MECHANISM OF GENE SILENCING SUPPRESSION BY THE GEMINIVIRUS PROTEIN TrAP

A Dissertation

by

CLAUDIA MARCELA CASTILLO GONZÁLEZ

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Xiuren Zhang
Committee Members,	Hays Rye
	Timothy Devarenne
	Jun-Yuan Ji
Head of Department,	Gregory D. Reinhart

August 2017

Major Subject: Biochemistry

ABSTRACT

The hosts-virus arms race reaches the epigenetic level, where silencing of viral chromatin can serve as an innate defense mechanism to restrict invading DNA viruses. However, viruses can code for suppressor proteins to counter epigenetic silencing and escape host surveillance. Thus, the virus-encoded suppressors offer an untapped source of tools for the understanding of pathogenesis and chromatin regulation.

TrAP is a transcription factor encoded by model DNA plant viruses of the family Geminiviridae, which is required for the expression of the virus late genes and for suppression of gene silencing. TrAP is known to interfere with the transcriptional gene-silencing (TGS) pathway by obstructing the methyl cycle in the cytoplasm. Nonetheless, multiple metabolic pathways other than chromatin regulation utilize the methyl donor, and TrAP mainly localizes to the nucleus; furthermore, TrAP is predicted to interact with the transcriptional machinery. Thus, we asked whether TrAP directly suppressed TGS.

We first generated TrAP-stable transgenic plants, and through transcriptome and biochemical assays, we demonstrated that TrAP hampered TGS. We then identified TrAP-interacting partners using a proteomics approach, confirmed by protein interaction experiments *in vivo* and *in vitro*. To determine whether these interactions were physiologically relevant, we performed virus infection assays in various host genetic backgrounds.

We demonstrated that TrAP interacts with multiple SET-domain proteins in Arabidopsis. Particularly, the H3K9me2 histone methyltransferase, Su(var)3-9 homolog 4/Kryptonite (SUVH4/KYP) is a *bona fide* cellular target of TrAP. TrAP expression phenocopies several TGS mutants, reduces the repressive H3K9me2 mark and CHH DNA methylation, and reactivates many endogenous KYP-repressed loci *in vivo*. KYP binds to the viral chromatin and controls its methylation to combat virus infection. We conclude that TrAP attenuates TGS by inhibiting KYP activity.

Furthermore, we show that TrAP interacts with other proteins, such as the methyl cycle enzymes SAMe-synthetase 2 (SAM2) and the S-adenosyl homocysteine hydrolase 1 (SAHH1), the RNA processing Enhanced silencing phenotype 3 (ESP3), and the chromatin remodeler Relative of early flowering 6 (REF6).

Our findings provide new insight in the host antiviral defense and virus counter-defense at an epigenetic level and provide a model system to study chromatin regulation, and virus infection.

DEDICATION

For the myriad times in which I saw this journey as the impossible, and the many more in which I was proven wrong, I dedicate this work to my beloved family.

To my parents Gabriel and Cristina, for their unconditional love and their exceptional example. To my sister Diana, for her unyielding spirit, her loyalty, and her coherence in life. To my godson Jaguar, for showing me the way to happiness, and for letting me be part of his discovery of the world. And to my husband Jeremy, the love of my life and my adamant accomplice, who can turn any challenge into an adventure, and brings perspective to my existence.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my mentor and committee chair, Dr. Xiuren Zhang, for pushing my boundaries, for believing in me sometimes more than I believed in myself, and for making his laboratory a nurturing environment to grow as a scientist. I would like to acknowledge Drs. Hays Rye, Tim Devarenne, Jun-Yuan Ji, Keith Maggert and Feng Qiao for serving as my committee members throughout these years, and for all their helpful advice and contributions towards my research. I would also like to express my deepest gratitude and appreciation to Dr. Dorothy Shippen, for her invaluable guidance and support during my graduate years; to Dr. Craig Kaplan for teaching me the discipline of thought; and, to Dr. Robert Klein for the stimulating discussions on science and life.

Thanks also go to our collaborators Drs. Xueping Zhou, Paul de Figueiredo, Hisashi Koiwa, and Joshua Yuan; and to the scientists, Drs. Linda Hanley-Bowdoin, Judith Brown, David Bisaro, and Garry Sunter, in this field who so kindly shared their materials, knowledge, and expertise to make this project possible.

I want to thank all former and current members of the Zhang laboratory, especially, Drs. Hongliang Zhu, Zhonghui Zhang, and Changjun Huang, for contributing uniquely to my graduate career. Finally, I thank my friends, colleagues, faculty, and staff of the Department of Biochemistry & Biophysics and the Institute for Plant Genomics and Biotechnology, for making Texas A&M University a second home.

NOMENCLATURE

ABRC	Arabidopsis Biological Resource Center
DCL	Dicer-like protein
DNA	Deoxyribonucleic acid
dsRNA	double-stranded RNA
ESP3	Enhanced Silencing Phenotype 3
HEN1	HUA enhancer 1
KYP	Kryptonite
miRNA	microRNA
mRNA	messenger RNA
NPC	Nuclear pore complex
pre-miRNA	precursor miRNA, consists only of the hairpin structure
pri-miRNA	primary miRNA, it is the mRNA from the *MIRNA* gene
PTGS	Post-transcriptional Gene Silencing
RdDM	RNA directed DNA Methylation
RDR	RNA-dependent RNA Polymerase
REF6	Relative of Early Flowering 6
RITS	RNA Induced Transcriptional Silencing complex
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic acid
RNAi	RNA interference

SAM	Shoot Apical Meristem
SAMe	S-Adenosyl Methionine
SE	Serrate
SET	Suppressor of variegation / Enhancer of Zeste / Trithorax
siRNA	small-interfering RNA
sRNA	small RNA
ssRNA	single-stranded RNA
tasiRNA	trans-acting small-interfering RNA
TGS	Transcriptional Gene Silencing
TrAP	Transcription Activation Protein
TrIP	TrAP-interacting protein

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Xiuren Zhang (chair and advisor), Hays Rye and Timothy Devarenne at the Department of Biochemistry and Biophysics, and Professor Jun-Yuan Ji at the Department of Molecular and Cellular Medicine.

Professor Xiuren Zhang provided the microarray data analyzed in Chapter 2 as well as the TrAP-stable transgenic plants. The raw data processing of the microarray data depicted in Chapter 2 were conducted in part by Dr. Xiuying Liu at the Chinese Academy of Sciences and were published in 2015 in eLife. The raw data processing of the bisulfite sequencing depicted in Chapter 5 were conducted in part by Dr. Zeyang Ma at the Department of Biochemistry and Biophysics and were published in 2015 in eLife.

The student completed all other work conducted for this dissertation independently.

Funding Sources

This work was supported by grants from the US National Science Foundation (NSF CAREER) (MCB-1253369), the US National Institutes of Health (R21AI097570) and the Welch foundation (A-1777) awarded to Professor Xiuren Zhang.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION*

Invasion of a host by a pathogen is a very complex process in which both organisms exhibit a splendid array of molecular weapons to fight for survival. Since viral pathogens rely on the host for replication, their weapons are targeted to the subversion of the host replication machinery and the restrain of the host defenses; hence, viruses are not only important because of their effects in the host, but they are also great tools for studying many important biological processes in the host.

1.1 RNA silencing

RNA silencing, also known as RNA interference (RNAi), co-suppression or quelling, was originally described in petunia flowers that were engineered to overexpress the enzyme chalcone synthase (CHS). The expectation was that overexpression of the CHS gene would produce purple flowers; alas, the transgenic plants were white because of the silencing of both the endogenous and the introduced copies of CHS (Napoli et al., 1990). About ten years later, research on the nematode *Caenorhabditis elegans* showed that injection of double-stranded RNA (dsRNA) caused gene silencing in a sequence-specific manner (Fire et al., 1998). These two findings paved the way to the discovery of a plant RNA-dependent RNA polymerase (RDR) able to produce double-stranded RNA from very abundant messenger RNAs (mRNAs) (Dalmay et al., 2001); thus, explaining the co-suppression of CHS in the engineered petunias, and unequivocally demonstrating

* Part of this chapter is reprinted with permission from "The functions of plant small RNAs in development and in stress responses" by Shengjun Li, Claudia Castillo-González, Bin Yu, and Xiuren Zhang. 2017. Plant J. *90*: 654–670. doi:10.1111/tpj.13444, Copyright [2017] by John Wiley and Sons. License number: 4134351448309.

the conservation of the RNA-mediated regulation of gene expression throughout eukaryotes. The RNAi is essential to regulate multiple biological processes, including stem cell development, maintenance and differentiation, stress response, symbiosis, and pathogenesis.

RNAi is elicited by small double-stranded RNAs (dsRNAs) that interfere with multiple steps of the informational flow in the cell; specifically, 20-30nt small RNAs (sRNAs) are loaded into an Argonaute (AGO)-containing protein complex to serve as guides for gene silencing. RNAi can occur at two stages: (1) Post-Transcriptional Gene Silencing (PTGS), which regulates the expression of target genes through an RNA-induced silencing complex (RISC) to direct translation inhibition or mRNA cleavage in a sequence-specific manner (Baulcombe, 2004; Li et al., 2013; Vaucheret, 2006). Notably, the same mRNA target can be subjected to both modes of PTGS; it has been proposed that translation inhibition is predominant at an early stage, and is followed by mRNA cleavage (Brodersen et al., 2008; Iwakawa and Tomari, 2013; Wilson and Doudna, 2013). (2) Transcriptional Gene Silencing (TGS), which regulates the expression of target genes through an RNA-induced transcriptional silencing (RITS) complex to direct DNA methylation and drive the conformation of heterochromatin, and precluding the transcription of target genes (Cui and Cao, 2014; Jones and Sung, 2014; Pikaard and Scheid, 2014; Zhang and Zhu, 2011).

Based on their biogenesis, there are two main classes of sRNA: (1) the micro RNAs (miRNAs), and (2) the small-interfering RNAs (siRNAs). In plants, all sRNAs have a 2-nt overhang and 2'-O-methylation at the 3'-ends (Li et al., 2005; Yu et al., 2005). The

methyltransferase HUA ENHANCER 1 (HEN1) methylates the 2'hydroxyl groups at the 3'-overhangs of the sRNA duplexes to protect them from trimming or tailing, and eventual degradation (Sanei and Chen, 2015; Yang et al., 2006b).

miRNAs originate from *MIRNA* genes which are transcribed by RNA Polymerase II (RNA Pol II) to produce primary miRNA transcripts (pri-miRNAs). Pri-miRNA is capped, polyadenylated, and folded onto itself to form a hairpin-like structure. The pri-miRNA hairpin, which contains an imperfect dsRNA stem and an ssRNA loop, is recognized and processed in the nucleus by an RNase III protein in a microprocessor complex. In plants, the core microprocessor excises the miRNA from the pri-miRNA in two consecutive steps, entailing the RNase III enzyme DICER-LIKE 1 (DCL1) and the RNA-binding protein HYPONASTIC LEAVES 1 (HYL1) (Figure 1). It was originally believed that the zinc-finger protein SERRATE (SE) was also part of the microprocessor (Iwata et al., 2013; Machida et al., 2011; Yang et al., 2006a), but recent studies have suggested that SE serves as a scaffold for the protein complexes that exert tight regulation on the processing of pri-miRNAs into miRNAs (Zhu et al., 2013). Notably, CPL1 (Manavella et al., 2012), NOT2 (Wang et al., 2013) and CDC5 (Zhang et al., 2013b) interact with the RNA Pol II and the microprocessor through SE, indicating the coupling of transcription and pri-miRNA processing. SE also interacts with the CAP-binding complex subunits CBP20 and CBP80, which are important for the pri-miRNAs into miRNAs (Gregory et al., 2008; Laubinger et al., 2008). Other accessory proteins, such as the RNA-binding TOUGH (TGH), DAWDLE, and many others, have also been reported

to bind to the pri-miRNA and control the microprocessor activity (Mateos et al., 2011; Szarzynska et al., 2009; Voinnet, 2009; Yang et al., 2006a; Zhu et al., 2013).

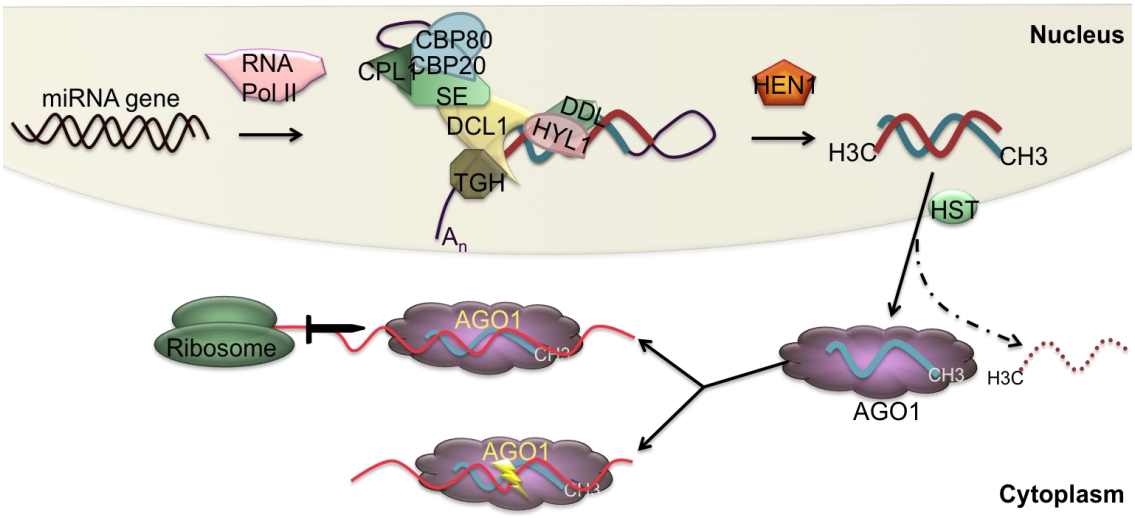


Figure 1. The miRNA biogenesis pathway.
RNA Pol II transcribes a *MIRNA* gene to produce a pri-miRNA transcript, which is subsequently processed by the DCL1/HYL1 microprocessor into 21-22bp miRNA/* duplexes. Then, HEN1 catalyzes the 2'-O-methylation of the miRNA/* duplexes at their 3'-ends to protect them from degradation. The mature miRNA/* duplexes are exported to the cytoplasm through HASTY; the miRNA (guide strand) is then loaded onto an AGO1-RISC to inhibit translation or to cleave the target transcript.

After miRNAs are produced (Figure 1) HEN1 methylates the miRNA duplexes in the nucleus (Yang et al., 2006b). The methylated miRNAs are then exported to the cytoplasm likely through the nuclear transportin HASTY (HST) (Park et al., 2005), where they are loaded into AGO1 as ssRNA, to activate the RISC and guide PTGS (Voinnet, 2009). In the miRNA duplex (miRNA/*), the AGO1-loaded strand is called "guide" or

miRNA, while the complementary strand is named "passenger" or miRNA* (Chen, 2005; Fang and Qi, 2016; Ji et al.; Meister, 2013; Voinnet, 2009; Wu, 2013). miRNAs can direct both mRNA cleavage and translation repression in a non-exclusive way, and are essential for plant development.

The small-interfering RNAs (siRNAs), on the other hand, are derived from perfect dsRNAs, which can originate from antisense transcripts, inverted repeats, viral genomes and transcripts, and from the activity of RDRs, among others. Notably, the expansion of the RDR, DCL, and AGO protein families contributed to the functional diversification of siRNAs in plants (Ahlquist, 2002; Chen, 2005; Tang et al., 2003; Tretter et al., 2008). Similarly to pri-miRNAs, a DCL protein processes the long dsRNAs into 21-24nt siRNAs, which are subsequently methylated by HEN1 and loaded into an AGO-containing silencing complex. However, unlike miRNAs, the biogenesis of siRNAs does not need to occur in the nucleus (Figure 2, Figure 3) (Ding, 2010; Meister, 2013; Pattanayak et al., 2013; Yang et al., 2006b). Of the four DCL proteins in Arabidopsis, DCL1 is mainly dedicated to the production of miRNAs, while siRNAs are predominantly generated by DCL2, 3 and 4.

There are three main classes of siRNAs, namely: heterochromatic-siRNAs (hetsiRNAs), natural antisense siRNAs (natsiRNAs), and secondary siRNAs, which are further classified into trans-acting siRNAs (tasiRNAs), phased-siRNAs (phasiRNAs), and epigenetically-activated siRNAs (easiRNAs) (Axtell, 2013; Borges and Martienssen, 2015).

