle numbers were used depending on the primer set: 21 cycles for the BCTV IR, and 26 for BCTV CP in primary infected tissue, 23
cycles for the BCTV IR and 28 for the BCTV CP in secondary infected tissue, 35 cycles for IGN5, Actin and Ta3. PCR products were diluted 2 fold and analyzed by gel electrophoresis. Quantitative real time PCR was performed with SyBR green primers sets, annealing temperatures, and a melt-curve analysis was executed. Cycle threshold values were normalized to input, and IgG background was subtracted. All PCR products were sequenced at The Ohio State University Plant Microbe Genomics Facility to ensure specific amplification of target DNA.

4.2.5 Small RNA analysis

Small RNAs were analyzed as described (Qu et al. 2008), beginning with ~0.2 g of symptomatic floral tissue harvested from a pool of 3-4 plants. Oligonucleotides used as size markers are listed in Fig. B.7 of Appendix B, as are sense and antisense intergenic region and coat protein oligonucleotides that were used as hybridization probes.

4.3 Results

4.3.1 Methylation pathway components Pol IV, Pol V, CLSY1, and DRD1 are involved in defense against geminiviruses.

Previous work has shown that Arabidopsis plants deficient for RdDM pathway components, cytosine and histone methyltransferases, and methyl cycle enzymes are hypersusceptible to geminivirus infection (Raja et al. 2008). Hypersusceptibility is characterized by enhanced symptoms (increased stunting and curling). Because a Pol IV/V double mutant (nrpd/e2) was used in earlier studies, we asked whether Pol IV and
Pol V, and associated remodeling enzymes, are individually involved in defense against geminiviruses.

Arabidopsis Pol IV (nrpd1), Pol V (nrpe1), Pol IV/V (nrpd/e2), clsy1, and drd1 mutant plants where inoculated with BCTV. The nrpd/e2 double mutant is deficient for the second largest subunit that is shared by Pol IV and Pol V. We also generated a second Pol IV/V double mutant lacking the catalytic subunits of both polymerases by crossing nrpd1 and nrpe1 plants to generate an nrpd1/e1 mutant line. In all cases, the null mutant lines used in these studies were homozygous, and none exhibited phenotypes that complicated the observation of symptom development. Virus inoculation was performed under conditions that normally produce mild symptoms to allow for symptom enhancement.

For each treatment, 32 were inoculated and monitored for symptoms of viral disease 21 days post-inoculation. All of the infected mutants proved hypersusceptible to BCTV compared to wild type plants (Fig. 4.1A). Disease ratings ranged from 3 to 4, on a subjective scale where 1 indicates symptoms typically observed with BCTV-infected wild type plants, and 4 indicates very severe leaf curling, deformation of floral tissues, and stunting. Symptoms were most severe in the nrpd1/e1 plants, which lack the catalytic subunits of Pol IV and Pol V.

Quantitative PCR (qPCR) was performed to assess viral DNA levels in infected plants using 18S rDNA control primers and BCTV IR primers to amplify the viral strand. Consistent with disease phenotypes, all of the infected mutant plants contained elevated levels of viral DNA (Fig. 4.1B). Levels in Pol IV (nrpd1) mutants were moderately increased (>1.5-fold), whereas all other mutants contained at least three-fold more viral
DNA than wild type plants. Viral DNA levels were highest (more than 10-fold greater than wild type) in Pol IV/V (nrpd1/e1) mutant plants. Surprisingly, only about a 3-fold increase was evident in the nrpd/e2 double mutant, a smaller increase than was observed in Pol V (nrpe1) mutant plants (~5-fold). These discrepancies suggest that the nrpd/e2 mutant might retain some Pol IV and/or Pol V function, possibly by substitution of the paralogous NRPB2 subunit from Pol II, or that the intact Pol IV/V catalytic subunits in nrpd/e2 plants interfere with other activities involved in viral DNA methylation. In any event, these results indicate that Pol IV, Pol V, CLSY1, and DRD1 participate in defense against geminiviruses, and suggest that Pol IV and V independently contribute to viral chromatin methylation.

4.3.2 Pol IV and Pol V amplify viral DNA methylation, but are not required for its establishment.

BCTV encodes an L2 protein (also known as C2) that suppresses methylation and TGS by interacting with adenosine kinase and inhibiting the methyl cycle that generates S-adenosyl methionine, an essential methyltransferase co-factor (Wang et al. 2003). Prior studies have shown that plants inoculated with L2-deficient BCTV (BCTV L2') exhibit a novel recovery phenotype (Hormuzdi and Bisaro 1995). Recovery occurs when shoots that arise after the establishment of a systemic infection show symptom remission and contain small amounts of viral DNA. Experimentally, plants inoculated with BCTV or BCTV L2' show similar symptoms of systemic infection by ~21 days. Infected tissues (called primary infected tissue) are harvested, and new shoots (called secondary infected tissue) are allowed to develop from axillary meristems. Secondary tissues of *Nicotiana*
*benthamiana* and Arabidopsis plants rarely recover from infection with wild type BCTV, but invariably recover from infection with BCTV L2'. Viral DNA in recovered tissue is hypermethylated in the IR. In addition, *ago4* mutants cannot recover from BCTV L2' and cannot hypermethylate the IR, confirming that the RdDM pathway is required for recovery (Raja et al. 2008). We have found recovery analysis to be a sensitive method for the identification and study of methylation pathway components (Raja et al., unpublished). Thus we assessed the recovery phenotypes of Pol IV (*nrpd1*), Pol V (*nrpe1*), Pol IV/V (*nrpd/e2* and *nrpd1/e1*), *clsy1*, and *drd1* mutants.

Following infection of wild type plants with BCTV L2', all secondary shoots that developed after the primary infection showed symptom remission (recovery) as expected, but Pol V (*nrpe1*) and Pol IV/V double mutants (both *nrpd/e2* and *nrpd1/e1*) failed to recover and secondary shoots continued to show severe symptoms. Interestingly, both Pol IV (*nrpd1*) and *clsy1* mutants displayed a novel, delayed recovery phenotype: all secondary shoots that developed after the establishment of infection at first displayed severe disease symptoms, but over time showed symptom remission (Fig. 4.2A). The similar delayed recovery phenotypes are consistent with Pol IV and CLSY1 acting cooperatively in the methylation pathway. By contrast, *drd1* mutants showed a partial recovery phenotype that was intermediate between recovered wild type plants and non-recovered *nrpe1* mutants. This suggests that while DRD1 may be involved, other factors are also able to recruit Pol V to the viral genome (Fig. 4.2A). These results underscore the sensitivity of the recovery system for methylation pathway analysis, and further indicate that while Pol V plays a unique and essential role in recovery, Pol IV function can be redundantly provided by another polymerase (e.g. Pol II).
Because recovery is typically accompanied by hypermethylation of the IR, we used bisulfite sequencing to assess BCTV L2- methylation status in the various mutants. In order to fully characterize the delayed recovery phenotype observed in Pol IV (nrpd1) and clsy1 mutants, non-recovered and recovered tissues were separately collected and pooled (as indicated by the white line in Fig. 4.2A). DNA was isolated and treated with sodium bisulfite, PCR was performed to amplify the viral strand of the BCTV IR, and PCR products were cloned and sequenced. Cytosine methylation in plants occurs in all sequence contexts: CG, CNG, and CHH (where H = A, T, or C), and our previous studies have shown that individual BCTV genomes are either mostly methylated in all sequence contexts, or mostly unmethylated. Thus, to ensure accuracy, rigorous conversion controls are included in all bisulfite experiments, and data were accepted only when control DNAs are completely converted (see Materials and methods).

Bisulfite sequencing of viral DNA obtained from secondary recovered tissue of wild type plants, and partially recovered tissue from drd1 plants, showed that the IR was hypermethylated, with most viral genomes methylated at most cytosine residues (>60% of cytosine residues methylated) (Fig. 4.2B and 4.2C). The high density of CHH methylation observed is indicative of de novo methylation. Similar hypermethylation (>60%) was also detected in the recovered portions of nrpd1 (Pol IV) and clsy1 shoots, but not in non-recovered portions of the same shoots (~45% of cytosine residues methylated), confirming that IR hypermethylation is a unique characteristic of recovered tissue. The nrpe1 (Pol V) and nrpd/e1 (Pol IV/V) mutant plants, which did not recover, also failed to hypermethylate the IR (~40 to 45% of cytosine residues methylated) (Fig. 4.2B and 4.2C).
While differences in methylation levels between recovered and non-recovered samples were statistically significant, an additional analysis was performed in which all bisulfite clones from BCTV-infected mutant plants classified as non-recovered (Col-0 background, obtained from this experiment and others, n = 72) were grouped together and mean cytosine methylation values determined. This process was repeated for all bisulfite clones from recovered wild type Col-0 or mutant plants (n = 84). The analysis confirmed that methylation levels in the recovered and non-recovered groups were significantly different with 99% confidence, with minimal variation within groups (Fig. B.8). Thus our data are not greatly impacted by sampling error.

The results of experiments presented in Fig. 4.2 allow several important conclusions concerning Pol IV and Pol V. First, the distinct phenotypes genetically confirm the independent involvement of both polymerases in viral genome methylation, and that Pol IV acts upstream of Pol V. In addition, because viral IR hypermethylation either does not occur (nrpe1) or is delayed (nrpd1), these polymerases act primarily to amplify methylation. Finally, because methylation nevertheless occurs in nrpd1 (Pol IV), nrpe1 (Pol V), and nrpd/e2 (Pol IV/V) mutants, neither polymerase is required to establish viral genome methylation.

4.3.3 Pol II, IV and V differentially associate with viral DNA.

Chromatin immunoprecipitation (ChIP) was performed to investigate polymerase occupancy in non-recovered and recovered tissues. To determine whether and where Pol IV and Pol V associate with viral DNA, transgenic Arabidopsis lines containing FLAG epitope-tagged Pol II, Pol IV, or Pol V catalytic subunits were obtained (Pontes et al.
The transgenes in these lines complement corresponding null mutations in the respective polymerase catalytic subunits. Pol II was included as a positive control for polymerase association with BCTV because it is expected to occupy both the promoter-containing IR and coding regions. Plants were inoculated with BCTV or BCTV L2', and only those in which the FLAG epitope-tagged transgene complemented the mutant phenotype (i.e. plants did not show hypersusceptibility and recovered from BCTV L2') were used in these experiments (Fig. B.9). Infected tissue was pooled from at least three plants, and three biological replicates of ChIP were performed with FLAG antibody (α-FLAG). Precipitated DNA was analyzed using both semi-quantitative PCR (sqPCR) and quantitative real-time PCR (qPCR), each of which were performed at least three times. Representative experiments are presented in Fig. 4.3. Actin served as a positive control for Pol II, while IGN5 was a positive control for Pol V (Wierzbicki et al. 2009). That Pol II and Pol IV did not associate with IGN5, and Pol V did not associate with actin, verified the specificity of ChIP (Fig. 4.3). Primers to amplify the BCTV IR, and two sets of primers for the downstream coat protein coding region (CP1 and CP2), were used to monitor Pol II, Pol IV, and Pol V association with viral DNA. The IR and CP2 are separated by more than 1 kbp, and the diagram in Fig. 4.3 shows the relative locations of the amplified regions. Prior to ChIP, DNA was sheared by sonication to an average length of ~600 bp to ensure that observed polymerase associations with the IR and the CP region primarily reflect independent events.