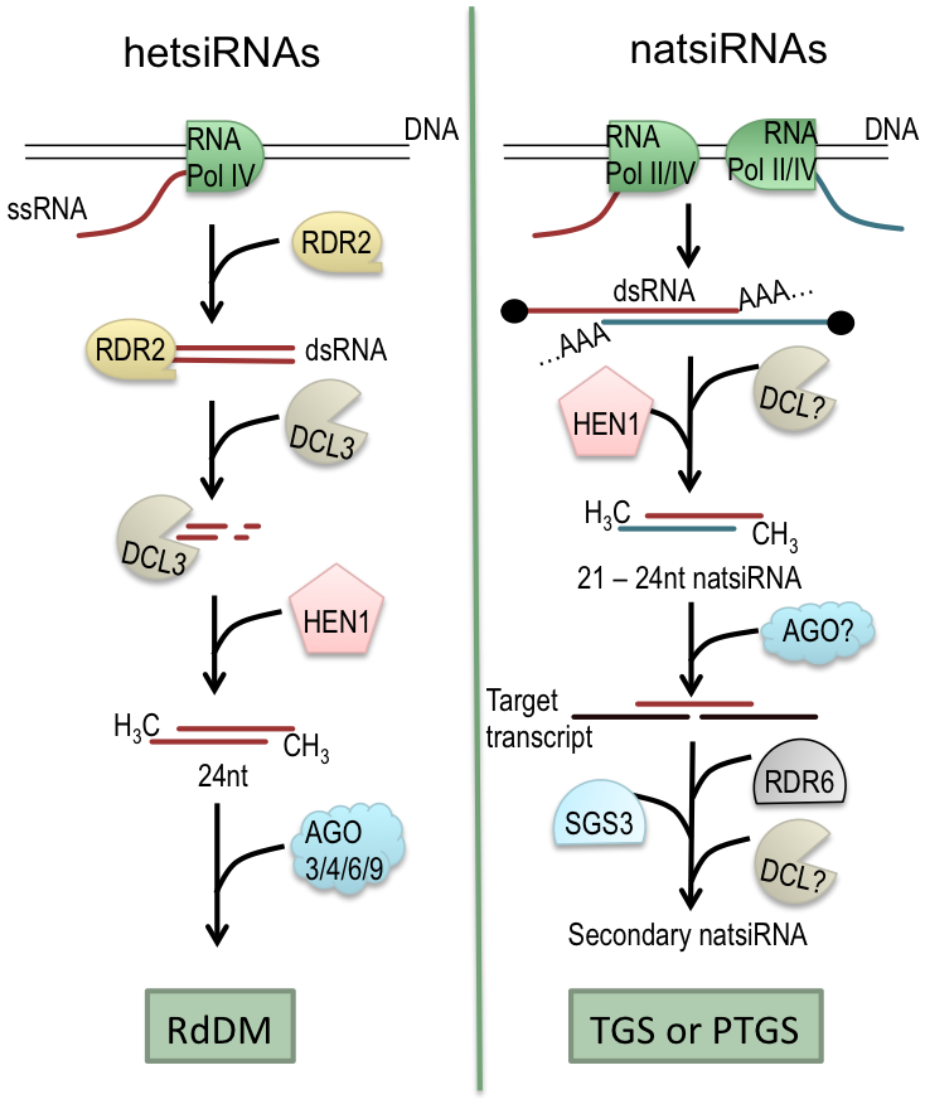


Figure 2. The primary siRNA biogenesis pathways.
There are two classes of primary siRNAs: hetsiRNAs (left panel), and natsiRNAs (right panel). hetsiRNAs are produced in the nucleus. RNA polIV synthesizes short transcripts to serve as substrates for RDR2 and produce short dsRNAs to be processed by DCL3 into 24nt hetsiRNAs and methylated at their 3'ends by HEN1. hetsiRNAs are loaded into an AGO3/4/6/9-containing RITS complex to direct TGS. natsiRNAs result of the transcription of complementary mRNAs, forming dsRNAs, which are processed by a DCL protein into 21-24nt natsiRNAs. After methylation by HEN1, natsiRNAs are loaded into an AGO-containing RISC complex to direct PTGS and subsequently trigger the secondary siRNA production by recruiting RNA polIV, RDR6, SGS3, and several other DCL proteins.

hetsiRNAs are the most abundant sRNAs, are 24nt long, and mediate the TGS of heterochromatic regions (i.e. pericentromeric regions and TE-rich regions) through a process called RNA directed DNA methylation (RdDM) (Matzke and Mosher, 2014; Pikaard and Scheid, 2014). Briefly, the plant-specific RNA polymerase IV (RNA polIV) synthesizes short transcripts that subsequently serve as substrates for RDR2 to produce short dsRNAs, which are processed by Dicer-like3 into 24nt hetsiRNAs (Figure 2). The hetsiRNAs are then loaded into an AGO3/4/6/9-containing RITS complex to recruit a de novo DNA methyltransferase to the DNA (Zhang et al., 2016). The methylated genome is recognized by chromatin remodeling proteins that catalyze the deposition of repressive marks to the nucleosomes (Chan et al., 2004; Havecker et al., 2010).

The natural antisense siRNAs, or natsiRNAs, are 21-24nt long and result from double stranded RNAs produced by complementary mRNAs in the cell (Figure 2). Those can occur in *cis*, when produced from overlapping regions of convergent transcripts; or in trans, when the complementary transcripts are produced from different genomic regions (Bologna and Voinnet, 2014; Vaucheret, 2006). Although the details of the natsiRNAs biosynthesis are not yet fully understood, it is widely accepted that environmental conditions or a developmental program induces them (Borges and Martienssen, 2015). As a general rule, one of the transcripts is constitutively expressed, while the other is induced (Borsani et al., 2005). The resulting dsRNAs will be substrates of a DCL protein into natsiRNAs, which are methylated by HEN1 and loaded in an AGO-containing RISC complex (Figure 2). Subsequently, the sRNA machinery, including the RNA Pol IV subunit NRPD1, the RNA-dependent RNA polymerase RDR6, SGS3, and several DCL

proteins (mostly DCL1 and DCL2) establish a reinforcement loop to produce secondary natsiRNAs and amplify the silencing signal (Figure 3) (Borsani et al., 2005; Martínez de Alba et al., 2013).

The secondary siRNAs mediate the amplification of the PTGS signal. Briefly, a mRNA becomes the substrate of an RDR to produce dsRNA, which is subsequently processed by a DCL protein and methylated by HEN1 to produce mature secondary siRNAs. The substrate mRNAs are typically aberrant mRNAs, mRNAs that have been cleaved by a miRNA-loaded RISC, or invading mRNAs (i.e. TEs, viroids, etc). There are three main classes of secondary siRNAs: tasiRNAs, phasiRNAs and easiRNAs (Figure 3).

The trans-acting siRNAs originate from a tasiRNA-precursor mRNA, which is encoded by a TAS gene. The Arabidopsis genome encodes for eight TAS genes grouped into four families. mRNAs produced from genes in the *TAS* families 1 and 2 are targeted by miR173 (Yoshikawa et al., 2016), while miR390 targets the mRNAs from *TAS3* genes (Endo et al., 2013; Fahlgren et al., 2006; Montgomery et al., 2008), and miR828 targets those of the *TAS4* family (Chen et al., 2010; Fei et al., 2013). Notably, miR390 is loaded almost exclusively onto AGO7-containing RISCs (Endo et al., 2013), while miR173 and miR828 are loaded to AGO1-RISCs. tasiRNAs are produced through dedicated machinery consisting of SGS3/RDR6/DCL4 that acts downstream of the miRNA-AGO cleavage of the TAS transcripts. In the case of the *TAS3* genes, AGO7 binds preferentially to miR390 directing the cleavage of tasiRNA-inducing transcript *TAS3* (Endo et al., 2013; Fahlgren et al., 2006; Liu et al., 2009; Montgomery et al., 2008; Zhang et al., 2013a).

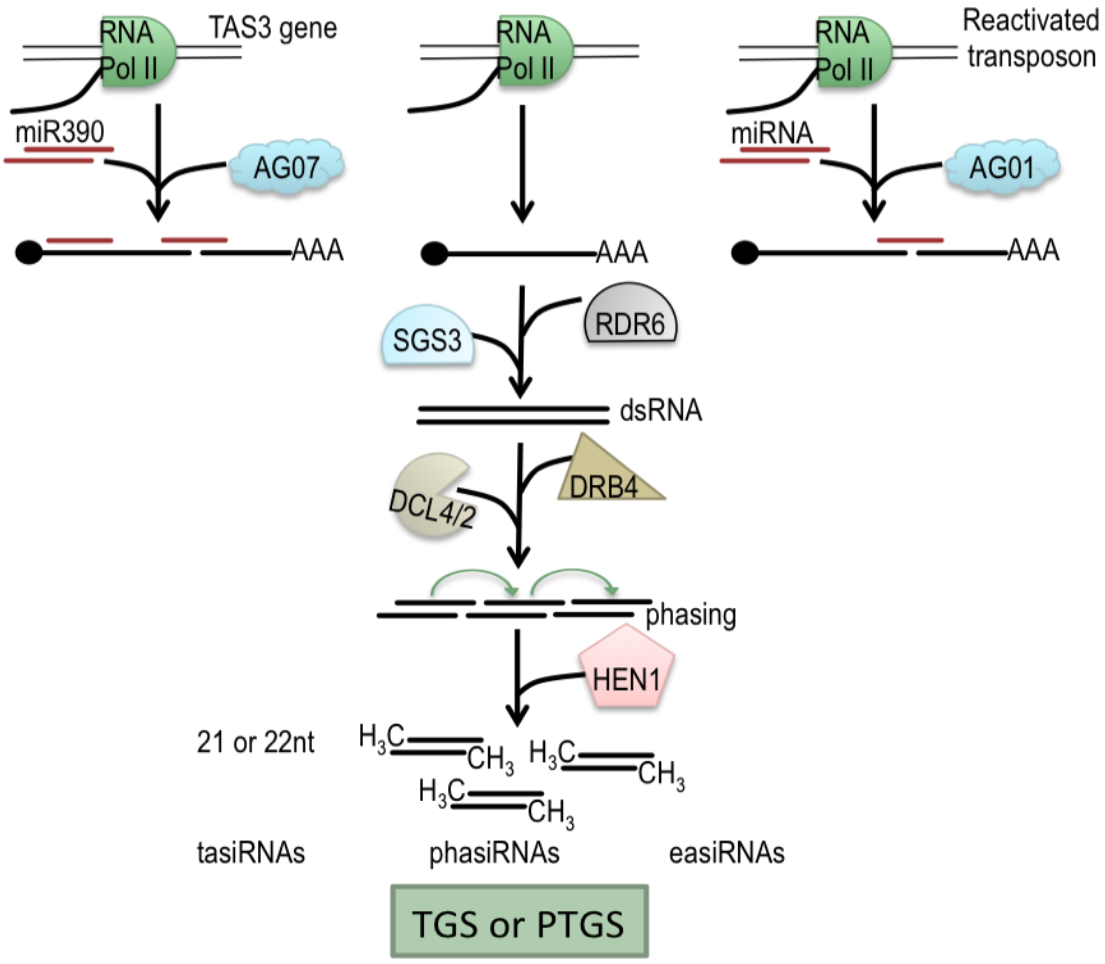


Figure 3. The secondary siRNAs biogenesis pathway.
RDR6 and SGS3 synthesize the dsRNA that serves as substrate for DCL2 and DCL4, which produce 22 or 21nt secondary siRNAs, respectively. Both DCL2 and DCL4 are highly processive enzymes which produce phased siRNAs by consecutive slicing along the dsRNA substrate. According to their precursor RNAs, there are three main classes of secondary siRNAs: tasiRNAs (top left), easiRNAs (top right) and phasiRNAs (top center). tasiRNAs are produced from a *TAS* gene which encodes a miRNA substrate. The phase tasiRNAs are produced fom the miRNA-RISC cleaves TAS transcript and act in the silencing of other genes. easiRNAs result from the targeting of miRNAs to transcripts of active retrotransposon, while phasiRNAs can be produced from coding and non-coding RNA and are independent of miRNA-directed cleavage.

The cleaved *TAS* products, stabilized by the Zn-finger protein SGS3, serve as substrates for RDR6 to make dsRNAs. The resultant dsRNAs are further processed by DCL4 or DCL1 into 21-22nt ta-siRNAs that are associated with AGO1-RISCs to down-regulate the expression of their target genes in trans (Figure 3) (Allen et al., 2005; Fahlgren et al., 2006; Montgomery et al., 2008; Peragine et al., 2004).

Due to the processivity of the DCL2 and DCL4 enzymes, the secondary siRNAs are phased; that is, they are produced consecutively every 21 or 22 nucleotides depending on whether they are processed by DCL4 or DCL2, respectively (Allen et al., 2005; Fei et al., 2013). Unlike tasiRNAs, the phasiRNAs are independent of miRNA-directed cleavage and do not have a clear function. The precursor phasiRNA can be coding or non-coding RNA (Figure 3).

Similarly to tasiRNAs, the easiRNAs result from the targeting of miRNAs to transcripts of active retrotransposons; particularly 22nt miRNAs loaded onto an AGO1-RISC (Chen et al., 2010). Mainly DCL4 catalyzes the processing of the dsRNA into 21nt easiRNAs; however, DCL2 is recruited to those substrates to produce 22nt easiRNAs in the absence of DCL4 or when the transcript levels surpass the catalytic capacity of the DCL4 enzyme (Figure 3). Notably, DCL3 can also be recruited to these substrates and produce 24nt easiRNAs that can direct *de novo* DNA methylation and establish TGS (Borges and Martienssen, 2015; Clavel et al., 2016; Creasey et al., 2014; Cui and Cao, 2014). Other invasive nucleic acids, such as viruses and viroids, as well as transgenes, are also substrates of DCL2 and DCL4 (Mlotshwa et al., 2008, 2010). Indeed, DCL2 and DCL4 play a major role in antiviral defense and produce virus-derived siRNAs (vsiRNAs)

from dsRNA produced by RDR1, which can be further amplified by RDR6 (Deleris et al., 2006; Mlotshwa et al., 2008; Qu et al., 2008; Wang et al., 2010b).

There has been significant progress in the understanding sRNA-mediated gene regulation; however, we still ignore the factors that regulate the efficiency of sRNA-mediated target regulation, the determinants for translational repression or transcript cleavage, and the specific contribution of each of these mechanisms to gene regulation.

1.2 RNAi in development

All plant tissues and organs arise from structures containing pluripotent stem cells called meristems, which can be determinate or indeterminate depending on whether they are exhausted during development. Plants specify their tissues from two main meristems, the shoot apical meristem (SAM) and the root apical meristem (RAM). The SAM is established during embryogenesis to maintain the population of stem cells, to provide cells to organ primordia, and to specify the primary axis of growth. After embryogenesis, axillary meristems produce branches and secondary growth axes. Finally, in the transition to the reproductive phase indeterminate shoot meristems develop into determinate inflorescence meristems (IM) (Wang *et al.*, 2016; Soyars *et al.*, 2016). Here I discuss the role of small RNAs (sRNAs) in the establishment, maintenance, and maturation of shoot meristems.

In plants, the SAM contains three radial layers (L1-L3) (Figure 4). L1 (epidermis) and L2 (sub-epidermal layer) are the two outer layers that are both one cell thick, whereas L3 (corpus) is a multilayer of cells that make up the rest of the internal cells. These layers of cells also constitute three different zones in the SAM: the Peripheral Zone (PZ) which

forms the lateral primordia; the Central Zone (CZ) which maintains the population of pluripotent cells in a subregion called Organizing Center (OC); and the Rib Zone (RZ) which produces the stem. The OC is made exclusively from L3 cells (Soyars et al., 2016). In the Arabidopsis SAM, stem cell fate is determined by the homeodomain transcription factor WUSCHEL (WUS) (Mayer et al., 1998). WUS protein is synthesized in the OC and migrates into the CZ to activate the negative regulator, *CLAVATA3 (CLV3)* (Brand et al., 2000; Schoof et al., 2000; Zhou et al., 2015a). CLV3 further restricts WUS to the OC via a receptor kinase signaling cascade (Brand et al., 2000; Schoof et al., 2000). This local feedback loop controls development and maintenance of stem cell population. A family of KNOX I transcription factors, mainly, SHOOT MERISTEMLESS (STM), and BREVIPEDICELLUS (BP), positively regulates *WUS* expression. In monocots, the establishment of stem cells also requires the transcription factors of the SAM-restricted KNOX I family; the founding member of this family is the maize *KNOTTED 1* (*KN1*) gene, and it is necessary for SAM establishment and maintenance (Bolduc et al., 2012). Unlike dicots, SAM regulation is likely independent of *WUS*: while no *wus* mutant has been found in maize, the rice *WUS* ortholog *TAB1* only seems to affect axillary meristem formation (Tanaka et al., 2015). Thus, while *KNOX I* regulation is sufficient for SAM development in monocots, *WUS* and *KNOX I* are critical for SAM development in eudicots. Numerous genetic pathways that control SAM development, converge in the interplay with these two key transcription factors (Galli and Gallavotti, 2016; Zhou et al., 2015a).

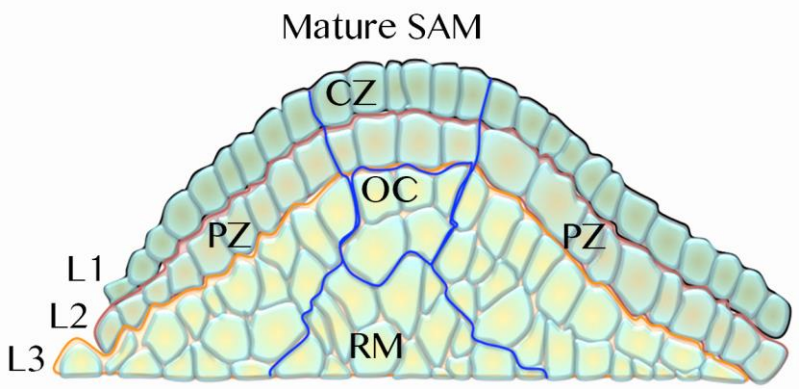


Figure 4. Shoot apical meristem organization.
The meristem is organized in concentric layers of different genetic lineages, named L1 or protoderm, L2 or subepidermal tissue, and L3 or inner corpus. The interactions among them give rise to the Peripheral Zone (PZ), which forms the lateral primordia; the Central Zone (CZ), which maintains the population of pluripotent cells, is made exclusively from L3 cells.

SAM development is regulated by plant hormones. It is known that a high ratio of cytokinin (CK) to auxin ratio is critical for meristem establishment and maintenance (Vanstraelen and Benková, 2012; Azizi *et al.*, 2015). While CK enhances the expression of *WOX* and *KNOX I* genes, auxin promotes cell growth; however, auxin also induces CK accumulation through the transcriptional inhibition of *ARR7* and *ARR15*, two CK repressors. In Arabidopsis, WUS directly represses *ARR7/15*, whereas STM and BP upregulate expression of the CK synthase *IPT7*, hence establishing a positive feedback loop to maintain the meristem (Jasinski et al., 2005; Leibfried et al., 2005; Yanai et al., 2005).

The effect of auxin on SAM development is exerted through transcriptional regulators in the Auxin Response Factor (ARF) family, and their regulation is paramount

in the establishment of embryonary stem cells (Seefried et al., 2014). The induction of somatic embryos requires the accumulation of miR167, which targets and directs the patterning of *ARF6* and *ARF8*. Loss of *ARF6* and *ARF8* results in somatic embryo arrest (Su et al., 2016). SAM establishment is also regulated by *ARF2*, *ARF3* and *ARF4* but they are downregulated by *TAS3*-derived tasiRNAs. Notably, the accumulation of *TAS3* transcripts is regulated during embryogenesis. Throughout the globular stage, *TAS3* is mostly detected in the apical region; by the torpedo stage *TAS3* accumulation is restricted to the adaxial region; and in the bent-cotyledon stage it is expressed mainly in the SAM (Liu et al., 2009). Such spatio-temporal patterning of *TAS3* underscores the regulatory role of tasiRNAs in meristem development through the modulation of *ARFs* (Figure 5) (Fei et al., 2013; Petsch et al., 2015; Rajeswaran and Pooggin, 2012). Importantly, the critical regulatory role of the tasiRNA pathway in SAM establishment and maintenance is conserved in mono- and eudicots. The *SGS3* maize ortholog *LEAF BLADELESNESS 1* (*LBL1*) regulates the meristem master regulator *KN1* (Nogueira *et al.*, 2009); and loss-of-function mutants in the rice orthologs of *RDR6*/*SHOOTLESS2* (*SHL2*), *AGO7*/*SHOOT ORGANIZATION 2* (*SHO2*), and *DCL4*/*SHO1* completely lack SAM (Nagasaki et al., 2007).

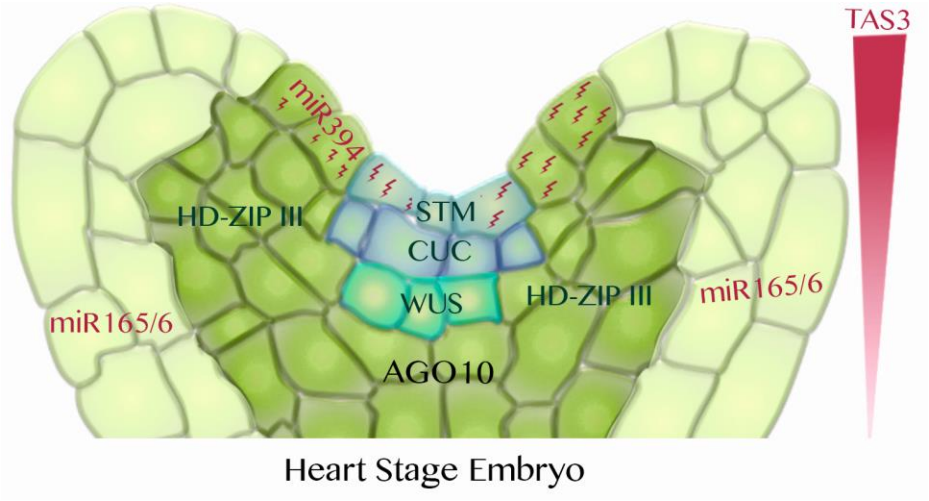


Figure 5. Functional model of miR165/166 and miR394 at the heart stage of embryogenesis. miR394 is strongly expressed from the L1 layer and moves inwards to regulate the downstream activity of WUS and defining the identity of the inner layers of the SAM. AGO10 restricts miR165 and miR166 to the outer and abaxial side of the embryo. This enables expression of HD-ZIP III transcription factors and the correct patterning of the SAM and vascular tissues.

Concurrent with the establishment of stem cells, the SAM surges from a boundary zone during embryogenesis. The boundary zones not only propitiate a local environment for meristematic activity but also separate pluripotent cells from the regions of active cell differentiation. The organ boundaries are set by several groups of transcription factors including NAC, MYB, LBD, and GRAS families (Wang *et al.*, 2016). The restriction of the *NAC* genes *CUP SHAPED COTYLEDON1* and *2 (CUC1 and CUC2)* in the boundary regions is accomplished through miR164 (Figure 5, Figure 6). miR164, which directly targets *CUC1* and *CUC2*, is accumulated in the PZ so that *CUC1* and *CUC2* transcripts are degraded in the PZ but not in the boundary regions (Figure 4, Figure 5) (Larue et al., 2009; Nikovics et al., 2006). In a similar way, miR164 is also essential in the formation of axillary meristems throughout the plant (Fouracre and Poethig, 2016; Laufs et al., 2004;

Wang et al., 2016). Besides regulating the boundary establishment, CUC1 and CUC2 act with the MYB transcription factor LOF1 to further induce STM and SAM formation during the late globular stage (Aida et al., 1999). Mutants lacking both CUC1 and CUC2 activities fail to establish the SAM and die during embryogenesis (Mallory et al., 2004; Takada et al., 2001).

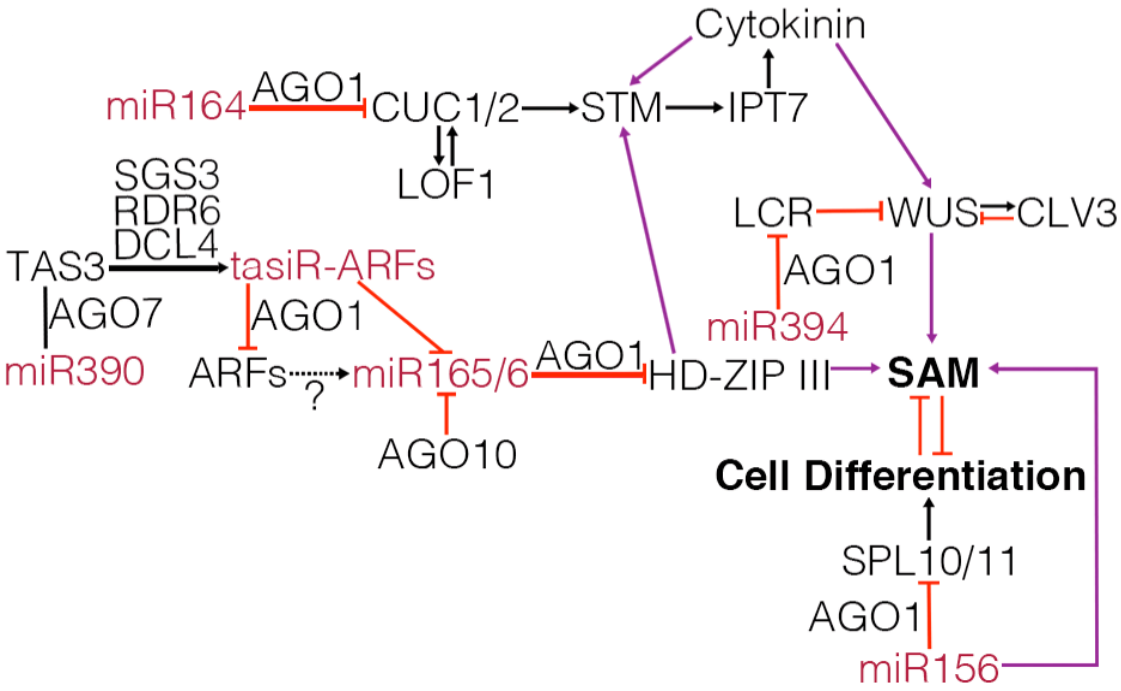


Figure 6. Multiple sRNA regulatory modules converge in the regulation of the SAM.
Black arrows represent validated positive regulation; dotted black arrow represents a hypothesized positive regulation; red blunt arrows represent validated downregulated targets; purple arrows indicate downstream positive regulation.

The abaxial/adaxial regulatory network originated from the leaf primordium impacts the SAM establishment. The network consists primarily of two sets of transcription factors; and at least one set regulates the SAM through modulation of expression of the downstream *KNOX I* and *WOX* genes (Roodbarkelari and Groot, 2016). One set includes the KANADI and YABBY transcription actors and promotes abaxial leaf identity, whereas the other set includes the class III HOMEODOMAIN LEUCINE ZIPPER (HD-ZIP) transcription factors (specifically PHABULOSA (PHB), PHAVOLUTA (PHV), REVOLUTA (REV), CORONA (CNA), and ARABIDOPSIS HOMEOBOX GENE 8 (ATHB8)) and promotes adaxial leaf fates. Interestingly, the spatial restriction of HD-ZIPIII adaxial factors is mediated by miRNAs that are also spatiotemporally distributed (Figure 5, Figure 6) (Williams *et al.*, 2005; Sakaguchi and Watanabe, 2012; Roodbarkelari and Groot, 2016).

The first indication of miRNA involvement in SAM maintenance was the discovery that *WUS* depends on AGO10 (also known as ZWILLE) for the activation of its downstream target *CLV3* in Arabidopsis (Bohmert et al., 1998; Lynn et al., 1999; Tucker et al., 2008). In Ler ecotype background, *ago10* mutants can initiate but not maintain the SAM (Lynn et al., 1999; Moussian et al., 1998). This defect is clearly ecotype-specific as Col-0 *ago10* mutant rarely displays shoot meristem termination. In the SAM, AGO10 recognizes structural properties in the duplexes of miR165/6 and their complementary strands (*), and outcompetes their binding by AGO1. However, unlike AGO1, AGO10 decoys miR165/6 and protects the *HD-ZIP III* transcripts by mechanisms yet to be unveiled (Zhang and Zhang, 2012; Zhu *et al.*, 2011; Zhou *et al.*, 2015; Liu *et al.*, 2009).

Arabidopsis has nine loci able to produce miR165/6: *MIR165a*, and *b*, and *MIR166a-g*. Despite their individual contributions to the accumulation of miR165/6, only four genes seem to be critical for the SAM development (*MIR165b* and *MIR166a*, *b*, and *g*) (Miyashima *et al.*, 2013; Zhou *et al.*, 2015). Such differences may be attributed to variation in the promoter activities (Miyashima et al., 2013); alternatively, the secondary structures of some of primary miR165/6 transcripts (pri-miR165/6) cause poor production of miR165/6, as shown for pri-miR166c, d, and e (Zhu et al., 2013).

Expression of *AGO10* is detected in embryo as early as the 8-cell stage of embryogenesis. Unlike the ubiquitously expressed *AGO1*, *AGO10* continues to accumulate only in the adaxial region and throughout the SAM, where it neutralizes mobile miR165/6 that are produced in the abaxial sites and enables the HD-ZIP III function in the adaxial regions (Figure 5, Figure 6) (Lynn *et al.*, 1999; Zhou *et al.*, 2015). In rice and maize, miR166 also regulates *HD-ZIP III* genes and SAM maintenance (Itoh et al., 2008; Juarez et al., 2004). Rice *AGO10* ortholog *OsPNH1* is also necessary for SAM maintenance and organ development, and exhibits a expression pattern similar to Arabidopsis AGO10; however, OsPNH1 regulation of miR166 or its mode of action are still unknown (Nishimura et al., 2002).

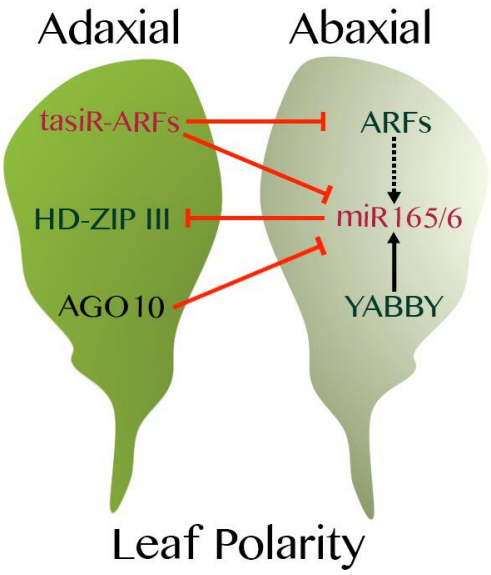


Figure 7. Interplay of tasiRNA and miRNA pathway in the determination of leaf polarity. Black arrows represent validated positive regulation; dotted black arrow represents a hypothesized positive regulation; red blunt arrows represent validated downregulated targets.

Reciprocally to the AGO10-mediated spatial decoy of miR165/6, determination of organ polarity in the PZ requires AGO1 to execute the tasiR-ARFs silencing and consequential down-regulation of miR165/6 accumulation (Figure 7) (Sakaguchi and Watanabe, 2012). Importantly, the rice tasiRNA machinery, SHL2/RDR6, SHO1/DCL4, and SHO2/AGO7, negatively regulates miR166 accumulation (Nagasaki et al., 2007). Likewise, maize loss-of-function mutants of *SGS3/LBL1* have abaxialized leaves as a result of ectopic distribution of pri-miR166 (Nogueira et al., 2007); and lack of *AGO7/RGD2* results in increased miR166 accumulation, although with normal leaf polarity (Douglas et al., 2010). In Arabidopsis, while *TAS3* expression is localized in the SAM apical region and adaxial side of the leaf primordia, *ago10 rdr6* and *ago10 ago7*

double mutants exhibit a stronger SAM phenotype and an even higher accumulation of miR165/6 compared to *ago10* single mutant. This observation further suggests that *AGO10* and tasiR-ARFs act in parallel to regulate miR165/6 and *HD-ZIP III* genes (Liu et al., 2009). Since AGO7, SHO2 and RGD2 preferentially bind to miR390 to target *TAS3* and produce tasiR-ARFs, it is likely that the target *ARFs* are positive regulators of *MIR166* genes (Figure 7). This would be in accordance with *HD-ZIP III* regulation by auxin (Itoh et al., 2008) and the presence of ARF-regulatory sequences in the promoters of several *MIR166* genes in maize and *MIR166c* in Arabidopsis (Nogueira and Timmermans, 2007). Alternatively, and consistent with the opposite leaf phenotypes of *lbl1* and *rgd2* in maize, yet undiscovered tasiRNAs might target pri-miR166 transcripts impairing miR166 accumulation. Additionally, the abaxial-determining genes YABBY, regulate the *MIR165A* gene expression, resulting in the rigid determination of the adaxial–abaxial boundary in leaf primordia (Figure 7) (Tatematsu et al., 2015).

The study of the spatiotemporal regulation of *HD-ZIP III* and their regulators miR165/166 during embryogenesis highlights that non-cell-autonomous function of sRNAs underlies SAM maintenance and organ patterning (Seefried *et al.*, 2014; Nodine and Bartel, 2010; Zhou *et al.*, 2015). sRNAs move cell-to-cell through plasmodesmata for approximately 15 cells in a path of decreasing concentration (Marín-González and Suárez-López, 2012; Sparks *et al.*, 2013). Indeed, the requirement of L1 cells for the meristem determination engages a mobile L1-derived miRNA, miR394 (Figure 4, Figure 5, Figure 6). The role of miR394 in SAM determination was discovered in a genetic-screening for enhancers of an *ago10-1* mutant in Col-0 background; there, a mutant that dramatically

increased shoot meristem termination was recovered and identified as a mutant in *MIR394b* (Knauer et al., 2013). The *ago10-1 mir394b-1* double mutant displays a single leaf-like organ or a filamentous structure at the stem cell position, reminiscent of strong *ago10* mutant alleles in Ler background. miR394 appears to regulate SAM development through modulating downstream *WUS* activity. As mentioned above, expression of *CLV3* in the CZ depends on the expression and mobility of WUS from the OC (Brand et al., 2000; Schoof et al., 2000); However, *mir394b-1 ago10-1* mutants lacked CLV3 even when the WUS expression domain expanded. miR394 targets transcripts of the F-box gene *LEAF CURLING RESPONSIVENESS (LCR)*, and its function is proposed to degrade proteins through the 26S proteasome. Thus, lack of miR394 presumably allows over-accumulation of LCR in the meristem and causes degradation of an unknown WUS cofactor necessary to induce *CLV3* in the CZ (Figure 6). miR394 moves inwards from the L1 layer, acting as a polarizing signal that confines the stem cells to the OC, defining the identity of the inner layers of the SAM (Figure 4, Figure 5) (Knauer et al., 2013).

sRNAs also govern the branching from the axillary meristems. Shoot branching in plants takes place by the formation of new meristems in the axils of leaves, which develop into secondary axes of growth and axillary buds. *CUC* and *LOF1* genes are required for the boundary formation between the stem and the leaf, so that the two organs can continue with their developmental program. However, CUC and LOF1 also promote the expression of the meristem initiators *WUS* and *STM*. This mechanism is also conserved in monocots were expression of the rice *WUS* and *STM* orthologs, *TAB1* and *OSH1*, are necessary for axillary meristem formation and branching (Tanaka et al., 2015). The CUC transcription

factors also directly regulate the expression of the *LATERAL SUPPRESSOR* (*LAS*) genes in Arabidopsis, tomato and rice. The *LAS* genes are putative transcription factors belonging to the GRAS family, which specifically regulate the initiation of axillary meristems during the vegetative growth phase. miR164 controls branching through the regulation of the *CUC* genes and consequent repression of downstream *LAS* genes. This was evidenced by the abolition of axillary meristem formation in plants overexpressing miR164, and the development of accessory buds in leaf axils in *mir164* mutants and plants expressing miR164-resistant alleles of *CUC1* or *CUC2* (Bustamante et al., 2016; Larue et al., 2009; Liu et al., 2014c; Raman et al., 2008). Other miRNAs have also been involved in the regulation of branch formation. While miR171c prevents branching through targeting the GRAS transcription factors of the SCARECROW family (Wang et al., 2010a), miR156 promotes branching by targeting the *LAS* repressors *SPL9* and *SPL15* (Tian et al., 2014).

The shoot meristems dictate the formation of all aerial organs, including leaves, stems and flowers. miRNAs also regulate leaf growth and senescence (Figure 8). An exemplified case is the regulation of class II TEOSINTE BRANCHED1/CYCLOIDEA/PCF (TCP) transcription factors by miR319 (Efroni et al., 2008; Li et al., 2012; Nath et al., 2003; Palatnik et al., 2003). Arabidopsis has eight class II *TCP* genes, five of which (*TCP2, 3, 4, 5, 10,* and *24*) are the direct targets of miR319. Arabidopsis also contains three *MIR319* genes and they are expressed largely through non-overlapping regions (Reviewed in Lopez *et al.*, 2015). Notably, TCP3 is a transcriptional activator of miR164, *AS1*, and auxin response repressor *SHY2*. Loss of miR319 or increase

in *TCP3* accumulation leads to down-regulation of *CUC*, *KNOX*, and *WUS* genes; thus promoting cell differentiation. *TCP* genes also regulate cell growth by modulating the auxin accumulation and response. For example, TCP3 represses the auxin efflux genes *PIN1*, *PIN5*, and *PIN6*, and also targets the *ARF* regulators, *TAS3* and *AGO7* (Koyama et al., 2010). Furthermore, TCP4 protein regulates CK response and cell proliferation through the SWI/SNF chromatin remodeler BRM/CHR2. This observation suggests direct involvement of TCP proteins in chromatin regulation (Efroni et al., 2013; Schommer et al., 2014).