DNA was precipitated with α-FLAG from transgenic plants expressing FLAG-Pol II, FLAG-Pol IV, or FLAG-Pol V, and similar results were obtained when DNA was analyzed by sqPCR or qPCR. We first examined BCTV genomes during primary
infection, when methylation is mostly confined to promoter regions (Yang et al. 2011), and found that all three polymerases were associated with the IR. By contrast, significant signal was observed only with Pol II when CP coding region primers were employed (Fig. 4.3A and 4.3C). Polymerase occupancy was then analyzed in secondary recovered tissue from plants infected with BCTV L2’, when the IR is hypermethylated. Again, the presence of all three polymerases was evident in the IR although, as might be expected, Pol II signal was reduced compared to primary tissue. However, Pol IV and Pol IV were now seen to occupy the CP region, suggesting that methylation spreads away from the IR during recovery (4.3B and 4.3C).

4.3.4 Pol IV and V association with a coding region correlates with methylation spread.

We asked if Pol IV and Pol V occupancy of the CP coding region observed during recovery in ChIP experiments correlated with the spread of DNA methylation, and whether these polymerases are required for spread.

Cytosine methylation in the CP2 region was assessed in primary tissue from wild type and Pol IV/V (nrpd/e2) mutant plants infected with BCTV, and in secondary tissue of recovered wild type plants and non-recovered Pol IV/V (nrpd/e2) mutant plants. Infected tissue was harvested from a pool of three plants, and DNA was isolated and treated with sodium bisulfite to convert unmethylated cytosines to uracil. PCR with primers designed to amplify CP2 was performed, and products were cloned and sequenced.

Analysis of wild type plants confirmed that the BCTV CP coding region is sparsely methylated during primary infection (~5% of cytosine residues methylated). However, in
secondary recovered tissue CP methylation was dramatically increased to approximately 30% (Fig. 4.4). CP methylation levels were similarly low in primary shoots from Pol IV/V (nrpd/e2) mutant plants (~5%), but were also increased to ~15% in non-recovered secondary tissue. Thus coding regions are not significant methylation targets during primary infection, but methylation is increased in coding regions in secondary shoots independent of Pol IV and Pol V activity.

Considered together, the results of ChIP and bisulfite sequencing experiments indicate that Pol IV and Pol V preferentially occupy the methylated IR during primary infection, and spread along with DNA methylation into coding regions in secondary infected tissues. However, while Pol IV and V are needed to amplify methylation in the coding region, they are not required to initiate methylation or its spread.

4.3.5 Pol IV and Pol V are not required for biogenesis of virus-derived siRNAs.

It is well established that Pol IV and Pol V are involved in producing siRNAs. Pol IV plays a major role in 24 nt siRNA biogenesis, since this size class is lost at the vast majority of loci in Pol IV mutants (Huettel et al. 2006, Zhang et al. 2007). Pol V, on the other hand, is additionally required for 24 nt siRNA accumulation at a some loci (Pontier et al. 2005). Previous studies have shown that abundant siRNAs of all typical size classes (21, 22, and 24 nt), spanning virtually the entire viral genome, are generated during geminivirus infections. However, in keeping with the importance of methylation in antiviral defense, the 24 nt size class is the most abundant. To further characterize the
roles of Pol IV and V in antiviral defense, we set out to determine whether Pol IV and Pol V are required for the accumulation of geminivirus-derived siRNAs.

Total RNA was isolated from primary tissue of BCTV-infected wild type, Pol IV (nrpd1), Pol V (nrpe1), Pol IV/V (nrpd/e2 and nrpd1/e1), clsy1, and drd1 mutant plants. Again, tissue was harvested from a pool of at least three plants to minimize plant-to-plant variation. The siRNAs were then visualized by Northern blot with probes specific for the viral IR. The blots were then stripped and re-probed with labeled oligonucleotides specific for the CP coding region. The blot shown is representative of three individual experiments (Fig. 4.5A). Total small RNA levels from all three blots were quantitated for each polymerase mutant and normalized to the rRNA loading control (Fig. 4.5B).

All of the mutants showed reduced BCTV IR-specific siRNAs, and the 24 nt siRNA size class was most affected (Fig. 4.5). Because virus titers are increased in all of these mutants (Fig. 4.1), the diminished siRNA levels cannot be attributed to limiting DNA template. That significant reductions were observed in nrpd1 and nrpe1 mutants indicates that both polymerases are important for the accumulation of siRNA derived from the BCTV IR, and the similar reductions observed with clsy1 and drd1 mutants further associate these remodeling enzymes with Pol IV and Pol V. Surprisingly, siRNA levels in both of the Pol IV/V double mutants (nrpd/e2 and nrpd1/e1) were not as greatly impacted as the corresponding single mutants. Possible explanations might be that a double mutant could have some residual function (possible for nrpd/e2 but not nrpd1/e1), or that siRNA levels appear elevated as a consequence of high virus titer (more likely for nrpd1/e1 than nrpd/e2, see Fig. 4.1). Alternatively, the presence of either Pol IV or Pol
V alone might somehow interfere with the activity of a functionally redundant activity (e.g. Pol II).

By contrast, when blots were subsequently stripped and re-probed with CP-derived oligonucleotides, no reductions in viral siRNA levels were observed in the mutant plants, indicating that Pol IV, Pol V, CLSY1, and DRD1 have little or no role in generating siRNAs derived from this coding region (Fig. 4.5). This is consistent with the results of ChIP experiments, which showed that Pol IV and V are not associated with the CP coding region in primary infected tissue (Fig. 4.3).

The results of these experiments demonstrate that Pol IV and V are not required to generate BCTV-derived 24 nt siRNAs associated with the methylation pathway, although they independently contribute to siRNA production in a methylated region that contains viral promoters (the IR). Furthermore, because Pol II is associated with both the BCTV IR and CP regions in primary infected tissue, it is likely that Pol II transcription plays an important role in the biogenesis of viral siRNA of all size classes.

4.3.6 Pol IV and V are required for the establishment and spread of repressive histone methylation.

DNA methylation generally correlates with the presence of H3K9me2, and we showed previously that both active (acetylated H3) and repressive (dimethylated H3K9) marks are present on histone H3 associated with the viral IR, in both primary and secondary infected tissue (Raja et al. 2008, Raja et al. 2010). The presence of both chromatin marks is consistent with the observation of mostly methylated and unmethylated genomes in infected plants, and suggests that populations of active and
repressed genomes are present in infected plants. Consequently, we evaluated the roles of Pol IV and Pol V in the establishment and spread of H3K9me2 on the BCTV genome.

Chromatin modifications were analyzed by ChIP using tissue harvested from pools of at least three plants infected with BCTV (for analysis in primary infected tissue) or BCTV L2- (for analysis in secondary infected tissue). Antibodies to histone H3 acetylated at lysine 9 and 14 (α-H3Ac) or dimethylated H3K9 (α-H3K9me2) were used to assess the association of these marks with the viral genome. Actin served as a positive control for H3Ac and a negative control for H3K9me2, while transposon Ta3 was a negative control for H3Ac and a positive control for H3K9me2. Three independent experiments were performed with three replicates each, and data from a representative experiment are presented in Fig. 4.6.

qPCR was executed with IR and CP2 primers to detect and measure precipitated viral DNA. When ChIP was performed using α-H3K9me2 with extracts from primary infected tissue, this repressive mark was again found to be associated with the IR, but very little was detected in the CP coding region. However, H3K9me2 was associated with both regions in secondary recovered tissue (Fig. 4.6A). The appearance of H3K9me2 in the CP region correlates with the spread of Pol IV and Pol V during recovery. By contrast, ChIP using H3Ac antibody showed that this mark was associated with the BCTV IR and the CP coding region in both primary and secondary infected tissue of wild type plants (Fig. 6E).

ChIP was then repeated with tissue harvested from plants deficient for Pol IV (nrpd1), Pol V (nrpe1), and Pol IV/V (both nrpd/e2 and nrpd1/e1) and infected with BCTV or BCTV L2-. Remarkably, little or no dimethylated H3K9 was detected at either
the IR or the CP coding region in any of the mutants in primary or secondary infected tissues (Fig. 4.6B-4.6D). This is in sharp contrast to the endogenous Ta3 locus, where detection of relatively abundant H3K9me2, confirms that the relevant histone methyltransferase activities were not adversely affected by the polymerase mutations. Additionally, using the same extracts, ChIP with α-H3Ac found this positive mark on both the IR and CP coding region during primary and secondary infections in the mutant plants, confirming that histone H3 continues to associate with viral genomes (Fig. 6F-6H). These results allow us to conclude that Pol IV and Pol V are individually essential for de novo deposition and spread of H3K9me2 on the viral genome, but suggest they are not required to maintain H3K9me2 established at endogenous sequences.

4.4 DISCUSSION

The studies presented here examined the roles of RNA polymerases IV and V, and associated remodeling enzymes, in repressive methylation of geminivirus chromatin, which constitutes a powerful antiviral defense. In initial experiments, the hypersensitivity of mutant plants lacking Pol IV, Pol V, CLSY1, and DRD1 suggested individual roles for these enzymes in antiviral defense, and the more sensitive recovery assay using BCTV L2\(^-\) mutant virus uncovered some interesting phenotypes. That nrpe1 mutants are unable to recover or hypermethylate the viral IR following infection with BCTV L2\(^-\) indicates an essential role for Pol V in antiviral RdDM. However, while DRD1 can recruit Pol V to some loci (Wierzbicki et al. 2008), the intermediate, partial recovery and IR hypermethylation observed in drdl plants suggest that other factors can
redundantly recruit Pol V to the viral genome. The novel delayed recovery displayed by
nrpd1 plants indicates that Pol IV also has an important role in antiviral DNA methylation,
although another activity (likely Pol II) can eventually compensate for its absence. The
strikingly similar delayed recovery observed with clsy1 and nrpd1 mutants additionally
supports previous work showing that CLSY1 functions with Pol IV in the production of
24 nt siRNAs (Smith et al. 2007). In all experiments, viral IR methylation levels
correlated with disease outcomes and, consistent with previous studies using our bisulfite
treatment protocol and BCTV IR primer set, confirmed that recovery phenotypes are
associated with >60% cytosine methylation, while ~40 to 45% of cytosines are
methylated in non-recovered secondary tissues. Because Pol IV (nrpd1), Pol V (nrpe1),
and Pol IV/V (nrpd/e2) mutants retained this 40-45% base level of DNA methylation in
non-recovered tissues, we conclude that these polymerases are required to amplify
antiviral RdDM, but are not required for its initiation. Assuming that transcription is an
initial event in RdDM, another polymerase activity (likely Pol II) must be responsible for
the observed base level methylation.

Cytosine methylation profiles showed that individual BCTV genomes are either
mostly methylated or unmethylated, with a greater proportion of methylated genomes in
recovered compared to non-recovered tissue. The largely methylated viral genomes
contain a high level of non-symmetrical CHH methylation (up to 60%), which cannot be
recognized by maintenance methyltransferases and hence is a signature of de novo
methylation. This pattern also suggests that during the interval in which the methylation
apparatus interacts with and releases a DNA segment, nearly all cytosine residues can
become methylated. The extent of these methylation tracts was not directly investigated
in this study. However, we previously found that cytosine methylation is largely confined to regions encompassing and adjacent to promoters in primary infected tissue (Yang et al. 2011).