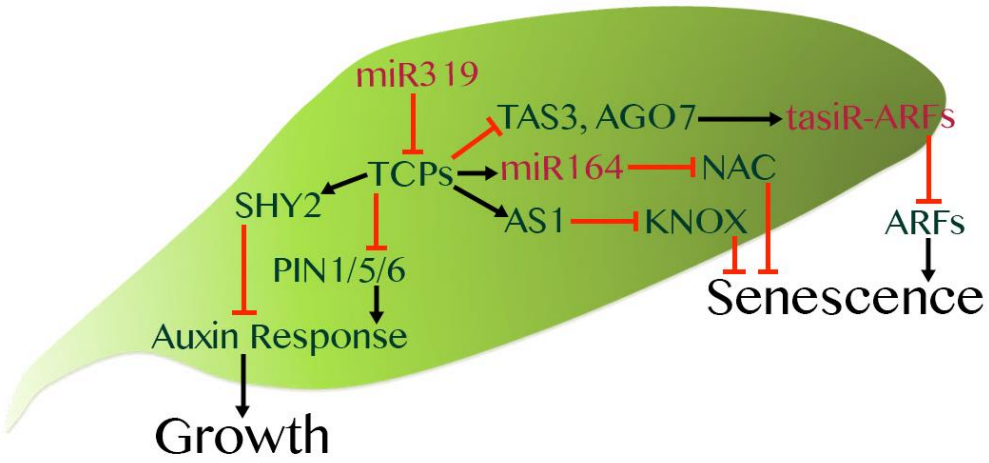


Figure 8. Convergent sRNA pathways determine the balance of cell proliferation, differentiation and senescence.
Black arrows represent validated positive regulation; red blunt arrows represent validated downregulated targets.

sRNAs also regulate plant developmental transition from juvenile to adult (Figure 9) (Xue et al., 2014). As understood today, such regulation consists primarily of the interplay of two miRNAs, miR156 and miR172. miR156 targets members of SQUAMOSA PROMOTER BINDING-LIKE (SPL) transcription factors (Xie, 2006), which are positive regulators of phase change and flowering time. Overexpression of miR156 delays flowering time (Zhang et al., 2011a), while the expression of miR156-resistant SPL transcripts induces early flowering. Furthermore, SPL9 acts as a transcriptional activator of *MIR172* genes, which in turn prevents expression of *APETALA 2* (*AP2*) transcription factors (Aukerman and Sakai, 2003; Chen, 2004; Zhu et al., 2009). The AP2 proteins regulate flowering time, organ identity and floral meristem fate. Arabidopsis has four AP2 transcription factors, TOE1, TOE2, SCHNARCHZAPFEN, and SCHLAFMUTZE. Overexpression of any *AP2* gene delays flowering, while overexpression of miR172 or loss-of-function mutation in any of the *AP2* genes produces an early flowering phenotype. As a surrogate pathway, miR169 can replace the function of *AP2* genes in Petunia and Antirrhinum because miR169 targets a family of NF-YA transcription factors, which regulate expression of downstream flowering genes (Cartolano et al., 2007). The decision to transit from vegetative to reproductive growth depends on the gradients of miR156 and miR172 (Wu *et al.*, 2009; and reviewed in Teotia and Tang, 2015; Nag and Jack, 2010); this was first evidenced in the maize dominant mutant *Corngrass* (*Cg*). *Cg* mutant displays enhanced expression of two tandem *MIR156* genes, and thus have lower accumulation of miR172. As such, the mutant retains juvenile features even in the reproductive phase (Chuck et al., 2007). Such phenotype not only

involves miRNAs in the phase change, but also uncouples the juvenile-to-adult transition from the vegetative-to-reproductive growth. The expression of miR156 originates in the SAM and is highest at the early stages of growth, but it declines as the plant ages. Conversely, the levels of miR172 rise and the reproductive phase transition occurs (Reviewed in Zhu and Helliwell, 2011). Recent works indicate that nutrient availability (i.e. sugar) plays a major role in the accumulation of miR156 and might serve as an environmental cue underlying phase transition (Yang et al., 2013; Yu et al., 2013).

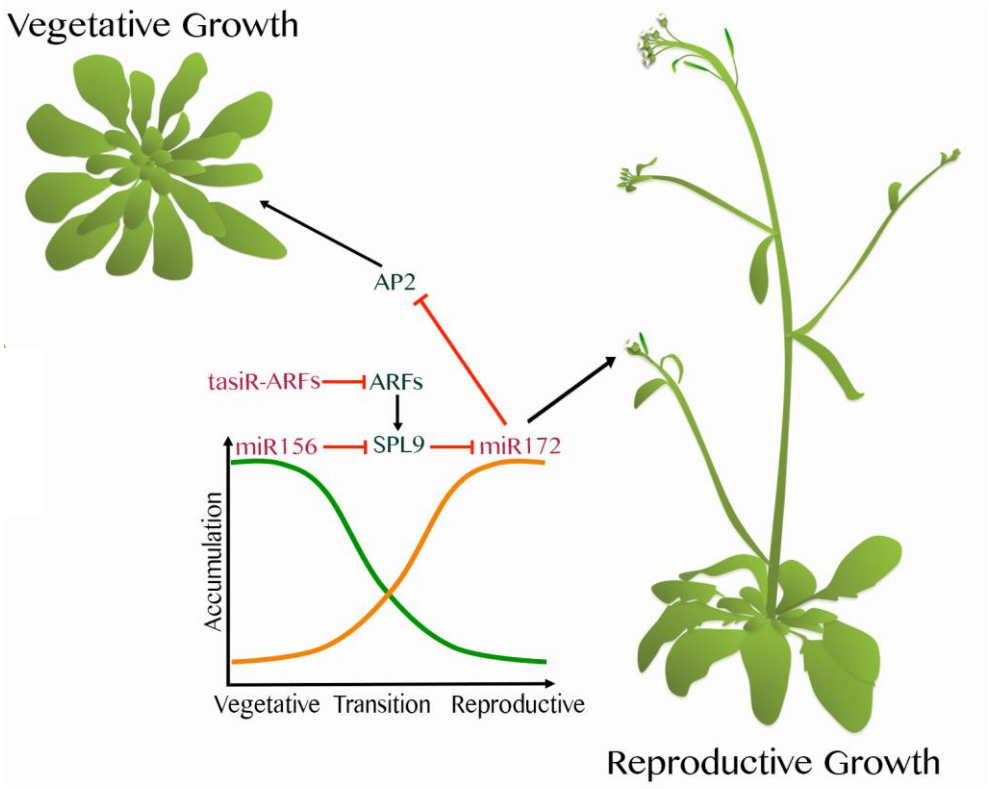


Figure 9. miR156 and miR172 modulate the juvenile-to-adult and vegetative-to-reproductive phase transitions.
Black arrows represent validated positive regulation; red blunt arrows represent validated downregulated targets

This kind of interplay between miRNA pathways seems to be a general rule and not the exception in plant development. In an elegant study in *Cardamine hirsuta*, it was evidenced that miR319-regulated TCPs compete with miR156-regulated SPLs for interaction with miR164-regulated CUC transcription factors. Such interplay determines leaf complexity as a function of plant age (Rubio-Somoza et al., 2014). Interestingly, miR156 not only regulates phase transition but also enables the appropriate embryo development. Involvement of miR156 in embryonic development appears to be through preventing the expression of transcription factors that promote differentiation later in the plant life cycle. This has been shown in the molecular characterization of the embryogenesis of *dcl1* mutants in Arabidopsis, where the expression of the miR156-regulated *SPL10* and *SPL11* genes is abnormally induced in embryos as early as 8-cells stage; and this up-regulation is largely responsible for the phenotypic defects in the embryo patterning in *dcl1* the mutant (Nodine and Bartel, 2010).

miRNAs also regulate seed maturation. The seed maturation program only takes place during late embryogenesis, and repression of the process is essential for seedling development. In Arabidopsis, *LEC2* is a key regulator of seed maturation; and its transcription is directly regulated by *PHB* and *PHV*. Through a sensitized genetic screening of suppressors of seed maturation in vegetative growth stage, a weak allele of *ago1* mutant was recovered. The repression function of *AGO1* in seed maturation program is through miR165/6 as AGO1 binds to miR165/166 to silence *PHB* and *PHV* transcripts, leading to ectopic expression of *LEC2* mRNA and activation of the downstream cascade of seed maturation (Tang et al., 2012). In maize, a severe growth-impaired mutant, *fuzzy*

tassel (*fzt*), was also identified. Loss-of-function *fzt* mutant contains a missense mutation in the RNase IIIa domain in DCL1 and negatively affects the accumulation of miRNAs (Thompson et al., 2014).

Transcriptome characterization of loss-of-function allelic series of genes involved in sRNA biogenesis uncovered an unforeseen depth of the regulation exerted by these small molecules. sRNAs are involved in the establishment and maintenance of the pluripotent state of cells, and also in the regulation of cell fate at the organ primordia (Nodine and Bartel, 2010; Seefried et al., 2014). Furthermore, sRNAs can control their targets by mRNA cleavage and/or translation inhibition, further increasing the complexity of the regulation exerted. Of note, miRNA-mediated translational inhibition occurs at the endoplasmic reticulum; and this regulation entails the membrane-bound protein ALTERED MERISTEM PROGRAMMING 1 (AMP1), which interacts with the AGO1-miRNA complex (Li et al., 2013). Interestingly, the *amp1* mutant was initially recovered in a mutagenesis screening for genes that affected leaf morphogenesis in Arabidopsis. Loss-of-function *amp1* allele develops an extraordinarily large shoot meristem during the globular embryo stage, well before than wild-type embryos (Conway and Poethig, 1997). *amp1* mutants also exhibit enlarged leaf primordia, leaves with striking similarities to cotyledons, ectopic stem cells niches, abnormal hormone regulation, seed dormancy, and many other pleiotropic phenotypes (Huang et al., 2015). Such pleotropic effect suggests a much broader effect of sRNA-mediated translation inhibition in the overall plant growth and development.

The detailed study of *MIRNA* gene families and their targets has substantially

increased our understanding of the regulation of development and also underscored the complexity of such intertwined regulatory systems. Numerous questions still need to be addressed concerning SAM control in particular and sRNA function in general. How do tasiR-ARF regulate miR165/6? How is the expression of miR156 regulated? Indeed, how is the expression of the MIRNA genes regulated? So far, it is possible to envision the gradients of sRNAs in the meristem functioning in a similar fashion as morphogens in animals (Skopelitis et al., 2012). However, it is unknown if mobility is an inherent characteristic of sRNAs. To date, few mobile sRNAs have been characterized, but this could be due to lack of sensitivity of our current methods. On the other hand, if not all sRNAs move, then what makes an sRNA mobile? Furthermore, sRNA gradients seem to determine a myriad of processes including but not limited to organ polarity and phase transition; nonetheless, in some cases it appears that the sRNA movement is detrimental for the recipient organ function (i.e. miR165/6 in SAM) and that complex adaptive mechanisms have been put into place to quench their effects (i.e. AGO10).

1.3 sRNAs as regulators of biotic interactions

Plants are the primary producers in the ecosystem; thus, constituting the base of the trophic chain and serving as food sources for microbes, invertebrates and higher organisms alike. Such trophic interactions can be established through parasitism, herbivory, or mutualism, and must be regulated for the plant survival. The deployment of defense or special developmental programs implies a major deviation of resources otherwise allocated to growth (Tian et al., 2003); therefore, the decision to fight or enable a biotic interaction requires an accurate calculation based on the environmental conditions,

and the identification of the interacting organism. Here I discuss the underlying roles of sRNAs as key regulators of the defense response and propose their role as the "lingua franca" in biotic transactions (Li et al., 2017).

1.3.1 RNAi as an immune system against invasive nucleic acids

RNA silencing is elicited by dsRNAs that interfere with multiple steps of the informational flow in the cell; specifically, sRNAs affect translation and stability of target mRNAs; sRNAs can also direct DNA methylation, therefore precluding transcription of target genes. Because dsRNAs can result from intermediates in RNA virus replication, highly structured RNA virus genomes, viral transcripts, repetitive DNA sequences or from transposable elements (TEs), it is not surprising that RNAi has evolved as an efficient mechanism to prevent the proliferation of invasive nucleic acids. Viruses, viroids, satellite RNAs (sat-RNAs), and TEs are the most common invasive nucleic acids in plants. Some satellite RNAs are malignant: for example, Cucumber Mosaic Virus (CMV) Y-sat can trigger production of 22nt sRNAs against transcripts of photosystem component Chl I, causing chlorosis in the infected plant and worsening viral symptoms (Shimura et al., 2011; Smith et al., 2011). Moreover, CMV Q-satellite (Q-sat) can also interact with a bromodomain protein (BRP), suggesting that sat-RNAs can directly target effectors of epigenetic regulation and potentially affect chromatin regulation (Chaturvedi et al., 2014). However, some satellite RNAs are molecular parasites of viruses because they cause the reduction of their helper viruses, acting benignly in agriculture. Two modes of action have been proposed to mediate the benign effects of sat-RNAs in the infected plants: 1) direct competition for cellular machinery, and 2) induction of sRNAs that target the helper

viruses for cleavage, producing substrates for RDR6 to generate secondary sRNAs that in turn target the virus (Shimura and Masuta, 2016). Interestingly, it was recently shown that the genome of *Nicotiana tabacum* contains Y-sat sequences that produce 24nt siRNAs and direct DNA methylation (Zahid et al., 2015). Furthermore, it has been hypothesized that sat-RNAs evolved from RNAi byproducts because the sliced products of RNAi-targeted viruses or TEs can serve as substrates of TE- or virus-encoded RNA ligases for the production of chimeric RNAs. This has been shown to occur spontaneously in greenhouse experiments with CMV (Hajimorad et al., 2009).

Plants deficient in RNAi are hypersusceptible to virus infection; specific cases are *hen1, ago1, ago2, dcl4, drb4, ago1 ago2,* and *dcl2* mutants. However, this phenomenon is not seen in *dcl1* or *dcl3* mutants, implying that they do not affect virus accumulation (Deleris et al., 2006; Qu et al., 2008). The current model of RNAi antiviral defense is akin to the tasiRNA biogenesis (Figure 10). It consists of the production of 21nt virus derived small RNAs (vsiRNAs) through the cytoplasmic DRB4 and DCL4. The vsiRNAs are stabilized by HEN1 methylation and loaded into AGO1 or AGO2 effectors to guide cleavage of viral RNAs. The cleaved transcripts serve as substrates for SGS3, RDR6 and DCL2 to produce 22nt secondary vsiRNAs. The secondary vsiRNAs are in turn loaded into the antiviral AGO2 and reinforce the antiviral response (Pumplin and Voinnet, 2013). However, this model has been challenged by the discovery of the production of secondary vsiRNAs by SGS3 and RDR6 in the absence of AGO1 and AGO2. Furthermore, not all vsiRNAs are equally efficient at targeting the virus (Wang et al., 2011). Indeed, degradadome studies have shown that only few vsiRNAs cause viral RNA cleavage, and

DCL2 produced 22nt vsiRNAs are less efficient than the 21nt vsiRNAs produced by DCL4 (Wang et al., 2010b, 2011). Furthermore, viruses can evade this mechanism by localizing to different cellular compartments than the RNAi machinery, such as vesicles and chloroplasts (Laliberté and Sanfaçon, 2010).

1.3.2 Viral suppressors of RNA silencing

RNAi targets invasive nucleic acids, imposing a selective pressure that has resulted in essentially all plant viruses encoding for suppressors of RNA silencing (VSRs), thus enabling infection. In fact, both infection efficacy and severity of the disease caused by viruses correlate with the effectiveness of their encoded VSRs (Csorba *et al.*, 2015). Notably, the virus-RNAi-VSR systems were the first evidence for a role of sRNA in plant defense. They have also been a useful toolkit in the investigation of the RNAi mechanism itself because viruses have evolved VSRs to target virtually every step in the sRNA-mediated defense (Figure 10). Roughly speaking, VSRs can be grouped in three major classes, those that prevent sRNA biogenesis, those that inhibit AGO effectors, and the ones that preclude RNAi amplification. VSR-mediated impairment of sRNA biogenesis is exemplified by Cauliflower Mosaic Virus (CaMV) protein P6. This suppressor prevents the accumulation of vsiRNAs by binding to DRB4 in the nucleus and preventing its translocation to the cytoplasm. Cytoplasmic localization of DRB4 upon virus infection is essential for its interaction with DCL4 (Haas et al., 2008). Tobacco Mosaic Virus (TMV) protein P126 inhibits HEN1 and thus destabilize vsiRNAs (Vogler et al., 2007; Yu et al., 2006), while Sweet Potato Chlorotic Stunt Virus (SPCSV) produces an RNase III enzyme that specifically cleaves sRNAs (21-24nt) into inactive 14nt oligos (Kreuze et al., 2005).

Additionally, it is well-known that the Tombusvirus genus of plant viruses, including Tomato Bushy Stunt Virus (TBSV), encode the P19 protein, P19 functions as a "molecular caliper" by decoying vsiRNAs and preventing their loading into AGO effectors (Guo *et al.*, 2014; Ye *et al.*, 2003; Vargason *et al.*, 2003). Interestingly, P19 has a particularly low affinity for miR168, which directly targets AGO1 transcripts, producing downregulation of AGO1 to escape RISC attack (Várallyay et al., 2010). Aside P19, other VSRs can directly target AGO proteins. Polerovirus-encoded P0 is an F-box protein able to form a complex with SKP1/CULLIN1 to ubiquitinate AGO1 and induce its degradation through the autophagy pathway (Baumberger et al., 2007; Csorba et al., 2010). Similarly, Potato Virus X (PVX) protein P25 promotes degradation of both AGO1 and AGO2 through the 26S proteasome (Chiu et al., 2010). VSRs can also inhibit AGO silencing function. CMV protein 2b directly binds to AGO1 and inhibits its slicing activity (Zhang et al., 2006a), while Turnip Crinkle Virus (TCV) protein P38 associates with AGO1 and AGO2 through a glycine-tryptophan (GW) - AGO hook to prevent loading of siRNAs but not of miRNAs into AGO effectors (Azevedo et al., 2010; Zhang et al., 2012). DNA viruses also encode VSRs to target AGO1 activity. Mungbean Yellow Mosaic India Virus (MYMIV) AC2 protein binds to AGO1 and inhibits its slicing activity; however, it also interacts with and inhibits RDR6, therefore preventing accumulation of secondary vsiRNAs and the amplification step of the silencing signal (Kumar et al., 2015). A similar mechanism has been reported for Rice Yellow Stunt Virus (RYSV) P6 protein (Guo et al., 2013), which targets RDR6, and for Tomato Yellow Leaf Curl Virus (TYLCV) V2 protein, which interacts with SGS3, and prevent systemic silencing (Zrachya et al., 2007).

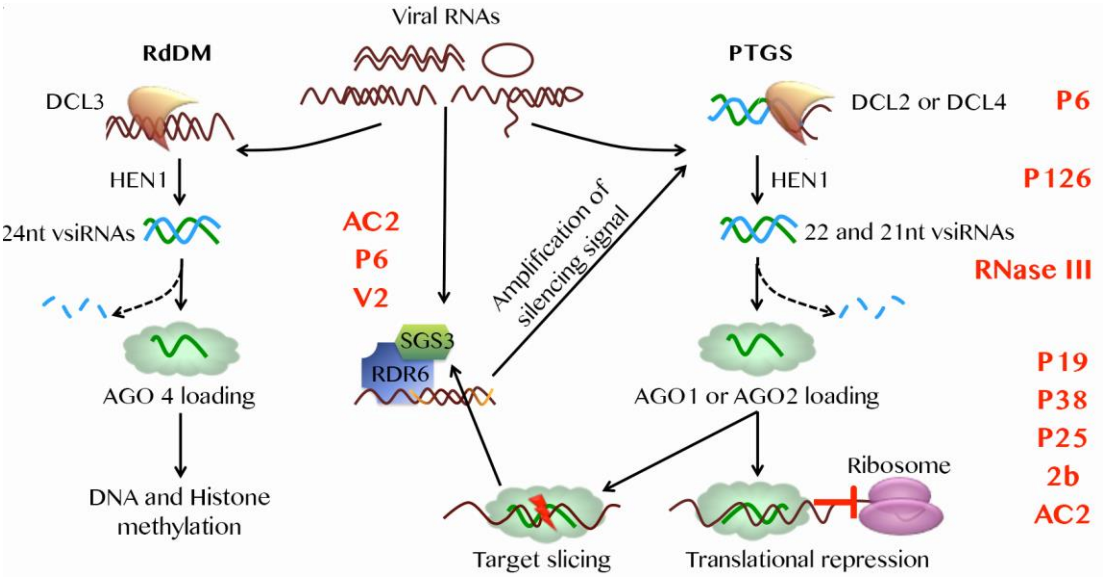


Figure 10. Mechanism of antiviral response mediated by sRNAs.
Viral suppressors of RNA silencing (VSRs) are depicted in red at the specific stage that they inhibit. Dotted black arrows represent the separation of the guide and passenger strands of the siRNA duplex, which ends up in the degradation of the passenger strand; red blunt arrow represents inhibition.

1.3.3 RNAi in non-viral threats

The plant defense against non-viral pathogens involves a complex signaling pathway to deploy a broad spectrum or targeted immune responses. Briefly, recognition of pathogen-associated molecular patterns (PAMPs), such as flagellin in bacteria or chitin in fungi, elicits a basal layer of defense known as PAMP-triggered immunity (PTI). Viral dsRNAs can be considered as virus-associated molecular patterns (VAMPs). Under this selective pressure, pathogens have engaged in a co-evolutionary arms race to overcome the PTI by producing specific virulence factors called effectors; this is in a similar fashion

as the selective advantage conferred by the viral production of VSRs. To counter those effectors, plants produce specific resistance proteins (R) during the targeted effector-triggered immunity (ETI). Previously, PTI and ETI have been considered to be independent from RNAi because they mostly regulate protein- and hormone-mediated immune responses, but recent discoveries have pointed to an sRNA-mediated regulation of both immune strategies.

The PTGS pathways have also been found to be involved in the response to non-viral pathogens. Arabidopsis plants lacking a functional miRNA pathway (i.e. *ago1*, *hen1*, and *hasty*) seem to prevent infection by the fungus *Verticillium*, evident in the reduced fungal growth. This observation implies that successful infection requires a functional miRNA pathway in the host. Importantly, mutants of *dcl4*, *rdr6*, *sgs3*, *and ago7* were all hypersusceptible to *Verticillium* but not to *Fusarium*, *Botrytis*, *Alternaria* or *Plectophaerella*, suggesting that an amplified endogenous siRNA pathway specifically restricts *Verticillium* infection. Furthermore, loss-of-function mutants in specific TGS components display similar hypersusceptibility to the fungi mentioned above; thus, different stages in the RNAi mechanisms play different roles in the fungal infection by Verticillium (Ellendorff et al., 2009).

1.3.3.1 RNAi is a constitutive repressor of the immune response

Two major breakthroughs were the discovery of miRNA-mediated constitutive repression of ETI and PTI components, and the finding that pathogenic bacteria, fungi and oomycetes also code for suppressors of RNA silencing among their effectors. The constitutive repression of ETI and PTI responses mediated by miRNAs is consistent with the requirement of the miRNA pathway for *Verticillium* infection. Regarding RNA silencing, the best-understood pathogenic system is the one comprising *A. thaliana* and the bacteria *Pseudomonas syringae* pv. tomato (*Pst*) (Figure 11). There, the flagellin peptide flg22 induces transcription of *MIR393* gene. miR393 represses *TRANSPORT INHIBITOR RESPONSE 1* (*TIR1*), which functions in the ubiquitin pathway to degrade transcriptional repressors in response to auxin and promote cell growth (Navarro et al., 2006). The role of miRNAs in regulation of PTI was further highlighted in a stronger miRNA response of wild-type Arabidopsis plants upon infection with *Pst hrcC*$^{-}$ *Pst hrcC*$^{-}$ is defective in the type-III secretion system and is therefore unable to deliver the its effectors to the host. Correspondingly, when Arabidopsis mutants in the miRNA pathway (i.e. *dcl1-9* and *hen1-1*) were challenged with *Pst hrcC*$^{-}$ the infection symptoms worsen (Navarro et al., 2008), implying that wild-type *Pst* delivers a suppressor of the miRNA pathway through the type III secretion system. *Pst* infection also induces the accumulation of AGO2, and *ago2* mutants are more susceptible to infection. Several miRNA* (passenger) strands accumulate and load into AGO2 during *Pst* infection, whereas these molecules are otherwise rapidly turned-over during loading of miRNA–miRNA* duplexes into AGO1 in uninfected plants. Among the miRNA* loaded into AGO2 is miR393*,

which targets the transcript of MEMBRIN 12 (*MEMB12*). MEMB12 is a SNARE protein that negatively regulates the secretion of PATHOGENESIS-RELATED 1 (PR1) protein. Accordingly, plants carrying mutations in *MEMB12*, or overexpressing miR393*, display enhanced secretion of PR1 and increased resistance to virulent and avirulent *Pst* (Zhang et al., 2011b). Thus, AGO2 and AGO1 are regulators of antibacterial immunity by binding to miR393* and miR393, respectively; thus de-repressing exocytosis of antimicrobial PR1 and preventing allocation of resources to growth (Figure 11). Similar de-repression of defense genes by miRNA regulation was observed during wheat powdery mildew infection by fungus *Erysiphe graminis.* In this case, downregulation of miR156 is required for the accumulation of the SPL target *Ta3711*, which is a positive regulator of plant defense response (Xin et al., 2010). Notably, it has been noticed that 10 conserved miRNA families are downregulated in the gymnosperm *Pinus taeda* (loblolly pine) upon fungal rust infection by *Cronartium quercuum* (Lu et al., 2007). miRNAs controlling *R* gene transcripts were also observed in soybean infected with the oomycete *Phytophthora sojae*.

The RNA decay machinery might compete or collaborate with the epigenetic machinery to potentiate or circumvent either transcriptional recovery via PTGS or memory formation via RdDM. For instance, stress is well known to induce very rapid production of transcripts, which conceivably could lead to an increase in aberrant mRNA that, in turn, would be potential substrates for RDRs. In line with this hypothesis, the RNA decay inhibitor 3′-phosphoadenosine-5′-phosphate is produced during drought and light stress, potentially impairing RNA decay. In this scenario, mRNA molecules could be more

readily susceptible to PTGS. During stress, this could lead to RdDM and potentially heritable changes in gene expression.

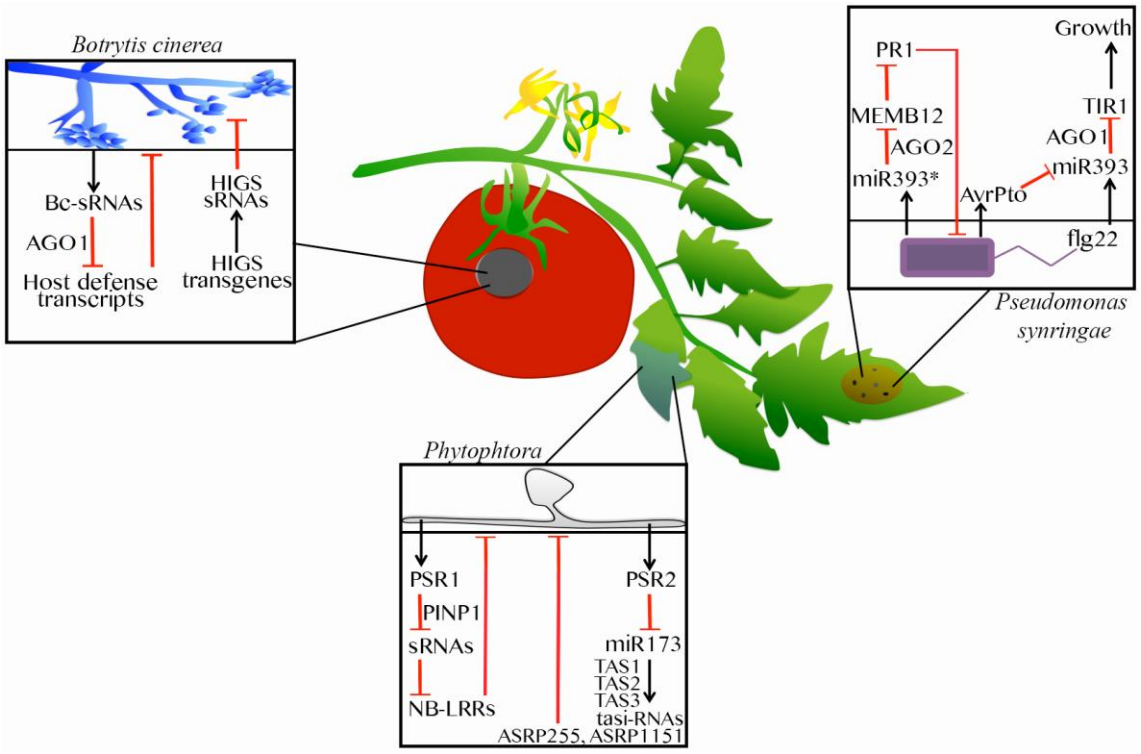


Figure 11. sRNAs regulate host-pathogen interactions.
Black arrows represent validated positive regulation or the release of pathogen effectors; Red blunt arrows represent validated inhibition or down-regulation of target genes.

1.3.3.2 Pathogen suppressors of RNA silencing

Consistent with the involvement of RNAi in the regulation of plant defense, further study of the *Pst*-Arabidopsis system led to discovery of bacterial suppressors or RNA silencing (BSRs) (Navarro et al., 2008). Specifically, *Pst* effectors include AvrPtoB,

which suppresses transcription of *MIR393a* and *MIR393b* genes (Figure 11); *AvrPto* also reduces accumulation of miRNAs in a posttranscriptional manner; and, Hop1-1 represses both slicing and translational inhibition activities of AGO effectors. Importantly, not only bacteria have been shown to produce silencing suppressors, the oomycete *Phytophtora sojae* delivers two RNA-interacting proteins to the host to prevent RNA silencing, namely PSR1 and PSR2 (from Phytophtora Suppressors of RNA silencing 1 and 2, Figure 11) (Qiao et al., 2013, 2015). PSR1 impairs accumulation of sRNAs by interfering with the activity of the nuclear PSR1-Interacting Protein (PINP1). PINP1 regulates the accumulation of several kinds of sRNAs likely mediating the assembly of dicing complexes. On the other hand, PSR2 negatively affects the abundance of specific sRNAs that targets some known defense response genes such as nucleotide-binding leucine-rich repeat proteins (NB-LRRs). Notably, PSR2 also influences the accumulation of specific tasiRNAs through targeting miR173, which in turn suppress the biosynthesis of ASRP255 and ASRP1151 tasiRNAs without affecting the miR390-dependent tasiRNAs.

1.3.3.3 RNA in the battleground: sRNAs as effector and resistance pawns

A second breakthrough came with the discovery of delivery of RNA molecules as effector/resistance factors. The host-pathogen interaction is bidirectional, which means that as pathogens are able to deliver effector molecules to their host, there are mechanisms for the host to deliver molecules to the pathogen. Weiberg and collaborators provided outstanding evidence for this phenomenon: the fungal pathogen *Botrytis cinerea* produces sRNAs that are transferred to the host and mediate the establishment of pathogenesis, by modulating the expression of components of the pathogen sensing systems (Figure 11)

(Weiberg et al., 2013). On the other hand, plants can inject sRNAs into pathogens and induce gene silencing in the invaders in a process called Host Induced Gene Silencing (HIGS) (Figure 11). This mechanism was first proposed and tested in barley and wheat infected with the fungus *E. graminis*, which causes powdery mildew. Briefly, siRNAs targeting the fungal effector Avra10 are produced in the grasses to trigger RNA silencing in the pathogen, which in turn renders the host immune to infection. The sRNA effect is abolished when the pathogen expresses an sRNA-resistant Avra10 transcript (Nowara et al., 2010). This artificial RNA-based crop protection has been successfully applied in many crops-pathogen systems since 2010, but so far it is unknown whether it occurs *in vivo* as well.

1.3.4 RNAi as a fine-tuned sensor of biotic threat

Another level of complexity is added when the RNAi mechanisms themselves serve as sensors of pathogen attack or disease. During virus infections, but also in other pathogenic processes, the RNAi effector AGO1 is modulated either by impairing its function or by altering its steady-state level in the cell. In any case the silencing effect is reduced, which in turn results in the deregulation not only of the "intended" virus RNA target but also of the endogenous AGO1-dependent transcripts. Among those endogenous transcripts is the *AGO2* mRNA, which is normally targeted by miR403 in Arabidopsis; however, as AGO1 is impaired, AGO2 repression is released and can accumulate to a higher level as such AGO2 can serve as a surrogate /or additional line of antiviral RNAi effector (Azevedo et al., 2010; Fátyol et al., 2016). The induction of secondary RNAi components upon infection has also been reported in rice, where virus infection induces

expression of AGO18, which comparable to AGO10 in Arabidopsis, works as a decoy for an endogenous miRNA, namely miR168. In normal conditions, AGO1 homeostasis is regulated by miR168; however, upon expression of AGO18, the steady-state level of miR168 is greatly reduced and the levels of AGO1 are increased to potentiate the antiviral response (Wu et al., 2015).

1.3.5 RNAi as mediator in amicable interactions

However, the life of plants involves more than just defending themselves from foes. In nature, plants often rely on biotic interactions for survival, the most widespread being the establishment of mycorrhizae and nitrogen fixing nodules (Gobbato, 2015). Mycorrhizae are mutualistic symbioses between fungi and plant roots, in which the fungal micellium acts as an extended rhizome providing the plant with essential nutrients (mainly phosphorus, but also water and other minerals) in exchange of sugars from the plant. It is estimated that around 80% of terrestrial plants form mycorrhizae, and it has been proposed that the establishment of these associations allowed plants to colonize terrestrial ecosystems. The mycorrhization can be extracellular or intracellular, denominated ectomycorrhizae and arbuscular mycorrhizae (AM), respectively. On the other hand, the nitrogen fixing symbiosis appeared later in evolution and has been described only in plants of the Rosid Clade I. It occurs under nitrogen-limiting conditions between nitrogen-fixing bacteria, known as rhizobia, and the plant root. Notably, the formation of rhizobia nodules and AM is not only a phenomenon of plant-microbe interactions but also implies the deployment of a developmental program. The AM colonized root cells specialize, and are physiologically modified for the exchange process. Similarly, nodulation occurs on stems

as root-derived organs, which depend on the deployment of the root developmental program (Franssen et al., 2015). Interestingly, both developmental programs –AM and nodulation- are reversible, since the interactions can be abolished upon changes in the environmental conditions, such as phosphorus and nitrogen replenishment.

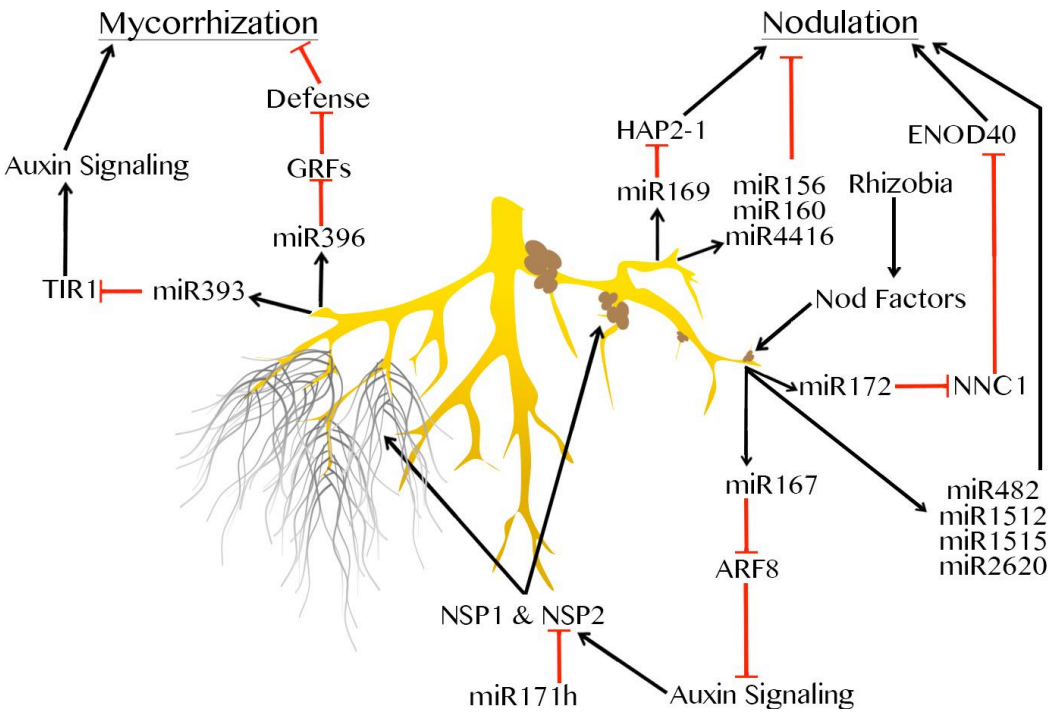


Figure 12. miRNAs regulate positive symbiosis.
Mycorrhization and nodulation of nitrifying bacteria are shown. Black arrows represent validated positive regulation; red blunt arrows represent validated downregulated targets or inhibition.