It should be pointed out that geminivirus DNAs exist in multiple forms in infected cells. Circular ssDNA, the encapsidated form of the viral genome, cannot be methylated. The relative abundance of viral ssDNA and dsDNA forms varies between plants, but as ssDNA is not enriched in methylation-deficient mutants (our unpublished observations), reduced methylation levels cannot be ascribed to increased levels of ssDNA. Both circular dsDNA replicative forms (which form minichromosomes) and heterogeneous linear dsDNA forms (presumably replication by-products) can be recognized by methyltransferases (Jeske et al. 2001, Pilartz and Jeske 2003), and it has been reported that the heterogeneous linear forms constitute the bulk of methylated geminivirus DNA (Paprotka et al. 2011). Our bisulfite sequencing protocol does not discriminate on the basis of DNA conformation. Nevertheless, we reproducibly observe reduced IR methylation in methylation pathway mutants (e.g. nrpd1 and nrpe1 in this study), and methylation levels are highest in asymptomatic, recovered tissue infected with a virus (BCTV L2) that is incapable of suppressing methylation. Thus we argue that our results accurately reflect the methylation levels of aggregate viral DNA populations.

Using anti-FLAG immunoprecipitation with extracts from transgenic plants expressing FLAG-Pol II, -Pol IV, and -Pol V transgenes we found that all three polymerases occupy the promoter containing viral IR, while only Pol II associates with the sparsely methylated CP coding region during primary BCTV infection. However, in recovered secondary tissue, Pol IV and V also associate with the CP open reading frame,
and this correlates with increased cytosine methylation. This again indicates that non-coding Pol IV and Pol V transcripts are primarily involved in amplifying pre-existing viral DNA methylation. Why methylation should spread into coding regions during recovery is not clear. However, in support of these observations, we found that in primary infected tissue of plants deficient for Pol IV, Pol V, or both polymerases, IR-derived siRNAs are reduced but not eliminated, while siRNAs originating from the CP coding region are not affected. Further, because 21-24 nt small RNAs derived from both the IR and CP regions were detected in Pol IV/V-deficient plants, Pol II transcription can clearly support the production of viral siRNAs of all size classes, including the methylation- and AGO4-associated 24 nt siRNAs. This is consistent with Pol II - AGO4 interaction (Zheng et al. 2009), but at odds with studies showing that the bulk of cellular 24 nt siRNAs are Pol IV-dependent (Onodera et al. 2005, Huettel et al. 2006, Zhang et al. 2007). We suggest this is because most 24 nt siRNAs are generated to maintain silencing of transposons and repeat elements, whereas the geminivirus system reflects de novo events. Additionally, the presence of abundant 24 nt siRNAs derived from the viral CP coding region, which is very lightly methylated in primary infected tissue, indicates that methylation levels can be regulated downstream of siRNA biogenesis.

Remarkably, the results of ChIP experiments showed that Pol IV and Pol V are individually required for the deposition of repressive H3K9me2 at both the BCTV IR and the CP coding region. Thus de novo cytosine methylation, likely carried out by DRM1/2, can occur in all sequence contexts in the absence of H3K9me2. However, in contrast to our finding that disruption of Pol IV or Pol V abolishes de novo H3K9 methylation of viral genome-associated nucleosomes, we observed little or no impact on the
maintenance of H3K9me2 associated with the endogenous Ta3 transposon. Thus Pol IV and V are not needed for H3K9me2 maintenance at this locus, which may be attributed to the linked activities of H3K9 methyltransferases such as KRYPTONITE/SUVH4 and the cytosine methyltransferase CHROMOMETHYLASE3 (CMT3), which is responsible for maintaining CNG methylation (Jackson et al. 2002, Lindroth et al. 2004). A recent study has elegantly elucidated the structural basis for CMT3 interaction with H3K9me2 peptides (Du et al. 2012). Thus our results are consistent with essential roles for Pol IV and V in initiating, but not maintaining, histone H3K9 methylation.

Together, our data suggest a model based on the role of RNA polymerase II in S. pombe. We propose that in Arabidopsis, Pol II initially generates non-coding transcripts leading to siRNAs that establish de novo viral DNA methylation. However, this process is self-limiting because Pol II transcription is likely inhibited by dense DNA methylation. Pol IV and V are subsequently recruited to methylated DNA where they initiate and spread dimethyl H3K9 modifications and reinforce, amplify, and spread DNA methylation. An attractive feature of this model is that it provides a ready explanation for initial targeting of methylation to viral promoters.

The delayed recovery phenotypes observed in nrpd1 and clsy1 mutants provide support for the model by suggesting a means to bypass Pol IV activity. Pol II is capable of producing viral siRNAs of all size classes (this study) and can engage AGO4 (Zheng et al. 2009). As AGO4-RISC complexes have slicer activity, they can both recruit methyltransferases and cleave scaffolding transcripts to generate RDR substrates for synthesis of homologous dsRNA that can be processed to secondary siRNAs (Qi et al. 2006). This amplification loop could eventually become sufficient to support viral IR
hypermethylation necessary for recovery. Recent reports from several groups provide additional support for the model. These have found that non-canonical pathways with similarities to post-transcriptional gene silencing, which is usually associated with Pol II, are capable of initiating RdDM in Arabidopsis (Garcia et al. 2012, Pontier et al. 2012, Wu et al. 2012, Nuthikattu et al. 2013). In addition, Pol II has been shown to produce scaffold transcripts that can recruit AGO4, Pol IV, and Pol V to cellular loci (Zheng et al. 2009), and to influence the cellular methylome independent of Pol IV and V (Stroud et al. 2013). Our studies clearly implicate Pol II in transcription leading to RdDM of geminivirus genomes and we favor the idea that, similar to the ancestral S. pombe system, non-coding Pol II transcription initiates antiviral DNA methylation in Arabidopsis. However, we cannot exclude the possibility that read-through transcription of the circular viral genome generates dsRNA that serves directly as DCL substrate, as was proposed in a recent study (Aregger et al. 2012). Regardless of the nature of these putative Pol II transcripts, Pol IV and Pol V may be targeted by them or by resulting DNA methylation, and appear specialized to amplify and spread cytosine methylation and to initiate H3K9 methylation.

In conclusion, the studies presented here highlight the utility of geminiviruses as models for de novo methylation of foreign DNA. We have demonstrated that: 1) Pol IV and Pol V are dispensable for viral DNA methylation, but are required for its amplification and spread; 2) Pol IV and Pol V play independent and essential roles in de novo H3K9 methylation of viral genome-associated nucleosomes, but are not necessary to maintain H3K9 methylation at an endogenous transposon locus; 3) Pol IV and V are not required to generate virus-derived siRNAs, suggesting that Pol II transcription can
generate viral siRNAs of all size classes. Current studies are focused on the role of Pol II in antiviral RdDM and histone methylation, and on features that recruit Pol IV and Pol V to the viral genome.

Fig. 4.1 Plants deficient for Pol IV, Pol V, Pol IV/V, CLSY1, and DRD1 are hypersusceptible to geminivirus infection. (A) Photographs taken under a dissecting microscope show floral heads from infected wild type (Col-0) plants, and clsy1, nrpd1 (Pol IV), drd1, nrpe1 (Pol V), nrpd/e2 (Pol IV/V), and nrpd1/e1 (Pol IV/V) mutant plants. Photographs of primary infected tissues were taken ~21 days post-inoculation, and are representative of 32 plants per treatment. Note increased floral deformation in mutant compared to wild type plants. (B) Viral DNA levels in infected plants were quantified by qPCR using BCTV IR primers and normalized to 18S genomic DNA. Viral DNA in wild type plants was set to 1. Samples consisted of symptomatic tissues collected from pools of three plants to minimize plant-to-plant variation, and DNA from at least three individual samples was analyzed in triplicate for each treatment. Asterisks
indicate that viral DNA levels were significantly increased in mutant plants compared to wild type plants with 95% (*) or 99% (**) confidence, as determined by Student's t test.
Fig. 4.2 Pol IV and Pol V are not required for viral genome methylation. (A) Dissecting microscope photographs show floral heads from infected wild type (Col-0) plants, and clsy1, nrpdl (Pol IV), nrpe1 (Pol V), nrpd/e2 (Pol IV/V), nrpd/e1 (Pol IV/V), and drd1 mutant plants. Photographs of secondary infected tissues were taken approximately 14 days after harvest of primary infected tissue (~35 days post-inoculation). Photographs are representative of all secondary shoots from all of the ~32 plants per treatment. Note the absence of symptoms indicating recovery in wild type plants, while Pol V (nrpe1), and Pol IV/V (nrpd/e2, and nrpd/e1) mutants do not recover and show severe symptoms. The Pol IV (nrpdl) and clsy1 mutants showed delayed recovery, and non-recovered (NR) and recovered (R) tissue was harvested as indicated (white line). The drd1 mutants have an intermediate partial recovery phenotype. (B) The BCTV L2 IR from secondary tissue was analyzed by bisulfite sequencing. In all cases, tissues were harvested from pools of three plants. Histograms compiled from data in (C) indicate the percentage of methylated cytosine residues in different sequence contexts in DNA obtained from secondary tissue of wild type and mutant plants. IR DNA isolated from recovered tissue is hypermethylated (>60% of cytosines methylated), while DNA from non-recovered tissue is relatively hypomethylated (40-45% of cytosines methylated). Asterisks indicate significant differences in total methylation between recovered and non-recovered tissues at the 95% (*) or 99% (**) confidence level, as determined by Student's t test. (C) Cytosine methylation profiles of sequenced clones obtained from wild type and mutant plants. Rows indicate individual clones (12 per treatment), and are organized from most to least methylated. Each circle represents a cytosine in the BCTV IR (42 total) with CG in red, CNG in blue, and CHH in green. A filled circle represents a methylated cytosine.
Fig. 4.3 Pol II, Pol IV and Pol V differentially associate with the viral genome. Representative ChIP experiments (of at least three independent trials) with extracts from BCTV-infected, transgenic plants expressing FLAG-Pol II, IV, and V catalytic subunits are shown. Extracts were obtained from pooled tissues of three plants, and DNA precipitated by α-FLAG was analyzed using both sqPCR and qPCR, each of which were performed at least three times. Primer sets to amplify the BCTV IR and two CP coding region primer sets (CP1 and CP2) were employed, as were control primers to amplify endogenous actin and IGN5. IgG was a background control. (A) PCR products (two-fold serial dilutions) generated with BCTV IR primers show that Pol II, IV and V occupy the viral IR, while CP1 primers suggest little or no Pol IV and V association with this region during primary infection. (B) Pol II, IV and V associate with both the BCTV IR and CP1 in secondary recovered tissue. (C) A representative qPCR experiment performed on precipitated DNA, using IR, CP2, or control primers. The graphs depict signal as % of input and support the semi-quantitative PCR results illustrated in (A) and
The diagram depicts the dsDNA replicative form of BCTV genome, and indicates the relative locations of IR, CP1, and CP2 regions. Open reading frames are indicated by filled arrows.