Not surprisingly, the miRNA machinery is also involved in developmental regulation of the interactions above (Figure 12). In *Medicago,* miR396 and miR393 are negative regulators of mycorrhization (Bazin et al., 2013; Etemadi et al., 2014). miR396

targets the bHLH and Growth Regulating transcription Factors (GRF), which play a role in synchronization of development and plant defenses (Liu et al., 2014a); while miR393 is important in auxin regulation by targeting the TIR1 family of auxin transporters (Navarro et al., 2006). Although it is expected that hormone, defense and development regulation converge in the establishment of mycorrhizae, the exact mode of action of these miRNAs in mycorrhization has been elusive. Establishment of nitrogen fixing nodules is also regulated by miRNAs (Figure 12). In *Medicago*, the downregulation of the *NF-YA* transcriptional regulator *HAP2-1* by over-expression of miR169 arrests meristem development and delays nodulation, which in turn causes the formation of non-fixating nodules. Interestingly, expression of miR169-resistant *HAP2-1* also produces defective nodules. These observations and the complementary expression pattern of miR169 and *HAP2-1*, imply the need for fine-tuning *HAP2-1* expression during nodulation (Combier et al., 2006). Thanks to the combined strategies of miRNA over-expression, target-mimicry and miRNA-resistant targets, many other miRNAs have been reported to be involved in nodulation. Initially, rhizobia release Nodulation Factors (Nod) that are sensed by the compatible host and enable nodulation. In the soybean-*Bradyhizobium japonicum* system, the rhizobia release Nod factors that subsequently induce expression of miR172c. miR172c positively regulates nodule initiation and nodule number by targeting the transcript of *NODULE NUMBER CONTROL 1 (NNC1)* for degradation. NNC1 protein directly inhibits the nodulin gene *ENOD40* upon binding to its AP cis-element, therefore preventing nodule initiation (Wang et al., 2014). Notably, the miR172-AP module has also been shown important in the nodulation in the common bean-*Rhizobium etli* system,

highlighting the idea of a conserved mechanism (Nova-franco et al., 2015). In soybean, *B. japonicum* Nod factors also induce expression of a positive regulator of nodulation, miR167c. Targets of miR167c in soybean are transcripts of auxin response factors *GmARF8a* and *b*, pointing toward the necessity for regulation of auxin response in the establishment of nitrogen fixing nodules. Interestingly, the miR167-*GmARF8* module seems to be essential under low inoculum conditions, and to act upstream the nodulation gene *NODULATION SIGNALING PATHWAY1 (NSP1),* and *ENOD40* (Wang et al., 2015a). Although miR172-*AP* and miR167-*ARF8* are the best understood modules involved in nodulation, miR482, miR1512, miR1515 and miR2606b are also known to positively affect nodulation (Li et al., 2010); while miR156, miR160 and miR4416 have a negatively effect (Turner et al., 2013). Given the similar impositions made by the mycorrhyzal fungus and the rhizobia on the plant hosts, it is perhaps not surprising that the establishment of these interactions include common regulators. In fact, the GRAS transcription factors NSP1 and NSP2 have been shown to be essential for the establishment of both symbioses (Delaux et al., 2013; Kaló et al., 2005; Lauressergues et al., 2012; Smit et al., 2005). In *Medicago*, *NSP2* is regulated by miR171h; consistently, over-expression of miR171h impairs fungal growth and mycorrhization (Lauressergues et al., 2012).

1.4 Chromatin modification

The eukaryotic nuclear DNA is packaged into nucleosomes to constitute the highly regulated chromatin. Nucleosomes are histone octamers around which the DNA wraps as wool in a spool; typically, one nucleosome wraps 146bp of DNA. Chromatin regulation includes modifications on the DNA (i.e. 5'methyl cytosine), post-translational modification of the histones, and positioning and density of the nucleosomes along the DNA. All these modifications are relevant to gene expression. Histone methylation is one of the most abundant posttranslational modifications found in chromatin. It takes place on lysine residues of histone amino-terminal tails, and serves as a second layer of information in the eukaryotic genomes. Histone methylation has variable effects on gene expression depending on the precise residues, contexts, and modification complexity (mono-, di-, or tri-methylation). For instance, histone 3 lysine 4 tri-methylation (H3K4me3) is almost uniquely associated with transcriptionally active chromatin, while H3K9me2 and H3K27me3 are used heterochromatin markers (Binda, 2013; Black and Whetstine, 2011; Du et al., 2015). Histone methylation is catalyzed by SET domain containing proteins such as the Arabidopsis Su(var)3-9 homolog 4, KYP, which is responsible of the H3K9me2 deposition (Jackson et al., 2002; Law and Jacobsen, 2010). In Arabidopsis, H3K27 methylation is mainly catalyzed by the PRC2 complex (Liu et al. 2010; Zheng and Chen 2011). Both H3K9me2 and H3K27me3 can spread over neighboring regions inducing a silent heterochromatic conformation. In plants, this spreading entails a positive feedback loop between non-CG methylation (catalyzed by *CHROMOMETHYLASE3 -CMT3-*),

recognized by KYP to direct H3K9me2, which further recruits CMT3 to boost or regulate DNA methylation (Berger, 2007; Du et al., 2015; Henderson and Jacobsen, 2007).

1.4.1 Chromatin methylation and genome stability

Chromatin methylation is important in the regulation of endogenous gene expression during developmental processes, but it is also critical for genome stability (Alvarez et al., 2010; Black and Whetstine, 2011; Borges and Martienssen, 2013; Chinnusamy and Zhu, 2009; Cui and Cao, 2014; Fischer et al., 2006; Liu et al., 2010; Mirouze et al., 2009; Pikaard and Scheid, 2014). TGS controls the exogenous invasive DNAs such as transposons and repetitive sequences. Since DNA viruses can associate with histones and form minichromosomes in the host cells (Hanley-Bowdoin et al., 2013), methylation is likely used to inhibit these pathogens. Indeed, previous work has shown that plants use methylation of viral chromatin to limit virus replication and transcription (Aregger et al., 2012; Brough et al., 1992; Pumplin and Voinnet, 2013). Specifically: 1) Viral DNA and associated histone proteins are methylated in infected plants; 2) Methylation-deficient *Arabidopsis* mutants are exquisitely sensitive to virus infection and show enhanced disease symptoms; and 3) Viral DNA methylation is reduced in methylation-deficient mutant plants that display enhanced susceptibility (Raja et al., 2008).

1.5 The Geminivirus

Viruses of the family Geminiviridae, or geminivirus, are circular ssDNA plant pathogens of great economical importance worldwide. They consist of one or two small (~2-5kb) genomic molecules encapsidated in twin icosahedral particles that are targeted

to the plant nucleus, where their genomes are replicated by rolling circle and conformed in a dsDNA intermediate. This intermediate is wrapped around nucleosomes, the viral minichromosome from where all transcripts and infective ssDNAs will be synthesized (Paprotka et al., 2015; Pilartz and Jeske, 1992; Shung et al., 2006). To do so, the virus must highjack the plant cell machinery with the 4-8 proteins encoded in its genome. The geminivirus life cycle, small genomes, and consequent limited coding capacity, have made Geminivirus especially useful in the study of DNA replication, transcription regulation, and defense in the hosts (Carrillo-Tripp et al., 2006; Fondong, 2013).

The family Geminiviridae is further classified in four genera: Begomovirus, Curtovirus, Mastrevirus and Topocovirus. This organization is based on genome conformation, host range and nucleotide similarity. Curtovirus have monopartite genomes, infect dicotyledonous hosts, and are transmitted by the beet leafhopper (*Circulifer tenellus*) (Hormuzdi and Bisaro, 1993, 1995); Topocovirus are so far represented by a single member named *Tomato pseudo-curly Top Virus* (TPCTV), which is a monopartite virus that infects dicotyledonous hosts through a treehopper vector (*Micrutalis malleifera*) (Briddon and Markham, 2001; Briddon et al., 1996; Fondong, 2013; Hanley-Bowdoin et al., 2013; Mansoor et al., 2003); Mastrevirus also have monopartite genomes and can infect monocotyledonous as well as dicotyledonous hosts almost exclusively in the Old World, and are transmitted by leafhoppers and dragonflies (Kammann et al., 1991; Matzeit et al., 1991). Finally, Begomovirus are the most prevalent and well studied of all geminivirus, they infect monocots and dicots, and are among the most devastating plant viruses in the world (Scholthof et al., 2011; Yadava et al., 2010).

They can have monopartite or bipartite genomes and are transmitted by whitefly (*Bemisia tabaci*) (Fondong, 2013; Gutierrez, 2000a, 2000b). The bipartite genomes are named DNA A and DNA B and do not share sequence similarity, except for a ~200nt 5' intergenic region named Common Region (CR), which contains the origin of replication and a conserved stem loop that is required for the assembly of the virus particles (Figure 13) (Bisaro et al., 1990a).

The geminivirus genomes have open reading frames coded in the virion sense –named V-, or in the virion complementary sense -named C-. The nomenclature of these genes specifies the coding strand followed by a number that indicates the order of discovery. In a begomovirus bipartite genome, the name of the DNA particle (A or B) precedes the strand and order determinants (Figure 13).

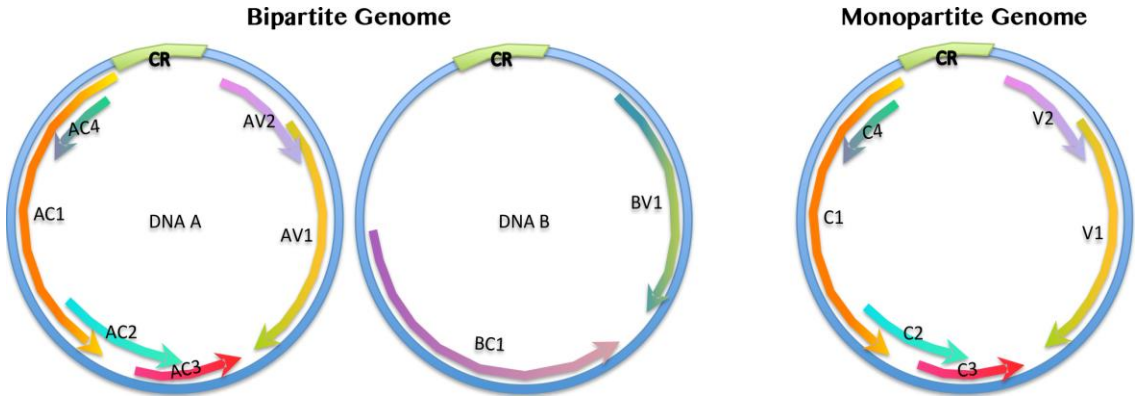


Figure 13. Genome organization of begomoviruses.
Geminivirus genomes code for proteins involved in: viral replication (AC1/C1), encapsidation (AV1/V1), cell-to-cell movement (BC1/V2), nuclear shuttling (BV1/V2), and silencing suppression (AC4/C4, AV2, AC2/C2). However, the limited coding capacity of these pathogens has lead to the evolution of multifunctional proteins, so most of them interact with multiple cellular components and have different roles in the pathogenesis.

1.6 TrAP: the begomovirus AC2/C2 protein

The AC2 is exquisitely multifunctional and unique among the geminivirus C2 positional analogues found in curtoviruses and topocoviruses. It contains a transcriptional activation domain, which is required for the expression of the late stage genes that code for the coat protein (AV1/V1) and the nuclear shuttle protein (BV1). Therefore, the begomovirus AC2/C2 is also known as TrAP for transcriptional activator protein, and it is essential for the virus assembly and spreading in the host. Here, we refer to begomovirus AC2/C2 as TrAP, and to the curtovirus and topocovirus positional analogues as C2.

TrAP is a 15KDa protein expressed from a polycistronic transcript that includes AC1/C1, AC4/C4 and extends to AC3/C3. TrAP is a Zn^{2+} binding phosphoprotein encoded by the begomovirus plant pathogens (Figure 14). From N- to C-termini TrAP contains a serine-rich stretch that has been proposed as a possible phosphorylation site to regulate its cellular localization, interestingly it contains a proline residue that can be recognized by Karyopherin α; immediately after, TrAP has a basic region that was proposed to serve in DNA binding, but that later was described as the Nuclear Localization Signal (NLS); then, the Cys-His Rich Region, which deviates from the canonical Zn2+ finger domains CCHH, is able to bind Zn^{2+} and is also the TrAP dimerization domain; it is followed by a stretch of 40aa with no known function, no sequence homology to known proteins, and no predicted secondary structure; finally at the C-terminal, there is an acidic region that has been related to transcriptional activation (Yang et al., 2007). Importantly, it has been reported that TrAP can be phosphorylated at serine 109 and that this modification affects the virus pathogenicity (Shen et al., 2014a). Interestingly, TrAP does

not bind to dsDNA but it relies on host proteins to target the viral promoters such as the *Arabidopsis thaliana* transcription factors PEAPOD2 and JDK (Chung and Sunter, 2014; Lacatus and Sunter, 2009; Lozano-Duran et al., 2011); however, the precise mechanism of transactivation mediated by TrAP remains elusive.

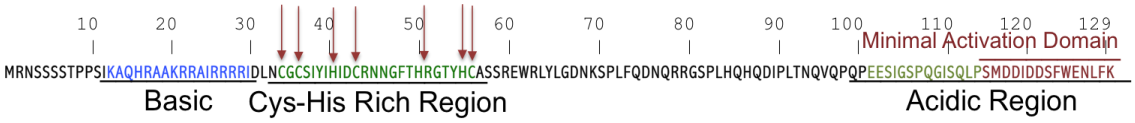


Figure 14. Primary structure of the tomato golden mosaic virus (TGMV) TrAP.
This is a model of the TrAP encoded by a bipartite begomovirus. The sequence is color-coded and functionally annotated. The nuclear localization signal (NLS) is located on the basic region and represented in blue. The dimerization and DNA binding domains are located in the Cys-His rich region and represented in blue, red arrows point to the cysteine residues; notably, C35 is essential for TrAP dimerization. The C-terminal acidic region codes for the transcriptional activation domain and it is represented in green and red, the red region represents the minimal activation domain.

Possibly due to the limited coding capacity of these viruses, most of their proteins are truly multi-functional. Particularly, TrAP has been shown to be a transcriptional activator, a silencing suppressor and a suppressor of basal defense.

1.6.1 TrAP is a transcriptional activator

As a transcriptional activator, TrAP is required for the expression of the virion-sense genes in both genomic particles. Promoter-reporter fusion and nuclear run-on assays revealed that TrAP regulates the transcription of the coat protein AR1 (Sunter and Bisaro, 1991; Sunter et al., 1990), and the nuclear shuttle protein BR1 (Sunter and Bisaro, 1992).

Importantly, there is no systemic infection caused by *trap* mutant begomoviruses because of the requirement of these two viral genes. Mutation/complementation analyses of *trap, ar1, bl1* and *br1* mutant viruses showed the absolute requirement of the TrAP-dependent gene products for successful systemic infection of the host and are summarized below:

1. WT DNA A genomic particle is self-sufficient for replication, but it depends entirely on the activity of both DNA B encoded proteins to infect adjacent cells as well as to move systemically (Von Arnim and Stanley, 1992).
2. WT DNA A cell-to-cell and systemic movements are sustained when both BL1 and BR1 proteins are provided in trans (Jeffrey et al., 1996).
3. *trap* mutant DNA A co-inoculated with WT DNA B does not sustain cell-to-cell nor systemic movement (Elmer et al., 1988). It can replicate in the inoculated cells but does not accumulate ssDNA (Hayes and Buck, 1989).
4. *trap* mutant DNA A can move to adjacent cells and systemically when BR1 and BL1 are provided, although to a greatly reduced extend as compared to the WT. No ssDNA accumulated (Jeffrey et al., 1996).
5. *trap* mutant DNA A cell-to-cell movement can be restored by BR1 alone but not by BL1. No ssDNA accumulated (Jeffrey et al., 1996).
6. *ar1* mutant DNA A co-inoculated with WT DNA B particles have cell-to-cell and systemic movement although at lower levels as compared to WT. This phenotypes is also accompanied by lower than WT ssDNA

accumulation (Gardiner et al., 1988; Sunter and Bisaro, 1992; Sunter et al., 1990).

7. *ar1* mutant DNA A shows recovered cell-to-cell movement when BR1 and BL1 are provided in trans. No ssDNA accumulated (Jeffrey et al., 1996).
8. *ar1* mutant DNA A systemic movement is not restored by BR1 and BL1 expression in trans. No ssDNA accumulated (Jeffrey et al., 1996).
9. *ar1* mutant DNA A expressing BR1 in place of AR1 shows cell-to-cell and systemic movement when BL1 is provided in trans. ssDNA accumulation (Jeffrey et al., 1996).

These discoveries evidenced a regulatory role for TrAP and it has since been proposed to serve as a temporal switch during the virus infection, following a common strategy among DNA viruses in which an early viral gene induces the expression of genes that are required for later stages of infection such as capsid and movement proteins. Notably, the rolling circle replication of geminivirus produces ssDNA in every round (Stenger et al., 1991). Therefore, probably there is a competition between the replication machinery and the virus proteins for encapsidation and movement, so TrAP can serve to tilt the balance from dsDNA amplification to ssDNA accumulation. This is consistent with the lack of ssDNA accumulation in *trap* and *ar1* mutants, as well as with the ssDNA binding capacity of both AR1 and BR1 proteins (Pascal et al., 1994).

Two main approaches have been taken to understand the TrAP-dependent transcriptional activation of viral genes: 1) by studying the promoter regions of the TrAP-activated genes (PAR1 and PBR1, Figure 15), and 2) by studying the structural

components in TrAP that are required for the transcriptional activation. The promoter studies have shown:

1. P_{AR1} truncations, but not WT P_{AR1}, are active in the phloem tissue, so TrAP possibly acts as a de-repressor of P_{AR1} in the phloem (Sunter and Bisaro, 1997). Indeed, many begomoviruses are restricted to the phloem, and P_{AR1} contains a phloem repressor element that has been mapped to a region within the TrAP gene.
2. P_{BR1} de-repression in the phloem tissue is also dependent on TrAP.
3. P_{AR1} and P_{BR1} of geminivirus that are not restricted to the phloem tissue have a conserved late element (CLE) (Argüello-Astorga et al., 1994).
4. P_{AR1} is activated by TrAP in mesophyll tissue (Sunter and Bisaro, 1997).
5. P_{AR1} and P_{BR1} have a bipartite arrangement of cis elements that are both necessary and sufficient for TrAP-dependent transcriptional activation, the sequences include but are not limited to the CLE and a CAAT region (Berger and Sunter, 2013; Sunter and Bisaro, 2003).

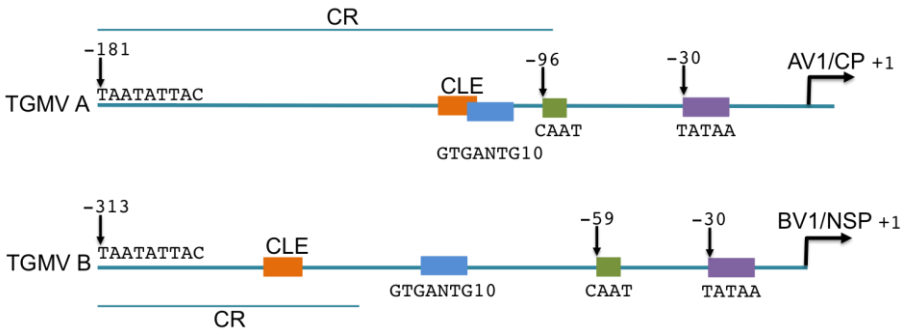


Figure 15. Structural components on the geminivirus genome required for TrAP-mediated transcriptional activation.
Schematic representation of the promoters of the TrAP-activated genes, AV1 and BV1, in the model begomovirus, tomato golden mosaic virus. CLE is conserved late element; CP is coat/capsid protein; NSP is nuclear shuttling protein.

The TrAP structural studies have shown:

1. TrAP proteins are not virus specific, which indicates a common transcriptional activation mechanism.
2. TrAP proteins can activate transcription of chromosomal insertions of PAR1-driven transgenes, and they can activate host genes (Trinks et al., 2005).
3. TrAP can bind to ssDNA and dsDNA. The binding is sequence independent, Zn^{2+}-dependent, and greatly favors ssDNA over dsDNA (Hartitz et al., 1999).
4. TrAP C-terminal acidic domain is a transcriptional activation domain capable of activating promoters in heterologous systems. In fact, a minimal activation domain consistent of the last 15aa of TrAP can drive gene expression in mice and yeast (Figure 14) (Hartitz et al., 1999).

5. TrAP interacts with the Arabidopsis transcription factor PEAPOD2, which recognizes both P_{AR1} and P_{BR1} promoters and can direct TrAP-dependent activation. PEAPOD2 does not have a transcriptional activation domain (Lacatus and Sunter, 2009).
6. TrAP but not C2 can interact with itself forming oligomers, this self-interactions is abolished by C35A mutation in the Cys-His Rich Region and is likely mediated by a disulfide bond. Viruses carrying the *trap* C33A mutation accumulate mostly in cytoplasm and cannot drive activation of P_{AR1}; nonetheless, it does not impair local silencing suppression (Yang et al., 2007).

1.6.2 TrAP is a silencing suppressor

TrAP/C2 have also been involved in symptom development and silencing suppression. These proteins have been involved in TGS suppression by interfering with the methyl cycle in the cytoplasm (Figure 16), depleting the cell from the methyl donor SAM and avoiding the viral chromatin methylation. Specifically, TrAP interacts with and impair the activity of adenosine kinase (ADK) (Mohannath et al., 2014; Wang et al., 2005). Briefly, the transfer of the methyl moiety from SAMe to a methyl acceptor produces S-adenosyl-homocysteine (SAH), which is subsequently hydrolyzed to homocysteine (Hcy) and adenosine by SAH-hydrolase. ADK catalyzes the synthesis of 5'AMP from adenosine and ATP, therefore promoting the flux through the methyl cycle, which regenerates SAMe from methionine (Met) (Moffatt et al., 2002; Weretilnyk et al., 2001). Thus, the inhibition of ADK likely impedes downstream trans-methylation reactions including chromatin

methylation in nuclei and metabolism of secondary metabolites including phytohormones in plants (Moffatt et al., 2002; Wang et al., 2003, 2005). Furthermore, C2 has been found to stabilize S-adenosyl-methionine decarboxylase 1 (SAMeDC1) by preventing its degradation by the 26S proteasome. SAMeDC1 is a negative regulator of the methyl cycle that produces dcSAMe from SAMe (Zhang et al., 2011c) (Figure 16). Targeting the methyl cycle may prevent not only the formation of heterochromatin, but also the deposition of important euchromatin marks (i.e. H3K4me3 and H3K36me3), which could result in suboptimal transcription and replication of the virus. Furthermore, SAMe is the most common methyl donor in the cell; its regulation not only serves in chromatin conformation, but also can have a plethora of metabolic effects in the cell.

RNAi can induce transcriptional and post-transcriptional gene silencing. The standard method to assess the silencing suppression activity of proteins in plants is as follows: a GFP stable transgenic *Nicotiana benthamiana* (*N. benthamiana*) plant is inoculated with a cocktail of three agrobacteria strains that drive the expression of GFP, anti-sense GFP RNA (These two constructs will provide dsRNA substrate to initiate RNAi against the GFP transgene), and suspect RNAi suppressor gene. The readout of silencing suppressor activity is the maintenance of GFP fluorescence in the plant. This system can reflect effects at any level of the RNAi mechanism and all TrAP, *trapΔAD* (TrAP protein lacking the 31 C-terminal aminoacids that constitute the transcriptional activation domain), and C2 proteins evaluated by this method showed RNAi suppression activity (Wang et al., 2005). This activity was (prematurely) attributed to the inactivation of ADK by TrAP/C2 because co-delivery of an inverted repeat of ADK gene (to produce ADK

dsRNA) and the addition of an ADK inhibitor also prevented silencing of the GFP transgene. It was also concluded that the silencing suppression activity is independent from the activation domain, but this is controversial because this activity in TrAP from Mungbean Yellow Mosaic Virus (MYMV) is completely abolished by mutations in the NLS, Cys-His- Rich Region, and *trapΔAD* (Trinks et al., 2005).

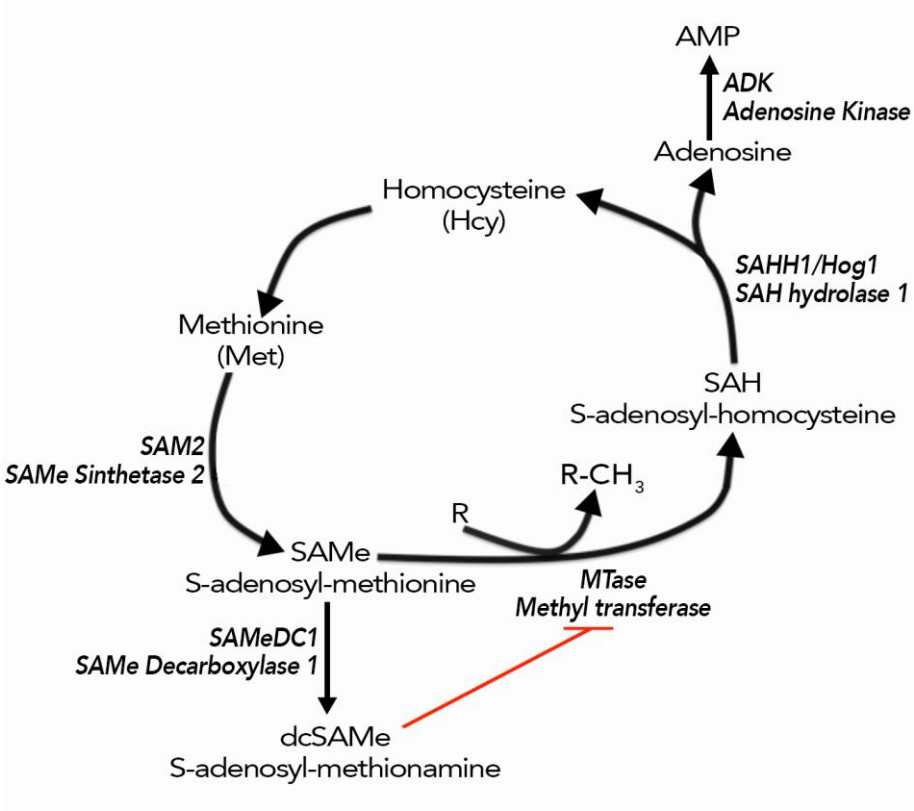


Figure 16. The cellular SAMe or methyl cycle.
R represents the substrate in the transmethylation reactions; enzymes are bolded and italic; red blunt arrow represents inhibition.

Other methods have been used to specifically evaluate the early events of RNAi, typically PTGS. The transient expression of genes in *N. benthamiana* is transient because of the establishment of RNAi against the delivered construct, so the co-delivery of a reporter gene together with a suspect RNAi suppressor serves to determine suppression of the RNAi establishment by measuring the expression time of the reporter gene. Typically a GFP reporter will be completely silenced after seven days, while when co-delivered with a PTGS suppressor it can last for several weeks. Neither TrAP nor C2 show silencing suppression by this method (Zrachya et al., 2007).

1.6.3 TrAP is a suppressor of basal defense and this activity is independent of the transcriptional activation domain

As a pathogenicity determinant, TrAP/C2 causes enhanced susceptibility when expressed in transgenic plants (Sunter et al., 2001). Both proteins interfere with host signaling pathways required for nutrient, cell cycle, development, and stress response regulation (Baliji et al., 2007, 2010; Yang et al., 2007). Until recently, TrAP/C2 interaction with the global regulator of metabolism SnRK1 kinase was proposed to be the effector of the virus metabolic control of the host cell (Hao et al., 2003; Shen and Hanley-Bowdoin, 2006). However, Dr. Hanley-Bowdoin's group reported that SnRK1 phosphorylates TrAP/C2 proteins to counter the virus threat (Shen et al., 2014a). In addition, C2 also affects the activity of COP9 signalosome (Lozano-Duran et al., 2011), suggesting its multiple functions in viral counter-defense. Although ADK, SAMDC1 and SnRK1 are TrAP/C2-interacting proteins, their cytoplasmic localization makes them unlikely responsible for the TrAP nuclear function. Furthermore, the loss-of-function mutants do

not phenocopy TrAP transgenic plants, suggesting that TrAP has additional targets in the nucleus.

Microarray analysis of the 6 hours transcriptome of protoplast transiently expressing TrAP protein from different begomoviruses, showed induction of 55 genes >2fold, while only 5 seemed downregulated; consistent with previous studies, there was no induction of host genes by TrAP mutated in their nuclear localization signal, in the Cys-His Rich Region, or lacking the activation domain (Trinks et al., 2005). This study confirmed the TrAP activation of host promoters by promoter-reporter fusions, but it did not show induction of genes normally associated with defense.

The constitutive expression of TrAP proteins lacking the activation domain results in plants that are more susceptible to virus infection. The same phenotype was observed in plants overexpressing the TrAP positional homolog C2; also, plants infected with *c2* mutant geminivirus exhibit enhanced recovery from infection. The mechanism is likely mediated by the interaction of C2 and TrAP with host proteins involved in pathogen response. Specifically, C2 and TrAP interact with the cytoplasmic proteins ADK and SnRK1, which are essential regulators of plant metabolism. ADK takes part in the methyl and ATP biosynthetic cycles, which affect a myriad of processes; among them are chromatin modifications and kinase activation. SnRK1 inactivation has shown to greatly enhance susceptibility to virus infection, while overexpression leads to higher resistance. C2 and TrAP inactivate ADK possibly directing the cell metabolism to a more permissive status for virus infection. Interestingly, SnRK1 can phosphorylate TrAP within the activation domain in S109, which is proposed to serve as a defense mechanism. Plants

infected with the phosphomimic TrAP S109D showed delayed and attenuated symptoms. Other works have shown that C2 interacts with and stabilizes SAMDC, further regulating the methyl cycle and greatly affecting the trans-methylation reactions in the cell; an important consequence of this is the demethylation of the virus genomes. Hyper-methylation of the virus genomes results in attenuated infection and higher recovery from infection.

In sum, sRNAs are central to gene regulation, which is not limited to development but underlies all interactions with the ecosystem. Among the biotic interactions, I have focused on the pathosystem established by begomoviruses and their host plants, and the dependence of their success on the function of the viral protein TrAP. Since TrAP has been previously described as a suppressor of sRNA-mediated gene silencing, the focus of my research was to determine the mechanism by which TrAP exerts this gene regulation in the host. I hypothesized that TrAP functions through the interaction with host factors and I used a combination of genetic and biochemical approaches to identify those TrAP-interacting proteins (TrIPs). Chapter 2 focuses on the interaction of TrAP with the histone methyltransferase KYP, while the global findings of proteomics studies are presented in Chapter 3. This work provides mechanistic insight for the TrAP-mediated suppression of TGS, provides valuable resources for the scientific community, and builds on the current understanding of the function of chromatin regulators. Specifically, we provide evidence of the role of KYP in the immune system and predict this to be a general mechanism employed by DNA viruses to overcome the host.

2. GEMINIVIRUS-ENCODED TrAP SUPPRESSOR INHIBITS THE HISTONE METHYLTRANSFERASE SUVH4/KYP TO COUNTER HOST DEFENSE*

2.1 Overview

Transcriptional gene silencing (TGS) can serve as an innate immunity against invading DNA viruses throughout Eukaryotes. Geminivirus code for TrAP protein to suppress the TGS pathway. Here, we identified an *Arabidopsis* H3K9me2 histone methyltransferase, Su(var)3-9 homolog 4/Kryptonite (SUVH4/KYP), as a bona fide cellular target of TrAP. TrAP interacts with the catalytic domain of KYP and inhibits its activity *in vitro*. TrAP elicits developmental anomalies phenocopying several TGS mutants, reduces the repressive H3K9me2 mark and CHH DNA methylation, and reactivates numerous endogenous KYP-repressed loci *in vivo*. Moreover, KYP binds to the viral chromatin and controls its methylation to combat virus infection. Notably, *kyp* mutants support systemic infection of TrAP-deficient Geminivirus. We conclude that TrAP attenuates the TGS of the viral chromatin by inhibiting KYP activity to evade host surveillance. These findings provide new insight on the molecular arms race between host antiviral defense and virus counter defense at an epigenetic level.

* Reprinted with permission from "Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense" by Claudia Castillo-González, Xiuying Liu, Changjun Huang, Changjiang Zhao, Zeyang Ma, Tao Hu, Feng Sun, Yijun Zhou, Xueping Zhou, Xiu-Jie Wang, and Xiuren Zhang. eLife 2015,4:e06671. doi: 10.7554/elife.06671, Copyright 2015 by eLife Sciences Publications Ltd.

2.2 Introduction

RNA silencing is a host defense mechanism to combat invading nucleic acids. One type of RNA silencing is referred to as post-transcriptional gene silencing (PTGS). In PTGS, double-stranded RNAs (dsRNAs) are processed by Dicer-like ribonucleases into small-interfering RNAs (siRNAs). Mature siRNAs are incorporated into an Argonaute (AGO)-centered RNA-induced silencing complex (RISC) to regulate expression of target genes through RNA cleavage or translational repression. PTGS has evolved as a universal defense response toward all viruses because dsRNAs can result from intermediates in RNA virus replication, highly structured RNA virus genomes, or from viral transcripts. To evade this surveillance mechanism, virtually all plant viruses are known to encode suppressor proteins that are able to block different key steps of the PTGS pathway (Ding and Voinnet, 2007).

While the host/virus battle at the PTGS level has been well appreciated, virus suppression at a transcriptional gene silencing (TGS) level is poorly understood. In eukaryotes, the nuclear DNA is wrapped onto histone octamers to constitute chromatin. The chromatin undergoes various DNA and histone methylations, and these modifications have variable effects on gene expression depending on the precise residues, contexts, and modification complexity. Histone methylation takes place on lysine and arginine residues of the amino-terminal tails (Greer and Shi, 2012; Kouzarides, 2007). The prevailing dogma is that histone 3 lysine 4 tri-methylation (H3K4me3) is mostly associated with transcriptionally active euchromatin, while H3K9me2 and H3K27me3 are repressive marks (Deal and Henikoff, 2011; Feng and Jacobsen, 2011). Histone methylation is

catalyzed by SET domain containing methyltransferases, specifically, H3K9me2 is deposited by *Arabidopsis* Su(var)3-9 homolog 4, Kryptonite (KYP) (Du et al., 2014), and its paralogs (SUVH5,6) in *Arabidopsis*, while H3K27 methylation is carried out by the PRC2, which includes Curly Leaf (CLF) (Liu et al., 2010; Zheng and Chen, 2011). Local H3K9me2 and H3K27me3 can spread over wide regions to elicit heterochromatin configuration. In animals, the propagation of histone methylations entails co-repressor heterochromatin protein 1 (HP1), whereas in plants, KYP acts synergistically with DNA methyltransferases (i.e., Chromomethylase 3 [CMT3]) to constitute a mutually reinforcing cycle of DNA and histone methylation to secure TGS (Du et al., 2012, 2014).

Histone methylation not only regulates endogenous gene expression but also invasive DNAs such as transposons and viruses (Narasipura et al., 2014). Plant DNA viruses, exemplified by Geminivirus, form minichromosomes in the host (Hanley-Bowdoin et al., 2013). Both Geminivirus DNA and associated histones are methylated in infected cells, whereas viral methylation is reduced in methylation-deficient hosts, methylation-compromised *Arabidopsis* mutants are hypersusceptible to Geminivirus infection and show exacerbated disease symptoms (Raja et al., 2008). Thus, plants appear to employ methylation of viral chromatin to limit viral replication and transcription (Aregger et al., 2012; Pumplin and Voinnet, 2013). On the other hand, geminiviruses encode a multi-functional protein called transcriptional activation protein (TrAP/AL2/AC2) that counters the epigenetic defense (Buchmann et al., 2009; Raja et al., 2008). It has been shown that TrAP inhibits adenosine kinase (ADK) (Wang et al., 2005). ADK catalyzes the synthesis of 5′ AMP from adenosine and ATP, a process that promotes

the regeneration of S-adenosyl-methionine (SAM), the major methyl donor in the cell (Buchmann et al., 2009; Moffatt et al., 2002). Consequently, the TrAP-mediated inhibition of ADK activity likely impedes downstream trans-methylation events, including viral chromatin methylation in the nucleus (Bisaro, 2006; Buchmann et al., 2009). In parallel, some Geminivirus encode a TrAP positional homolog, named C2, that is able to stabilize SAM decarboxylase 1 to downregulate the methyl group metabolism (Zhang et al., 2011c). It seems that interfering with the methyl cycle is a common suppression mechanism for Geminivirus-encoded TrAP/AL2/C2 proteins. In addition, C2 also subverts the activity of COP9 signalosome to inhibit jasmonate signaling (Lozano-Duran et al., 2011), suggesting its multiple functions in viral counter-defense.