Fig. 4.4 Spread of DNA methylation during recovery correlates with the association of Pol IV and V. (A) Histograms compiled from data in (B) indicate the percentage of cytosine residues methylated in different sequence contexts in the BCTV CP1 region from primary and secondary tissues of wild type (Col-0) and Pol IV/V (nrpd/e2) mutants. The BCTV CP is sparsely methylated during primary infection, but accumulates methylation in secondary tissue. (B) Cytosine methylation profiles of sequenced clones obtained from wild type and mutant plants. Rows indicate individual clones (12 per treatment), and are organized from most to least methylated. Each circle represents a cytosine in the BCTV CP (46 total) with CG in red, CNG in blue, and CHH in green. A filled circle represents a methylated cytosine.
Fig. 4.5 Pol IV and Pol V are not required to generate BCTV-specific siRNAs. (A) A representative Northern blot of virus-specific small RNAs from BCTV-infected primary tissue is shown. IR-derived siRNAs are reduced in Pol IV (nrpd1), Pol V (nrpe1), Pol IV/V (nrpd/e2 and nrpd1/e1), clsy1, and drd1 mutant plants, while CP-derived siRNAs are unaffected. Also note that neither Pol IV or V is required to generate the methylation-associated 24 nt species. (B) Graph comparing relative siRNA levels in wild type plants and plants harboring the indicated mutations in Pol IV, Pol V, or both polymerases (normalized to the rRNA loading control). Data are from three independent experiments. Bars indicate standard error.
Fig. 4.6 Pol IV and V are required for the establishment and spread of dimethylated H3K9 histone marks. Each histogram represents technical triplicates of ChIP qPCR performed in biological triplicate with antibodies to dimethylated H3K9 using extracts from wild type, Pol IV (nrpd1), Pol V (nrpe1), and Pol IV/V (nrpd/e1) mutant plants infected with BCTV (primary infected tissue) or BCTV L2- (secondary infected tissue). The graphs show the ChIP signal as % input. Actin was a negative control and transposon Ta3 was a positive control for the presence of H3K9me2. 

(A) ChIP performed in wild type (Col-0) plants revealed that H3K9me2 is present on the viral
IR, while the CP1 primers suggest little or no H3K9me2 association with this region during primary infection (light shading). H3K9me2 associates with both the BCTV IR and CP1 regions in secondary recovered tissue (dark shading). (B-D) No H3K9me2 was associated with the viral genome in primary or secondary tissue from \textit{nrpd1}, \textit{nrpe1}, or \textit{nrpd/e2} mutants, while this repressive mark remains associated with the endogenous Ta3 locus.
CHAPTER 5

DISCUSSION

5.1 Analysis of AL2 suppressor activities and developmental changes in the role of ADK in the methyl cycle

5.1.1 AL2 transcription-dependent suppressor functions

Previous work in the laboratory found that AL2 and L2 reversed TGS and global DNA methylation by interacting with and inhibiting ADK. This work also showed that in addition to ADK inhibition, AL2 utilized a transcription-dependent mechanism to reverse TGS (Buchmann et al. 2009). Chapter 2 of this thesis set out to explore the range of the AL2 transcription-dependent suppressor function. Initial studies analyzed the ability of AL2, AL2_{1-114}, AL2-C33A, and L2 proteins, as well as ADK inhibition to impact systemic spread of the silencing signal. Only AL2 was able to block systemic spread, while AL2_{1-114}, L2, and ADK inhibition were completely unable to suppress spread. AL2-C33A, which has diminished transcription activation activity, was able to prevent the spread of silencing at reduced levels, supporting the observation that this is a transcription-dependent suppressor function of AL2. Subsequent studies used over-expression of AL2 and L2 to assess their ability to reverse established PTGS of a GFP transgene in vegetative and reproductive plants, defined as before and after transition to

158
flowering. Both AL2 and L2 were able to reverse established PTGS in vegetative plants, but only a transcriptionally competent AL2 was able to do this after the onset of flowering. We speculate that AL2 likely activates transcription of endogenous genes that are involved in the regulation of systemic spread of silencing and PTGS. Thus, AL2 reverses PTGS in vegetative tissue likely through inhibition of ADK, as well as activation of host genes. Conversely in reproductive plants, only AL2 transcription activation of endogenous genes is able to reverse PTGS (discussed further in section 5.1.3).

Genes induced by MYMV and ACMV AC2 have been identified by microarray analysis in Arabidopsis protoplasts where >55 genes had significantly increased activation following transient AC2 expression (Trinks et al. 2005). One of these genes was 3'-5' Werner exonuclease-like (WEL-1) protein that contains a non-canonical DEDDy exonuclease domain. Interestingly, another canonical DEDDy domain containing protein, Werner-exonuclease (WEX), has been shown to be required for PTGS (Glavoz et al. 2003). Thus, it is possible that WEL-1 is activated by AC2/AL2 to negatively regulate PTGS by an as yet undetermined mechanism. Both WEX and WEL-1 belong to larger gene families and more work will be required to determine the role of these exonucleases in the regulation of both PTGS and TGS. While the WEL and WEX family genes appear to be the intriguing candidates, the other 55 genes induced by AC2 should not be forgotten because they may provide interesting insight into geminivirus biology and the regulation of host defenses.
5.1.2 Discovery of a transcription- and ADK-independent AL2 RNA silencing suppressor activity

Since AL2, but not L2, was able to reverse PTGS in vegetative and reproductive plants, it seemed obvious to assess TGS reversal in these developmental stages. Geminivirus infection mimicked the PTGS reversal in that only the AL2 containing CaLCuV and TGMV were able to reverse TGS in reproductive plants. Interestingly, although BCTV was able to inhibit ADK activity in both developmental tissues, this was only sufficient to reverse TGS in vegetative plants. Thus, the ability of CaLCuV and TGMV to reverse TGS in reproductive plants appears to be independent of ADK inhibition.

To determine the role of transcription activation in this reproductive TGS reversal the AL2, AL2\textsubscript{1-14}, and AL2-C33A were expressed from PVX viral vectors. Surprisingly, all of the AL2 containing PVX constructs were able to reverse TGS in both vegetative and reproductive plants, while PVX::L2 could only reverse TGS in vegetative plants. Therefore, TGS reversal by AL2 in reproductive tissue appears unrelated to its ability to activate transcription (required to reverse PTGS and block the spread of silencing) and also is independent of ADK inhibition. Therefore, a third uncharacterized suppressor function of AL2, beyond transcription activation and ADK inhibition, was identified by this study. Figure 5.1 models the previously characterized and newly identified suppressor functions of AL2.

Currently, the mechanism of this novel suppressor activity is unknown, but one could speculate that AL2, but not L2, could interact with and inhibit components of the TGS/RdDM pathway. Many viral silencing suppressors have been shown to directly
interact with and inhibit pathway components (reviewed in Bisaro 2006). Preliminary bimolecular fluorescence complementation data suggests that AL2 can interact with DCL3, DRB3, and AGO4 (P. Raja and J. Ostler, unpublished data). Whether these interactions are direct, indirect, or functionally important has not yet been tested. Furthermore, it will be important to test the interaction of these proteins with L2, AL2\textsubscript{114}, and AL2-C33A. The AL2-C33A mutant may be particularly interesting because although it sustained GFP expression, visual and Northern blot data suggested that reproductive plants had reduced TGS reversal. The AL2-C33A mutation lies within a conserved CCHC zinc-finger domain and such regions are frequently involved in protein-protein interactions. Thus, AL2-C33A and other CCHC mutants may be useful tools for characterizing the newly identified transcription-independent TGS suppression mechanism.

5.1.3 Changes in the role of ADK in the methyl cycle of reproductive plants

The observation that BCTV inhibits ADK activity but cannot reverse TGS in reproductive plants suggested that the role of ADK and/or the methyl cycle changes after flowering. To test this hypothesis, TRV viral vectors were used to knock-down ADK, SAHH, and MET1 enzymes in vegetative and reproductive plants. Although inhibition of ADK was verified, TRV::ADK was unable to reverse TGS after flowering, indicating that the role of ADK in the methyl cycle changes throughout development. Further analysis found that reproductive plants methylate DNA independent of ADK activity. Thus, the methyl cycle is able to function without ADK in reproductive \textit{N. benthamiana} plants.
In order for the methyl cycle to efficiently produce the methyl donor, SAM, adenosine must be salvaged. In plants, only two enzymes have been shown to have adenosine salvage activity, ADK and adenosine nucleosidases (AN). ADK converts adenosine directly to AMP, while ANs require the concerted action of adenine phosphoribosyltransferase (APT) to produce AMP. Work in *Avicennia marina* (mangrove) found that both of these pathways are active, thus it is reasonable to suspect that this could also be the case in *N. benthamiana* (Fig. 5.1). There are five nucleoside hydrolase genes in Arabidopsis. Future work will be required to identify these candidates in *N. benthamiana* and to test the nucleoside specificity, which would be required to identify the potential AN(s) that could function with or support the methyl cycle in reproductive plants.

5.2 Geminiviruses as models to examine host RdDM pathway components

5.2.1 Identification of DRB3 as an RdDM pathway component using the geminivirus model system

Previous work in the lab found that methylation deficient plants were hypersusceptible to geminiviruses and unable to recover from the infection. Loss of recovery in an AGO4 mutant was associated with hypomethylation of the viral intergenic region (IR) (Raja et al. 2008). In Chapter 3, the recovery phenomenon was demonstrated to be an extremely sensitive assay for identification of methylation pathway components. DCL3 is a well-known RdDM pathway component, but shares considerable redundancy
with DCL2 and DCL4. While *dcl2, dcl3, and dcl4* individual mutants showed hypersusceptible phenotypes, only *dcl3* mutants were unable to recover and methylate the viral IR. Each DCL partners with a DRB to process or load siRNAs into RISC complexes, but the DRB that functions with DCL3 in RdDM had yet to be identified. Thus, the sensitive geminivirus recovery system was used to determine which DRB genetically and physically partnered with DCL3 in DNA methylation.

A panel of *drb* mutants (*drb2-5*) was tested for hypersusceptibility and recovery from geminivirus infection. Plants deficient for *drb3* appeared to be the most susceptible to CaLCuV and BCTV and only *drb3* mutants were unable to recover from BCTV L2-infection. Bisulfite analysis found that both of these phenotypes correlated with decreased DNA methylation of the viral IR, genetically linking DRB3 with the RdDM pathway. DRB3 was then found to interact with DCL3 and AGO4 using bimolecular fluorescence complementation (BiFC) and co-immunoprecipitation assays. Together, these studies demonstrated physical associations between DRB3:DCL3 and DRB3:AGO4 that are consistent with their functional cooperation in the methylation pathway. DRB proteins are known to assist DCLs in either siRNA biogenesis or siRNA loading into AGO. Thus, viral 24 nt siRNAs, associated with the DNA methylation pathway, were analyzed in RNA extracts from *dcl3, drb3,* and *ago4* plants infected with BCTV. While *dcl3* mutants lost 24 nt siRNAs as expected, *drb3* and *ago4* mutants showed little to no change in siRNAs of all size classes. These results suggest that neither DRB3 nor AGO4 are required for the biogenesis or accumulation of geminivirus siRNAs.

These studies did not determine the function of DRB3 in the RdDM pathway, but the BiFC and siRNA data provide information for us to speculate a role for DRB3. The
BiFC studies showed that DRB3 interacts with DCL3 and AGO4 in two different subnuclear bodies and siRNAs continue to accumulate in the absence of DRB3. Therefore, we speculate that DRB3 carries DCL3-derived siRNAs and loads them into AGO4 containing silencing complexes. This hypothesis can be tested by immunoprecipitating AGO4 and determining the association of virus-derived 24 nt siRNAs in WT and drb3 extracts. Altogether, this work has not only laid the groundwork for studying the function of DRB3, but also highlights the sensitivity of the geminivirus model system for studying methylation and epigenetic regulation of cellular chromatin.

5.2.2 Geminiviruses as models to study Pol IV and Pol V in RdDM

In plants, Pol IV and Pol V have established roles for producing siRNAs and transcribing transposon DNA for silencing, respectively. Previous work in our lab has found that the pathway targets both geminivirus DNA and transposons for repressive DNA and histone methylation, but it is unknown how RdDM recognizes and targets DNA for silencing. Thus, in Chapter 4 of this thesis we set out to determine if geminiviruses could serve as models to study how Pol IV and Pol V target DNA for RdDM. To begin, *nrpd1* (*pol iv* mutant), *clsy1* (a putative chromatin remodeler associated with Pol IV), *nrpe1* (*pol v* mutant), *drd1* (a putative chromatin remodeler associated with Pol V), *nrpe2* (a *pol iv/v* double mutant deficient in a shared subunit), and *nrpd/e1* (deficient in catalytic subunits of both a *pol iv* and *pol v*) were tested for hypersusceptibility to BCTV and recovery from BCTV L2- infection. All of the mutants examined showed a hypersusceptibility phenotype and this was associated with an increase in viral titers.
Each of these mutants was then tested in the recovery assay. The only plants able to recover from BCTV L2- infection in a manner similar to wild type plants were *drd1* mutants, suggesting that other chromatin remodelers can recruit Pol V to viral DNA. However *clsyl* and *nrpd1* did not recover, but showed an interesting phenotype. Initially these plants showed viral symptoms, but over time they were able to recover. This phenotype was termed delayed recovery and the non-recovered symptomatic and recovered asymptomatic tissues were studied separately. These phenotypes suggested that CLSY1 and Pol IV work together in the pathway where they play an important role, but likely have redundant functions with another polymerase (likely Pol II). Finally, plants deficient for *nrpe1, nrpd/e2*, and *nrpd/e1* were completely unable to recover. Thus, Pol IV and Pol V play important and essential roles, respectively, in defense against geminiviruses. Moreover, this is the first time that *nrpd1* and *nrpe1* have shown phenotypic differences, highlighting the sensitivity and discrimination of the geminivirus model system.

To begin to determine how Pol IV and Pol V target DNA for methylation, ChIP analyses were performed to detect Pol II, Pol IV and Pol V association with viral DNA. Transgenic plants expressing Pol II-, Pol IV-, or Pol V-FLAG tagged proteins were a kind gift from Dr. Craig Pikaard. ChIP was performed with an anti-FLAG antibody followed by semi-quantitative and quantitative PCR with primers for the BCTV IR, a BCTV coat protein coding region (CP), as well as endogenous control genes Actin and IGN5. Pol II also served as a positive control since Pol II is known to associate with and transcribe viral DNA. During the primary infection all three polymerases were associated with the BCTV IR, while only Pol II was detected on the BCTV CP region,
consistent with the ability of Pol II to associate with viral promoters and transcribe viral mRNAs. Interestingly, in secondary recovered tissues all three polymerases were associated with both the BCTV IR and CP regions. Thus, Pol IV and Pol V appear to localize to the BCTV IR and later during infection spread to other regions of the viral genome. Future ChIP analysis will be required to determine the specific association of Pol IV and Pol V with promoter regions, as well as the extent of their spread. Altogether the phenotypic observations and ChIP data indicate that Pol IV and Pol V play important roles in geminivirus defense and specifically interact with geminivirus DNA. Therefore, the geminivirus model system will be very useful in studying how these polymerases recognize and target DNA for RdDM.

5.2.3 Pol IV and Pol V are not required for establishment of cytosine methylation, but are needed for its spread

Along with the phenotypic recovery analysis, Chapter 4 describes bisulfite sequencing that was performed on DNA from secondary tissue of WT, clsy1, nrpd1, drd1, nrpe1, and nrpd/e2. Recovered tissues generally contained ~70% cytosine methylation (WT, clsy1, drd1, and nrpd1), while non-recovered tissues had ~40% methylated cytosines (clsy1, nrpd1, nrpe1, and nrpd/e2). These results were consistent with previous work, but it was surprising that the nrpd/e2 mutants showed significant cytosine methylation. Contrary to current opinion in the field, Pol IV and Pol V alone are not required for the establishment of DNA methylation. This implicates Pol II as a likely
candidate for establishing this cytosine methylation because it associates with the viral IR and is known to carry out transcription that leads to methylation in *S. pombe*.

In collaborative work our lab has shown that in primary tissue, the TYLCCNV IR is targeted for DNA methylation, while coding regions are not (Yang et al. 2011). These data correlate well with the above ChIP data (described in 5.2.2) in that Pol II, Pol IV, and V associate with the viral IR during the primary infection, suggesting that the viral promoter region is targeted first by RdDM. The ChIP data also suggest that the CP coding region may be targeted for methylation in secondary recovered tissue. Consequently, we determined whether the association of Pol IV and Pol V with the CP coding region correlated with DNA methylation. Bisulfite sequencing analysis of the CP region in WT primary tissues found that this coding region is sparsely methylated (~6% total cytosine methylation) compared to the IR (previously reported as ~30%). But cytosine methylation spreads to the CP in WT secondary recovered tissues, increasing from 6% to 40% total methylation. When the CP was bisulfite sequenced in *nrpd/e2* non-recovered tissues the methylation decreased significantly to approximately 20%. This study indicates that Pol IV and Pol V play important roles in the spread of DNA methylation, but again are not required for basal levels of methylation. Taken together, these results imply that cytosine methylation is established by another polymerase (likely Pol II), while Pol IV and Pol V play critical roles in its amplification, maintenance, and spread.

These experiments have provided intriguing information on the individual roles of Pol IV and Pol V in RdDM, but much more work can be done. It will be interesting to determine if both Pol IV and Pol V are required for the spread of methylation to the CP or
if only one is essential. Moreover, bisulfite treatment of viral DNA in primary, secondary recovered and non-recovered tissues followed by Illumina sequencing could indicate where Pol II, Pol IV and Pol V initially target DNA for methylation and profile the extent to which it spreads. Furthermore, it will be critical to determine the precise role of Pol II in this process. Plants containing a weak nrpb2 mutation (second largest subunit of Pol II) have been infected with BCTV and bisulfite sequencing is in progress to assess the effect on DNA methylation.

5.2.4 Pol IV and Pol V are required for the establishment of dimethylated H3K9 repressive chromatin marks

The viral IR has been shown to contain both active (acetylated H3 or H3Ac) and repressive (dimethylated H3K9 or H3K9me2) chromatin marks in primary tissue (Raja et al. 2008). Furthermore, ChIP with antibodies to H3Ac and H3K9me2 followed by bisulfite sequencing (ChIP-BS) of the viral DNA associated with these marks confirmed that two populations of virus co-exist during an infection. ChIP performed with anti-H3Ac was able to slightly enrich for unmethylated viral DNA. Conversely, ChIP using anti-H3K9me2 showed significant enrichment of methylated viral genomes. This supports previous observations that cytosine methylation and dimethylated H3K9 are frequently coincident marks (Bernatavichute et al. 2008, Lippman et al. 2004, Zhang, et al. 2006, Turck et al. 2007, Vaughn et al. 2007, Zhang et al. 2009, Roudier et al. 2011). In Chapter 4, along with DNA methylation, the role of Pol IV and Pol V in the establishment and spread of repressive H3K9me2 was assessed. ChIP performed with anti-H3Ac (acetylated at lysine 14 and 9) found that in WT plants, both the IR and CP in
primary and secondary recovered tissues are associated with this chromatin mark. Similar to DNA methylation, H3K9me2 was only associated with the BCTV IR in WT primary tissue and subsequently spread to the CP in secondary recovered tissues. Therefore, like DNA methylation, the H3K9me2 spreads to coding regions and correlates with the presence of Pol IV and Pol V. Consequently, to determine the role of Pol IV and Pol V in recruiting this repressive histone mark, these ChIPs were repeated in nrpd/e2 and nrpd/e1 primary and secondary tissues. Interestingly, no dimethylated H3K9 was associated with the viral DNA in the absence of Pol IV and V, while the endogenous transposon, Ta3, associated with these repressive chromatin marks. Alternatively, active acetylated H3 was localized to both the IR and CP without either Pol IV or V. Finally, ChIP was performed in nrpd1 and nrpe1 individual mutant tissues and found that both of these polymerases are necessary for the establishment of H3K9me2 on viral DNA.

Although previous work has shown that dimethylation at H3K9 correlates with the presence of Pol IV and Pol V, this is the first time that these polymerases have been shown to be required for the establishment of this repressive chromatin mark. Moreover, this was a surprising result considering that DNA methylation can be established in the absence of Pol IV and Pol V and without the presence of H3K9me2, therefore suggesting that DNA methylation is laid down first and H3K9me2 is recruited later. Interestingly, maintenance of CNG methylation and dimethylated H3K9 is coordinated by a positive feedback loop between CMT3 and KYP2, respectively (Jackson et al. 2002, Malagnac et al. 2002, Jackson et al. 2004, Lindroth et al. 2004). This feedback loops appears to maintain the heterochromatin marks on endogenous Ta3 without Pol IV and Pol V, while this is not the case on the viral DNA. Altogether this data suggests that viral DNA
methylation is independent of maintenance pathways and is targeted by \textit{de novo} RdDM (discussed in section 5.2.6). Future work will be required to determine if the feedback loop on endogenous DNA can be broken. Mutations that reduce DNA methylation, such as \textit{ddm1}, could be crossed with \textit{nrpd/e1} and ChIP analysis for dimethylated H3K9 could be performed. Moreover, it would be interesting to assess the association of other repressive histone marks, such as H3K27me1 and H3K27me3, on the viral genome in the absence of Pol IV and Pol V.

5.2.5. Identifying a role for Pol II in siRNA biogenesis, DNA methylation and RdDM

Pol IV and V are known to play important roles in siRNA production. Pol IV is required for 24 nt siRNAs at the vast majority of loci, but some loci depend entirely on Pol V. Knowing that both Pol II, IV, and V associate with the viral IR, but only Pol II localized to the CP in primary tissue, it was important to assess which polymerase produces viral siRNAs at these loci. Small interfering RNAs of all size classes were analyzed in WT, \textit{clsy1}, \textit{nrpd1}, \textit{drd1}, \textit{nrpe1}, \textit{nrpd/e2}, and \textit{nrpd/e1}. Radiolabeled oligonucleotides were used to probe for CP derived siRNAs. In support of the Pol IV and Pol V ChIP analysis (discussed in section 5.2.2), the siRNAs were unchanged in the various Pol IV and Pol V mutant plants. The blots were then stripped and re-probed with oligonucleotides complementary to the BCTV IR. Mutations in Pol IV and Pol V, as well as their putative chromatin remodelers, all resulted in reduced siRNAs of all size classes. But, surprisingly, siRNAs never disappeared, even in the \textit{nrpd/e2} and \textit{nrpd/e1} mutants. These results indicate that Pol IV and Pol V are not required for the biogenesis
of virus-derived siRNAs. Instead it is likely that another polymerase participates in the production of siRNAs of all size classes.

The delayed recovery phenotype observed in nrpd1 and clsy1, the persistence of DNA methylation and siRNAs in the absence of Pol IV and Pol V altogether imply that another polymerase is involved in RdDM. Many lines of evidence suggest that the best candidate polymerase is Pol II. RdDM in S. pombe is performed entirely by Pol II and Pol IV and V not only share subunits with Pol II, but also evolved from it. More direct evidence comes from our observation and that of others who have shown that Pol II transcribes the viral DNA, interacts with RdDM components (AGO4), and recruits Pol IV and Pol V to target loci (Zheng et al. 2009). Further work with the nrpb2 mutant plants and protoplasts will be critical for characterizing the exact functions of Pol II in RdDM.

5.2.6 A model for geminivirus de novo methylation

Altogether, the work presented in Chapter 4 of this thesis presents several novel findings. First, the sensitivity of the geminivirus recovery assay was able to phenotypically distinguish nrpd1 and nrpe1 mutant plants. While Pol IV and Pol V play an important role in the spread of DNA methylation and the establishment of repressive dimethylated H3K9 on the viral genome, they are not required for establishment of DNA methylation at least in some regions. Furthermore, the biogenesis of geminivirus siRNAs can occur independent of Pol IV and Pol V. Taken together, these studies indicate that another polymerase, such as Pol II, generates siRNAs and establishes DNA methylation.

The novel information uncovered in this work is likely due to the sensitivity and unique nature of the geminivirus model system. Unlike endogenous genes, transgenes,
and transposons, geminiviruses do not require several generations or cell divisions to assess RdDM. Instead, the rolling-circle replication system precludes DNA methylation and spread of the virus likely occurs in the ssDNA form. Moreover, the bisulfite results show that the BCTV IR contains extensive CHH methylation, which is a trademark for the de novo methylation pathway. Furthermore, recovery from geminivirus infection does not require DDM1, which is a critical factor for maintenance of RdDM at transposons (J. Jackel unpublished data). Thus, it is likely that the geminivirus model system directly interrogates the de novo RdDM pathway.

Taken together, our data suggest a model based on the ancestral role of RNA Polymerase II in S. pombe in TGS. In this model, another polymerase, possibly Pol II, initially generates nc-transcripts and siRNAs that establish de novo DNA methylation. However, this process could be self-limiting because transcription by this polymerase is likely inhibited by DNA methylation. Pol IV and V are subsequently recruited to methylated DNA to reinforce, amplify, and spread DNA methylation, as well as establish repressive H3K9me2 histone marks (Fig. 4.7). The Pol IV delayed recovery phenotype support this model in that Pol IV can be bypassed in this system through Pol II redundant functions. Pol II could transcribe viral DNA, produce siRNAs, which lead to DNA methylation through either Pol II or Pol V. Then, AGO4 association with either of these polymerases can slice the non-coding transcript and serve as a template for RDR2 and subsequently DCL3, thus, resulting in amplification of the siRNAs and DNA methylation.

The results and model proposed here, open the door for many future studies. First, it will be critical to assess the de novo nature of the geminivirus model system. Since it is
the ssDNA form that is packaged and believed to spread cell to cell, it will be essential to validate that this form is unmethylated. The above model also posits that recovery only requires DNA methylation and can occur in the absence of dimethylated H3K9. To this end, it has been determined that ddc mutants do not recover from BCTV L2- infection suggesting that DNA methylation is necessary for recovery. Additionally, kyp2 recovered from geminivirus infection, but KYP2 shares redundant functions with SUVH 5 and 6 (J. Jackel and P. Raja unpublished data). Therefore, suvh4,5,6 mutants will need to be tested for recovery and tissue obtained will be subject to bisulfite sequencing to analyze viral DNA methylation in the absence of these histone MTases. Subsequently, it will be exciting to determine how the de novo methylation pathway targets geminiviruses and foreign DNA (discussed in 5.3.1). Moreover, due to the high copy number of geminiviruses and the critical role of RdDM in defense, it would be interesting to assess the impact that viral infection has on the expression of various pathway components. Finally, it will be vital to evaluate the role of Pol II in RdDM by analyzing siRNA production and DNA methylation during geminivirus infection of nrpb2 mutant plants and protoplasts.

5.3 Geminiviruses as models to study chromatin targeting and reading

5.3.1 Determining how Pol II, Pol IV, and Pol V recognize and target DNA for methylation

Loss-of-heterozygocity and genome-wide epigenetic repression has been associated with many cancers (Reviewed in Jones and Baylin 2002, Sharma et al. 2010). Therefore,
determining how DNA is targeted for repressive histone and DNA modifications is a major question in both plants and animals. Plants offer a great model system to study this process because much more is known about the genes involved in TGS. Moreover, the potential role of Pol II in RdDM in plants could provide a basis for animal studies since Pol IV and V are plant specific. To begin to analyze how Pol II, IV, and V target DNA for methylation, it will be critical to identify and characterize the non-coding (nc)-transcripts produced by these polymerases. The nc-transcripts will provide important information on where initiation and termination occurs by each of these polymerases. This information could provide an opportunity to isolate and identify the non-coding form of these polymerases. Additionally, the differential localization of Pol IV and Pol V on the IR and CP regions during BCTV infection will provide a unique opportunity to determine which factors or DNA features target these polymerases to specific regions of the viral DNA. DNA methylation is one feature that is hypothesized to recruit Pol IV and Pol V. Thus, ChIP-BS will be performed to directly assess the cytosine methylation of the DNA associated with each of these polymerases. Further analysis would include in vitro methylation of viral DNA, transformation into protoplasts followed by ChIP to determine polymerase association. Altogether future experiments will provide novel information about how DNA is recognized by Pol II, Pol IV and Pol V and the process of non-coding transcription, which can be applied to other eukaryotic systems to increase our understanding of the establishment and maintenance of DNA methylation.
5.3.2 Geminiviruses as models to identify chromatin readers

The RdDM pathway writes the epigenetic code, but in order for the cell to respond accordingly, chromatin readers must recognize these marks. While a lot is known about the proteins that modify DNA and histones, very little is known about the proteins in plants that recognize such marks. In chapters 3 and 4 of this thesis, it was shown how the geminivirus model system identified and characterized RdDM pathway components. Therefore, geminiviruses could serve as sensitive models to discover and characterize chromatin readers as well. Mammalian proteins that recognize specific chromatin modifications belong to the "Royal family", which contain homologous Tudor, chromatin binding (Chromo), PWWP or MBT domains (Kim et al. 2006, Taverna et al. 2007, Wu et al. 2011). A plant-specific Agenet domain, belonging to this "Royal family", has been found in approximately 20 Arabidopsis proteins (Maurer-Stroh et al. 2003). Preliminary work on 4 of these proteins (named ACK1-ACK4) found that plants deficient for these proteins show extremely interesting phenotypes. CaLCuV infection of ack3 plants resulted in a severely hypersusceptible phenotype, while ack1, ack2, and ack4 plants appeared to be tolerant (with V. Fernandez unpublished data). Analysis of the viral titers, viral cytosine methylation, siRNAs and chromatin modifications associated with these phenotypes are in progress. Additionally, interaction with histones, specific histone modifications, and the viral genome will be assessed (T. Coursey unpublished data). Furthermore other plants deficient for Agenet domain containing proteins are currently being screened by CaLCuV infection. All together, these studies will provide novel information on how cellular chromatin is recognized and interpreted. Moreover, the
discovery of mutant plants that are more tolerant of geminivirus infection could be very beneficial for crop production.

Fig. 5.1 A new model for AL2 silencing suppression activities. As previously shown, geminivirus AL2 and L2 proteins interact with and inactivate ADK. This suppresses the methyl cycle because adenosine salvage is required for flux through the cycle, thus reducing S-adenosyl methionine (SAM) production and increasing the abundance of S-adenosyl homocysteine (SAH), a competitive inhibitor of SAM. Therefore, inhibition of the methyl cycle partially breaks down the maintenance of methylation, resulting in the reversal of endogenous TGS and a reduction of methylation within the genome. Moreover, since siRNAs are methylated to increase stability, ADK inhibition also suppresses PTGS. In N. benthamiana after flowering, the role of ADK in the methyl cycle changes and other adenosine nucleosidases likely salvage adenosine.
AL2 and L2 inhibition of ADK is sufficient for PTGS and TGS reversal in vegetative plants only. AL2 also has transcription-dependent suppressor activity due to the activation of a number of host proteins, including WEX1 and WEL1, which may act as endogenous suppressors of silencing. Although the mechanisms are unknown, these endogenous genes likely add to the ability of AL2 to reverse PTGS and allow it to block systemic spread of silencing. AL2 reverses TGS on silenced transposons like CACTA in an unknown transcription-dependent manner. Finally, since ADK inhibition is insufficient to reverse TGS in reproductive plants, AL2 has evolved a transcription- and ADK-independent suppressor activity. Preliminary data suggests AL2 could also reverse TGS by interacting and inhibiting TGS pathway components.


Pelissier, T., Clavel, M., Chaparro, C., Pouch-Pelissier, M.-N., Vaucharet, H., and Deragon, J.-M. 2011. Double-stranded RNA binding proteins DRB2 and DRB4 have an antagonistic impact on polymerase IV-dependent siRNA levels in Arabidopsis. RNA 17:1502-1510.


Schwach, F., Vaistij, F.E., Jones, L., and Baulcombe, D.C. 2005. An RNA-dependent RNA-polymerase prevent meristem invasion by Potato virus X and is required for the


2005. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in Arabidopsis. Genome Biol. 6: R90.01-R90.11.


APPENDIX A

SUMMARY OF PUBLICATION CONTRIBUTIONS

During my graduate career I have contributed to several research projects that are not included in the preceding chapters of this dissertation. The data obtained from these studies are being prepared, have been submitted, or have been accepted for publication in peer-reviewed journals. The following sections contain manuscript abstracts followed by a brief description of my contributions to each publication.

6.1 Geminivirus AL2 and L2 Proteins Suppress Transcriptional Gene Silencing and Cause Genome-Wide Reductions in Cytosine Methylation.


Geminiviruses replicate single-stranded DNA genomes through double-stranded intermediates that associate with cellular histone proteins. Unlike RNA viruses, they are subject to RNA-directed methylation pathways that target viral chromatin and likely lead to transcriptional gene silencing (TGS). Here we present evidence that the related
geminivirus proteins AL2 and L2 are able to suppress this aspect of host defense. AL2 and L2 interact with and inactivate adenosine kinase (ADK), which is required for efficient production of S-adenosyl methionine, an essential methyltransferase cofactor. We demonstrate that the viral proteins can reverse TGS of a green fluorescent protein (GFP) transgene in *Nicotiana benthamiana* when overexpressed from a *Potato virus X* vector and that reversal of TGS by geminiviruses requires L2 function. We also show that AL2 and L2 cause ectopic expression of endogenous *Arabidopsis thaliana* loci silenced by methylation in a manner that correlates with ADK inhibition. However, at one exceptional locus, ADK inhibition was insufficient and TGS reversal required the transcriptional activation domain of AL2. Using restriction-sensitive PCR and bisulfite sequencing, we showed that AL2-mediated TGS suppression is accompanied by reduced cytosine methylation. Finally, using a methylation-sensitive single-nucleotide extension assay, we showed that transgenic expression of AL2 or L2 causes global reduction in cytosine methylation. Our results provide further evidence that viral chromatin methylation is an important host defense and allow us to propose that as a countermeasure, geminivirus proteins reverse TGS by nonspecifically inhibiting cellular transmethylation reactions. To our knowledge, this is the first report that viral proteins can inhibit TGS.

As a third author on this publication, I performed experiments that confirmed TGS reversal caused by the over-expression of AL2 and L2 viral proteins. I found that BCTV and CaLCuV infection reversed TGS of the GFP transgene in *Nicotiana benthamiana*, while BCTV L2' infection was unable to reverse TGS. GFP expression was assessed visually under UV light and validated by Northern blot analysis. The data was presented
in Figure 1C and 1D. I also performed methylation-sensitive single-nucleotide extension assay control experiments which indicated that no significant incorporation of radioactive $\alpha^{32}\text{P-dCTP}$ occurred following a 1-hour incubation of genomic DNA with Taq DNA polymerase. Moreover, I showed that dexamethasone treatment of wild type *Arabidopsis thaliana* did not reduce genomic methylation compared to a water control. These control experiments were necessary to show that dexamethasone induction of AL2 and L2 gene expression in transgenic plants reduced global genomic cytosine methylation. The data from these experiments are presented in supplemental data Figures 3A and 3B.

6.2 Suppression of methylation-mediated transcriptional gene silencing by $\beta$C1-SAHH protein interaction drives geminivirus-DNA $\beta$ satellite association.

Xuiling Yang, Yan Xie, Priya Raja, Sizhin Li, Jamie N. Wolf, Qingtang Shen, David M. Bisaro, and Xueping Zhou. PLoS Pathogens (2011) Vol. 7: e1002329

DNA methylation is a fundamental epigenetic modification that regulates gene expression and represses endogenous transposons and invading DNA viruses. As a counter-defense, the geminiviruses encode proteins that inhibit methylation and transcriptional gene silencing (TGS). Some geminiviruses have acquired a betasatellite called DNA $\beta$. This study presents evidence that suppression of methylation-mediated TGS by the sole betasatellite-encoded protein, $\beta$C1, is crucial to the association of *Tomato yellow leaf curl China virus* (TYLCCNV) with its betasatellite (TYLCCNB). We show that TYLCCNB complements *Beet curly top virus* (BCTV) $L2^*$ mutants.
deficient for methylation inhibition and TGS suppression, and that cytosine methylation levels in BCTV and TYLCCNV genomes, as well as the host genome, are substantially reduced by TYLCCNB or βC1 expression. We also demonstrate that while TYLCCNB or βC1 expression can reverse TGS, TYLCCNV by itself is ineffective. Thus its AC2/AL2 protein, known to have suppression activity in other geminiviruses, is likely a natural mutant in this respect. A yeast two-hybrid screen of candidate proteins, followed by bimolecular fluorescence complementation analysis, revealed that βC1 interacts with S-adenosyl homocysteine hydrolase (SAHH), a methyl cycle enzyme required for TGS. We further demonstrate that βC1 protein inhibits SAHH activity *in vitro*. That βC1 and other geminivirus proteins target the methyl cycle suggests that limiting its product, S-adenosyl methionine, may be a common viral strategy for methylation interference. We propose that inhibition of methylation and TGS by βC1 stabilizes geminivirus/betasatellite complexes.

I conducted experiments that helped show that TYLCCNB is able to complement BCTV L2− mutants deficient for methylation inhibition and TGS suppression. Infection of *N. benthamiana* with BCTV in conjunction with TYLCCNB showed symptoms characteristic of both the virus and satellite DNA. In addition, plants were able to recover from BCTV L2− infection, but recovery was lost when BCTV L2− was inoculated simultaneously with TYLCCNB. I performed Southern blot analysis to show that in symptomatic secondary tissue, BCTV L2− viral DNA was increased in the presence TYLCCNB. I also detected TYLCCNB DNA using PCR followed by gel electrophoresis. These results indicated that TYLCCNB complements the methylation inhibition deficient BCTV L2−. These studies were published in Figure 2A-2D. To show
that TYLCCNB complements BCTV $L^2$ in TGS suppression, I inoculated *N. benthamiana* plants containing a transcriptionally silenced GFP transgene (line 16-TGS) with BCTV $L^2$ in the presence or absence of TYLCCNB. Visual observation under UV light and Northern blot analysis indicated that BCTV $L^2$ together with TYLCCNB reversed TGS and caused GFP expression, while BCTV $L^2$ alone was insufficient. These results are presented in Figure 4B and 4D of this publication. Finally, I prepared and modified the methyl cycle diagram shown in Figure 9 of the discussion.

6.3 Characterization of the RNA silencing suppression activity of the Ebola virus VP35 protein in plants and mammalian cells.


Ebola virus (EBOV) causes a lethal hemorrhagic fever for which there is no approved effective treatment or prevention strategy. EBOV VP35 is a virulence factor that blocks innate antiviral host responses, including the induction of and response to alpha/beta interferon. VP35 is also an RNA silencing suppressor (RSS). By inhibiting microRNA-directed silencing, mammalian virus RSSs have the capacity to alter the cellular environment to benefit replication. A reporter gene containing specific microRNA target sequences was used to demonstrate that prior expression of wild-type VP35 was able to block establishment of microRNA silencing in mammalian cells. In addition, wild-type VP35 C-terminal domain (CTD) protein fusions were shown to bind small interfering RNA (siRNA). Analysis of mutant proteins demonstrated that reporter activity in RSS assays did not correlate with their ability to antagonize double-stranded RNA (dsRNA)-
activated protein kinase R (PKR) or bind siRNA. The results suggest that enhanced reporter activity in the presence of VP35 is a composite of nonspecific translational enhancement and silencing suppression. Moreover, most of the specific RSS activity in mammalian cells is RNA binding independent, consistent with VP35's proposed role in sequestering one or more silencing complex proteins. To examine RSS activity in a system without interferon, VP35 was tested in well-characterized plant silencing suppression assays. VP35 was shown to possess potent plant RSS activity, and the activities of mutant proteins correlated strongly, but not exclusively, with RNA binding ability. The results suggest the importance of VP35-protein interactions in blocking silencing in a system (mammalian) that cannot amplify dsRNA (Zhu et al. 2012).

As co-author, I performed plant silencing suppression assays to test the RSS activity of VP35 without interferon. RSS activity was examined using the well-established two- and three-component assays. In the three-component system silencing is established by infiltrating N. benthamiana leaves with Agrobacterium containing plasmids expressing GFP and a long-hairpin form of GFP. Suppression of this silencing is tested by simultaneous infiltration of an additional plasmid containing a control or test gene. In the two-component system, silencing is initiated when a plasmid expressing GFP is infiltrated into a GFP-transgenic line (16c) of N. benthamiana. Again control or test genes are concurrently infiltrated to test suppression of silencing. GFP silencing suppression was assessed using a UV-light and Northern blot analysis to detect GFP expression at 3, 5, and 7 days post-infiltration. To determine if RNA binding ability impacted VP35 RSS activity in plants, I used constructs expressing VP35, as well as VP35 mutants that eliminate RNA binding activity (VP35 K309A, VP35 R312A, and
VP35 K309A/R312A) or mutations that do not change RNA binding (VP35 G333S). UV light and Northern blot assays showed that VP35 possessed strong RSS ability in plants and mutation of the RNA binding domain strongly compromised this activity. These data were published in Figure 4A through 4D.

6.4 A complex containing SNR1-related kinase (SnRK1) and adenosine kinase in *Arabidopsis*.

Gireesha Mohannath, Jamie N. Jackel, Youn Hyung Lee, R. Cody Buchmann, Hui Wang, Veena Patil, Allie Varner, and David M. Bisaro (manuscript submitted).

SNF1-related kinase (SnRK1) in plants belongs to a conserved family that includes sucrose non-fermenting 1 kinase (SNF1) in yeast and AMP-activated protein kinase (AMPK) in animals. These kinases play important roles in the regulation of cellular energy homeostasis, and in response to stresses that deplete ATP, they inhibit energy consuming anabolic pathways and promote catabolism. Energy stress is sensed by increased AMP:ATP ratios. When AMP levels are high, the SNF1/AMPK/SnRK1 kinases are activated, in part through allosteric effects mediated by AMP itself. In previous studies, we showed that geminivirus pathogenicity proteins interact with both SnRK1 and adenosine kinase (ADK), which phosphorylates adenosine to generate AMP. This suggested a relationship between SnRK1 and ADK, which we investigate in the studies described here. We demonstrate that SnRK1 and ADK physically associate in the cytoplasm, and that SnRK1 stimulates ADK in vitro by a novel, non-enzymatic mechanism. Further, altering the activity of either SnRK1 or ADK in transgenic plants...
altered the activity of the other kinase, providing evidence for in vivo linkage. This study establishes the existence of SnRK1-ADK complexes that likely play important roles in energy homeostasis and cellular responses to biotic and abiotic stress (G. Mohonnath unpublished).

As a co-author, I contributed data to show that ADK and SnRK1 activity are linked in vivo. It was observed that increased ADK activity resulted in an increase in SnRK1 activity, while reducing ADK activity had no effect on SnRK1 activity. I performed reciprocal experiments to examine how changing SnRK1 activity would impact ADK activity. I employed transgenic N. benthamiana lines that constitutively expressed Arabidopsis SnRK1.2 (sense lines) or antisense SnRK1.2 RNA (Hao et al., 2003). SnRK1 activity was measured after the addition of γ^{32}P-ATP to the SnRK1 recognized GST-SAMS peptide. SnRK1 over-expression lines displayed 1.5 to 2.5-fold increased activity, while the antisense lines reduced SnRK1 activity to 35-40% of wild-type. Soluble protein extracts obtained from these transgenic lines were then used to measure ADK activity, which was monitored by measuring the production of labeled AMP from the addition of γ^{32}P-ATP to adenosine. While over-expression of SnRK1 appeared to have no effect on ADK activity, reduced SnRK1 activity led to dramatic 7-fold increases in ADK activity. Together these experiments demonstrated that ADK and SnRK1 activities are linked in vivo. These results are presented in Figures 7C and 7D.
APPENDIX B

SUPPLEMENTAL DATA FIGURES

Fig. B.1 Arabidopsis drb3 and dcl3 mutants do not recover from infection with BCTV L2'. (A) Secondary shoots of wild-type (Col-0), drb2, drb3, drb4 and drb5 plants infected with BCTV L2' are shown. (B) Secondary shoots of wild-type (Col-0), dcl2, dcl3, and dcl4 plants infected with BCTV L2' are shown. Wild-type (Col-0) plants recover from infection, and present asymptomatic shoots. The drb3 and dcl3 mutants fail to recover and secondary shoots continue to display severe symptoms of virus infection. The drb2, drb4, drb5, dcl2, and dcl4 plants recover and secondary shoots show little evidence of disease. However, although secondary shoots lack floral deformation, drb4 and drb5 plants were sometimes stunted, and examples are shown in this figure.
Fig. B.2. Cytosine methylation analysis of the BCTV L2 genome in wild-type and mutant plants. Dot plots depict cytosine methylation profiles of the viral intergenic region (IR). The 44 cytosines in the IR were analyzed by bisulfite sequencing and 12 individual clones for each treatment are shown, arranged from most to least methylated. Each filled circle represents a methylated cytosine, while open circles represent unmethylated cytosines. The dots are color-coded based on sequence context: CG (red), CNG (blue), CHH (green). Data are shown graphically to the right. (A) Wild-type Col-0 plants recover from infection and the IR is hypermethylated (63% of total cytosines methylated). Most genomes are densely methylated. (B) drb3 plants fail to recover, and total methylation in the IR is reduced to ~34%, with a larger proportion of genomes unmethylated or hypomethylated. Methylation is reduced in all sequence contexts. (C) Viral genomes from recovered drb4 plants show some reduction in total methylation relative to Col-0 (54% compared to 63%), but most are densely methylated. (D) Methylation is substantially reduced (to 48%) in non-recovered dcl3 plants. (E) dcl4 plants recover and methylation is similar to wild-type Col-0 (~70%). A majority of viral genomes are hypermethylated. (F) Wild-type Ler-0 plants recover from infection and the IR is hypermethylated (77% of total cytosines methylated). (G) Methylation is substantially reduced (~38%) in non-recovered ago4 plants. The ago4 mutation is in the Ler-0 ecotype background.
Fig. B.3. Cytosine methylation analysis of the BCTV L2' genome in wild-type, *drb3*, and *dcl3* mutant plants. As in Figure B.2, except data are from a second, independent experiment involving 9 clones for each treatment. In this instance, total methylation levels were 62%, 35%, and 43% for wild-type, *drb3*, and *dcl3* plants, respectively.
Fig. B.4. Hypermethylated viral genomes are associated with H3K9me2. ChIP-BS was performed with extracts from BCTV L2 infected, recovered tissue using antibodies to H3Ac or H3K9me2. Input (extract without ChIP) and ChIP DNAs were bisulfite treated and PCR primers spanning the viral IR were used to amplify associated DNA. Products were cloned and sequenced. Data are presented as in Figure B.2. Total methylation levels were ~70%, 55%, and 90% for wild-type, H3Ac, and H3K9me2 samples, respectively. The difference between input and H3K9me2 fractionated samples was significant at the 95% (*) confidence level, as determined by Student's t test.
Fig. B.5. Interactions between DRB3, DCL3 and AGO4 proteins using BiFC analysis in *N. benthamiana* epidermal cells. Constructs expressing the indicated proteins were fused to the N- or C-terminal portion of YFP and fusion proteins were co-
expressed in *N. benthamiana* leaves. Constructs were delivered by agroinfiltration, and cells were photographed 36 hr later using a confocal laser scanning microscope. RFP-fibrillarin was used as a marker for the nucleolus. Protein combinations in each infiltration are indicated above each photograph. Panels show nuclei at 100X magnification. (A) DRB3:DCL3 complexes are concentrated in subnuclear bodies that sometimes co-localize with the nucleoli. (B) DRB3:AGO4 complexes are enriched in small bodies that do not associate with nucleoli.

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence (5' to 3')</th>
<th>Assay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 CaLCuV2556CF</td>
<td>GGGGATATGTTAAGAATATATTTTG</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>2 CaLCuV359CR</td>
<td>TCCCCACCATAAACAACACC</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>3 BCTV2640CF</td>
<td>GGGATATGTAAGAAATATG</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>4 BCTV147CR</td>
<td>TCTATTAACCAACAAACAATCT</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>5 BCTVCPF57</td>
<td>TCAAAGTGCCTGGAAGAAGAA</td>
<td>21nt marker</td>
</tr>
<tr>
<td>6 BCTVIRsR-2</td>
<td>TTCCTAGGAAGTCTCCCGCT</td>
<td>22nt marker</td>
</tr>
<tr>
<td>7 BCTVIRqRTTR</td>
<td>ACAGATCGCCCTATGAGAGGTTT</td>
<td>24nt marker</td>
</tr>
<tr>
<td>8 BCTVIRsR-2</td>
<td>TGAGCGGGAAACTTCTGAAGA</td>
<td>Probe</td>
</tr>
<tr>
<td>9 BCTVIRqRTTR</td>
<td>ACAACTCTCATAAGGCTCAGGTA</td>
<td>Probe</td>
</tr>
<tr>
<td>10 BCTVIRsF-1</td>
<td>GACACGTGGAAGGTTCTGTAAGA</td>
<td>Probe</td>
</tr>
<tr>
<td>11 BCTVIRsF-2</td>
<td>TAAGATTCTGTTGACTGGAATAGAGAAG</td>
<td>Probe</td>
</tr>
<tr>
<td>12 BCTVIRsR-1</td>
<td>GTTTTGATTGGAAGGCTTC</td>
<td>Probe</td>
</tr>
<tr>
<td>13 BCTVIRqRTF</td>
<td>ATCGACCTCCAGGACGTCGCT</td>
<td>Probe</td>
</tr>
<tr>
<td>14 BCTVCPF57</td>
<td>TCAAAGTGCCTGGAAGAAGAA</td>
<td>Probe</td>
</tr>
<tr>
<td>15 BCTVCPF1122</td>
<td>TTCCCGATAACGGTCAGGCTAT</td>
<td>Probe</td>
</tr>
<tr>
<td>16 BCTVCPF1273</td>
<td>TGGCATTGGAAGGCGCTTGT</td>
<td>Probe</td>
</tr>
</tbody>
</table>

**Fig. B.6. Sequences of oligonucleotides used in this Chapter 3.** Oligonucleotides 1 through 4 were used as primers for PCR amplification and cloning of *Cabbage leaf curl virus* (CaLCuV) or *Beet curly top virus* (BCTV) intergenic regions for bisulfite sequencing. Oligonucleotides 5 through 7 served as size markers for small RNAs, while 8 through 16 were end-labeled with $^{32}$P for use as hybridization probes to detect BCTV-specific small RNAs on gel blots.

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence (5' to 3')</th>
<th>Assay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 BCTV147CR</td>
<td>TCTATTAACCAACAAACAATCT</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>2 BCTV2640CF</td>
<td>GGGATATGTAAGAAATATG</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>3 bisBCTV1CP2F</td>
<td>ATGAGGAAATATAAAGAAATAAGTAT</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>4 BisBCTV362CP2R</td>
<td>ACCCACAATCCTTTTCTATTATAATC</td>
<td>Bisulfite</td>
</tr>
<tr>
<td>5 qAct2FA118</td>
<td>GAGAGATTCAGAGGCCCAGAAAGT</td>
<td>ChIP (sqPCR and qPCR)</td>
</tr>
<tr>
<td>No.</td>
<td>Oligonucleotide</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----</td>
<td>------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>6</td>
<td>qAct2RA119</td>
<td>TGGATTCAGCAGCTTCCA</td>
</tr>
<tr>
<td>7</td>
<td>qIGN5FA193</td>
<td>AAGCCCAAACCATAACAATAATCTAAT</td>
</tr>
<tr>
<td>8</td>
<td>qIGN5RA194</td>
<td>CCGAATAACAGCAAGTCTTTTAATA</td>
</tr>
<tr>
<td>9</td>
<td>BCTVIRqRTF</td>
<td>ATGGGACTCCGATGACGTGGGTTTA</td>
</tr>
<tr>
<td>10</td>
<td>BCTVIRqRTR</td>
<td>ACGGATGGCCCTTATGAGAGTTGT</td>
</tr>
<tr>
<td>11</td>
<td>BCTVCPF1122</td>
<td>TTCCCGATAACGGTCAGGCTAT</td>
</tr>
<tr>
<td>12</td>
<td>BCTVCPRI273</td>
<td>TGGCATGGAAGGAGCCCTTTGT</td>
</tr>
<tr>
<td>13</td>
<td>BCTVCP3F</td>
<td>ACCAATGGGAAGGAGCCTGTGACCA</td>
</tr>
<tr>
<td>14</td>
<td>BCTVCP3R</td>
<td>ATCCCTATACCAGCAGAGCCCAAACA</td>
</tr>
<tr>
<td>15</td>
<td>BCTVCPF57</td>
<td>TCAAAGTGCGTGCCCAGAAGAA</td>
</tr>
<tr>
<td>16</td>
<td>BCTVIRsR-2</td>
<td>TTCTTCAAGGAAGTTTCGCTCA</td>
</tr>
<tr>
<td>17</td>
<td>BCTVIRsRTR</td>
<td>ACGGATGGCCTTATGAGAGTTGT</td>
</tr>
<tr>
<td>18</td>
<td>BCTVIRsR-2</td>
<td>TGAGCGGAAAACCTTCTGAAGA</td>
</tr>
<tr>
<td>19</td>
<td>BCTVIRsF-1</td>
<td>GACACGTGGAAGGGTACTGTGAA</td>
</tr>
<tr>
<td>20</td>
<td>BCTVIRsF-2</td>
<td>TAAGATTGTGATCTGGTCAATAGAAGG</td>
</tr>
<tr>
<td>21</td>
<td>BCTVIRsR-1</td>
<td>GTTTTGATTGGTTGCCCATTTTGGTTTCG</td>
</tr>
<tr>
<td>22</td>
<td>BCTVCPF57</td>
<td>TCAAAGTGCGTGCCAGAAGAA</td>
</tr>
<tr>
<td>23</td>
<td>BCTVCPF1122</td>
<td>TTCCCGATAACGGTCAGGCTAT</td>
</tr>
<tr>
<td>24</td>
<td>BCTVCPF1273</td>
<td>TGGCATGGAAGGAGCCTTGGT</td>
</tr>
</tbody>
</table>

**Fig. B.7 Oligonucleotides used in this Chapter 4.** Oligonucleotides 1 through 4 were used as primers for PCR amplification and cloning of *Beet curly top virus* (BCTV) intergenic region or coat protein coding region for bisulfite sequencing. Oligonucleotides 5 through 8 were used to amplify Actin and IGN5 control regions of ChIP-sqPCR and ChIP-qPCR (Weirzbicki et al. 2009). Oligonucleotides 9 and 10 were used to amplify the BCTV IR for sqPCR and qPCR amplification after ChIP. Oligonucleotides 11 and 12 were used to amplify the BCTV CP region in ChIP-sqPCR, while 13 and 14 were used in ChIP-qPCR for the BCTV CP region. Oligonucleotides 15 through 17 served as size markers for small RNAs, while 18 through 24 were end-labeled with $^{32}$P for use as hybridization probes to detect either BCTV IR- or BCTV CP-specific small RNAs on gel blots.
Fig. B.8 Statistical analysis of bisulfite clones isolated from recovered and non-recovered tissues. The bisulfite clones from all mutant plants (obtained from this experiment and others) that were classified as non-recovered were grouped together and the mean value of methylation was determined. This was repeated for all bisulfite clones gathered from recovered plants. Standard error of the mean was determined. This evaluation suggested that there is little variation between bisulfite experiments. Thus suggesting that sequencing 12 viral clones is sufficient to represent the viral population. The difference between recovered and non-recovered fractionated samples was significant at the 99% (**) confidence level, as determined by Student's $t$ test.

Fig. B.9 Pol II-, Pol IV-, and Pol V-FLAG transgenic plants recover from geminivirus infection. Dissecting microscope pictures of secondary floral heads of BCTV $L2^*$ infected NRPB2-, NRPD1-, and NRPE1-FLAG tagged transgenic plants used for ChIP assays. Observe the lack of viral symptoms indicative of recovery from BCTV $L2^*$ infection. Thus, the FLAG epitope tagged transgene complemented the mutant phenotype.
Fig. B.10 Pol IV and V are not required for the establishment or spread of acetylated H3 histone marks. Each histogram represents technical triplicates of ChIP qPCR performed in biological triplicate with antibodies to acetylated H3 using extracts from WT, nrpd1, nrpe1, and nrpd/e1 BCTV 1° and BCTV L2- 2° infected tissue. (Only nrpd1 BCTV L2- recovered tissue was used for the ChIP analysis.) The graphs show the ChIP signal as % Input. H3Ac is present on the viral IR and CP in both primary and secondary infected tissue (blue bars and red bars, respectively) of Col-0 WT (A), nrpd1 (B), nrpe1 (C), or nrpd/e2 (D) mutants. Although the relative abundance of H3Ac appeared to change in the mutant, Pol IV and V individually do not play important roles in establishing and spreading H3Ac on the viral genome.