Here, we investigated the suppression mechanism of TrAP proteins, encoded by two Geminivirus members, *Tomato Golden Mosaic Virus* (TGMV) and *Cabbage Leaf Curl Virus* (CaLCuV). We found that constitutive expression of TGMV-*TrAP* in *Arabidopsis thaliana* caused morphological abnormalities that mimic loss-of-function mutants of numerous TGS components including *lhp1* (*like-heterochromatin1*) and *clf*. Microarray analyses of *TrAP* transgenic plants and *lhp1* mutants revealed a substantial overlap in reprogrammed host genes at a genome-wide level. Through biochemical screening, we identified KYP as the bona fide target of TrAP. We demonstrated *in vitro* that TrAP binds to the catalytic domain of KYP and inhibits its enzymatic activity; while *in vivo*, *TrAP* decreases the repressive H3K9me2 marks and H3K9me2-dependent CHH methylation in gene-rich regions. We also found that KYP directly associates with the Geminivirus minichromosome and deposits H3K9me2 marks

on viral chromatin. In addition, *kyp* mutants but not wild-type plants sustain low systemic infection of CaLCuV lacking TrAP protein. Taken together, we propose that KYP-catalyzed H3K9me2 is a line of the innate immunity against invading DNA pathogens, and Geminivirus TrAP functions to inactivate KYP to counter host defense. Thus, this study provides new insight into the host–virus interaction at the TGS level.

2.3 Results

2.3.1 TGMV-encoded TrAP causes developmental abnormalities in Arabidopsis but not through miRNA pathway

To study the suppression mechanism of TrAP, we generated 235 *Arabidopsis* transgenic lines overexpressing full-length TGMV TrAP, with or without Flag-Myc4 (FM), 3HA, or CFP epitopes. These transgenic plants were confirmed by northern (data not shown) or western blot assays (Figure 17A, Figure 21A). Importantly, the majority of the transgenic lines exhibited developmental abnormalities consisting of short statues, strongly upward curled cotyledons and true leaves (Figure 17B). Moreover, these overexpressing lines exhibited early flowering compared to wild-type (WT) plants. These phenotypes were morphologically distinct from loss-of-function mutants of *ADK1 ADK2* (Moffatt et al., 2002; Weretilnyk et al., 2001), *SnRK1* (Shen et al., 2009, 2014a), *PEAPOD2* (Lacatus and Sunter, 2009), and *rgsCaM* (Chung et al., 2014), a calmodulin-like protein, which are also targets or partners of TrAP. This result indicated that TrAP exerts some novel cellular function (s).

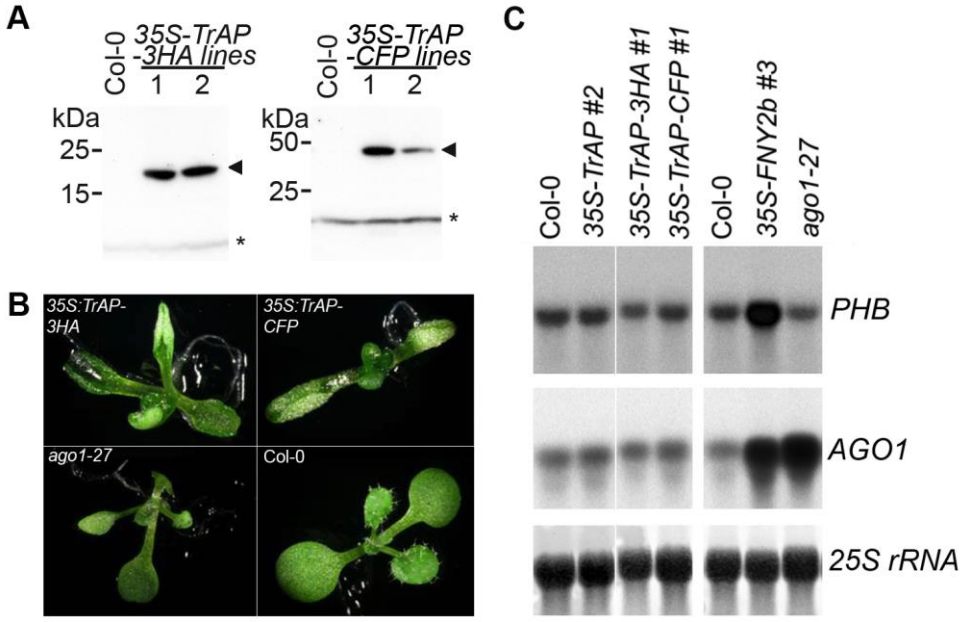


Figure 17. TrAP caused developmental abnormalities in Arabidopsis, but not through miRNA pathway.

(A) Western blot analysis of *35S-TrAP-3HA* and *35S-TrAP-CFP* transgenic lines. Arrows indicate the locations of the tagged TrAP proteins; * Cross-reaction band serves a loading control. (B) Morphological defects of Arabidopsis transgenic plants expressing *35S-TGMV TrAP*. Photographs were taken of 10-day seedlings. (C) RNA blot analysis of *PHB* and *AGO1* transcripts in the *TrAP* overexpression transgenic plants using gene-specific random-labeled probes. *25S rRNA* is a loading control. (D) sRNA blot analysis of miRNA and miRNA* in the *TrAP* overexpression transgenic plants. Total RNA was prepared from a pool of T2 transgenic plants (n>50 for each line). sRNA blots were probed using 5' end ^{32}P-labeled oligonucleotide probes complementary to the indicated miRNA or miRNA*. *5S rRNA* and *tRNA* are a loading control. All the samples were run in the same gel; but the lane order of miRNA*s was rearranged to match that of miRNAs. (E) sRNA blot analysis of miR165 loading into Arabidopsis RISCs. RNA was extracted from flowers or Flag-AGO1 immunoprecipitates of transgenic plants harboring *35S-TrAP* or *35S-TrAP-3HA* in *ago1-36; P_{AGO1}-Flag-AGO1* background and *ago1-36; P_{AGO1}-Flag-AGO1* control plants (Baumberger and Baulcombe, 2005). (Top panel), total RNA; (Middle panel), each lane contained sRNA associated with Flag-AGO1 immunoprecipitated from 0.4 g of flowers. (Bottom panel), the input and immunoprecipitate of Flag-AGO1 were analyzed by Western blot assays in the same samples for sRNA blots. A cross-reacting band (*) was used as a loading control.

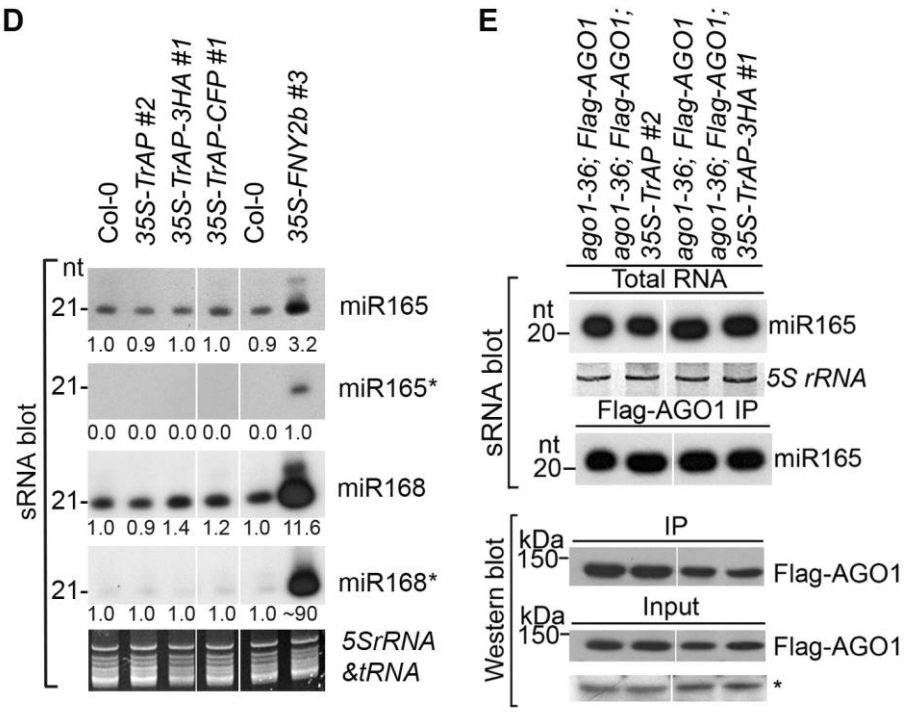


Figure 17 Continued.

Developmental anomalies of transgenic plants expressing viral suppressors have been associated with interruption of the miRNA pathway. To test this, we compared expression levels of several miRNAs and their targets between Col-0 wild-type and *TGMV TrAP* transgenic plants. Plants expressing *Cucumber mosaic virus*-encoded 2b suppressor (Zhang et al., 2006a) and *ago1-27*, a hypomorphic allele of *ago1* (Morel et al., 2002), were used as controls. We observed that the accumulation of miR165 and miR168 and their targets, *PHB* and *AGO1* transcripts, in the *35S-TGMV TrAP* lines was comparable to the amount in wild-type plants (Figure 17C, D). We further confirmed that loading of miRNAs into AGO1-centered RISCs was not affected by TrAP (Figure 17E). The same

results were obtained with miR167, miR159, and ta-siRNA480/255 and their corresponding targets (Figure 18). Thus, unlike most of previously reported viral suppressors, TrAP does not act on the miRNA pathway.

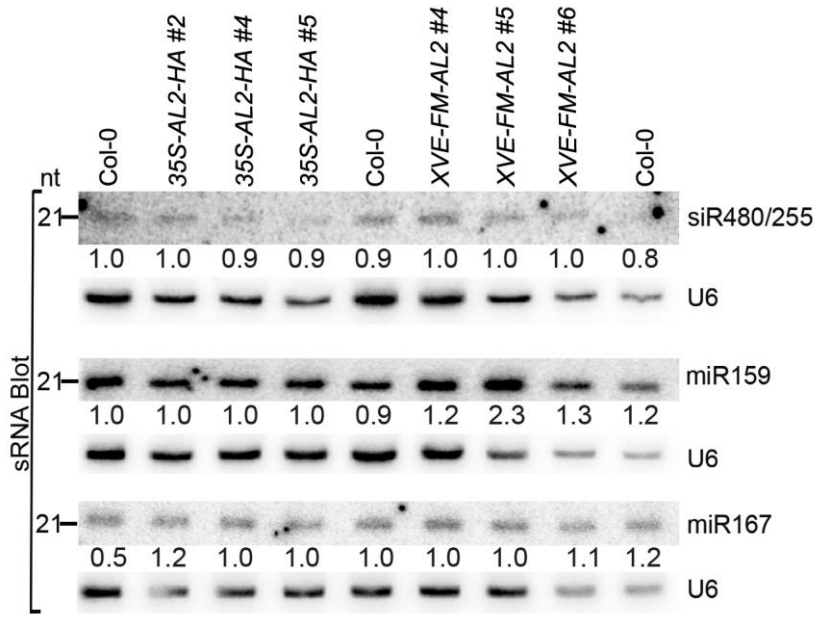


Figure 18. sRNA blot analysis of additional miRNAs and siRNA in the TrAP overexpression transgenic plants.
Total RNA was prepared from a pool of T2 transgenic plants (n>50 for each line). sRNA blots were probed using 5' end 32P-labeled oligonucleotide probes complementary to the indicated miRNAs or siRNA. U6 serves as loading control.

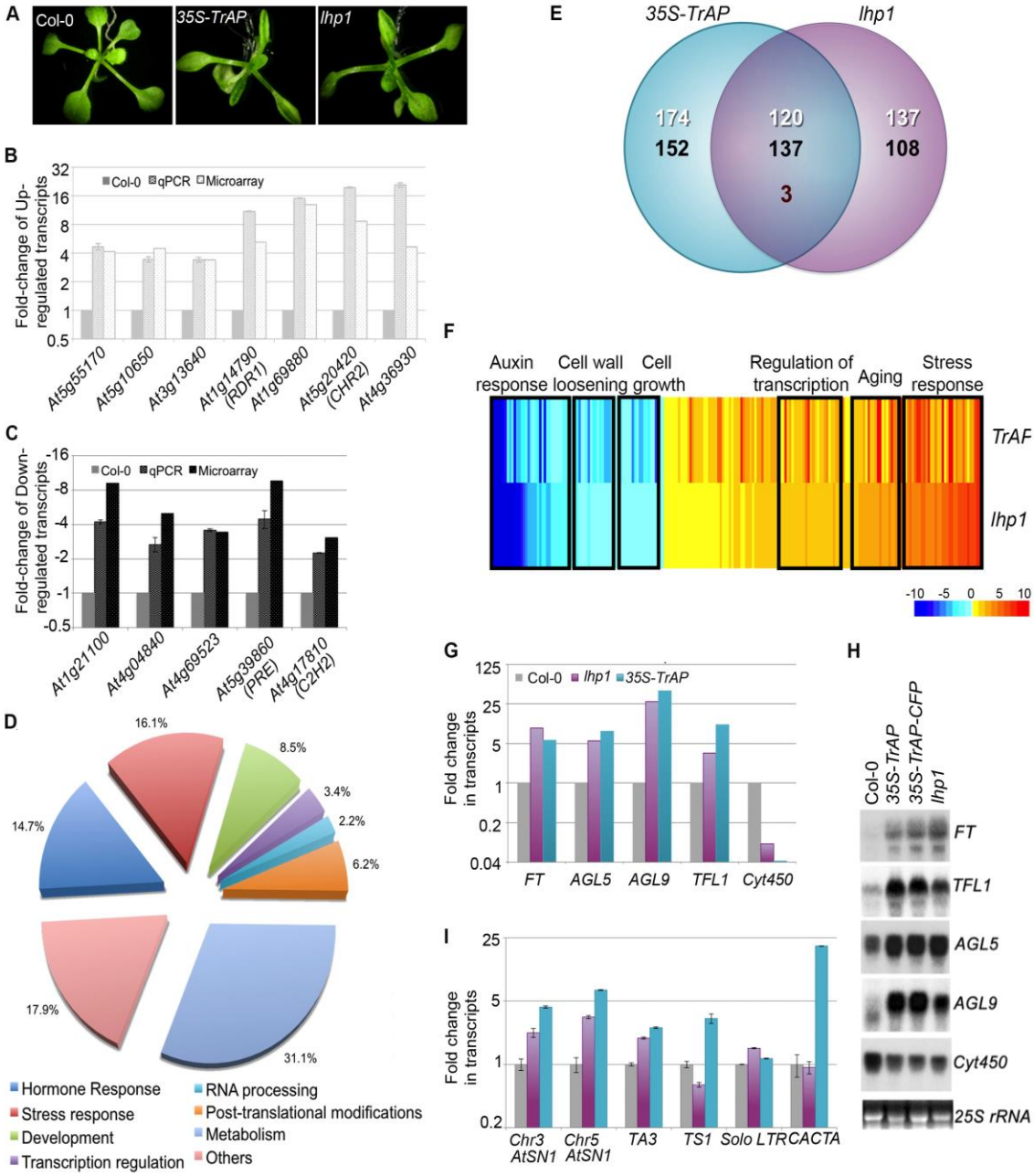
2.3.2 TrAP genetically interrupts the TGS pathway

To study how *TrAP* altered plant development, we mined publicly available databases and literature for the molecular and morphological phenotype of *35S-TrAP* lines. We found that*35S-TrAP* transgenic lines phenocopied several mutants in the

epigenetic pathway including *LHP1* (Kotake et al., 2003) (Figure 19A) and *clf* mutants (Chanvivattana et al., 2004), with respect to the early flowering and upward curling of leaves. CLF belongs to PRC2, a complex that catalyzes the deposition of H3K27me3 marks. LHP1 (Nakahigashi et al., 2005), on the other hand, associates to silent genes in euchromatin and directs the spreading of the silent status to adjacent loci (Farrona et al., 2008; Turck et al., 2007; Zhang et al., 2005, 2007a; Zheng and Chen, 2011). Thus, coordinate activities of CLF and LHP1 result in chromatin methylation and transcriptional repression (Farrona et al., 2008).

We examined global expression profiles of 7-day-old *35S-TrAP* transgenic plants compared to Col-0 wild-type using an Affymetrix ATH1 GeneChip and identified 586 genes that are differentially expressed in the *35S-TrAP* transgenic plants ($q < 0.005$). Of these, 295 transcripts were elevated whereas 291 were reduced (Figure 19E). We performed real-time PCR and RNA blot assays to validate the microarray results for the differentially expressed genes (DEGs). Among 25 genes randomly tested, we confirmed the ATH1 results for 24, indicating that the microarray results were reliable (Figure 19B, C; data not shown). Gene-Ontology (GO) analysis placed the DEGs into seven functional categories (Figure 19D; Supplementary file 1): hormone response (86 genes), stress response (94 genes), development regulation (50 genes), transcriptional regulation (20 genes), RNA metabolism (13 genes), post-translational modification (36 genes), and general metabolism (182 genes), plus a set of 105 un-annotated genes (Figure 19D; Supplementary file 1).

Figure 19. TrAP is genetically involved in the TGS pathway.
(A) 35S-TrAP transgenic plants phenocopied lhp1 mutants. Photographs were taken of 15-day seedlings. (B and C) Microarray results were validated by qRT-PCR analysis. Only 12 randomly selected loci were shown. (D) Gene ontology analysis of the TrAP-regulated DEGs. The numbers adjacent to the pies represent the ratio of genes in each category over the total DEGs (E) Genome-wide overlapping of the genes regulated by TrAP and loss-of-function lhp1. White and black numbers correspond to upregulated and downregulated genes, respectively. Maroon number indicates the genes that are differentially deregulated in both genotypes. (F) Heatmap of the commonly deregulated genes in the 35S-TrAP and lhp1 lines. The typical gene-ontology categories are shown on top. (G and H) Microarray and RNA blot analyses of epigenetically regulated flowering genes in the TrAP transgenic lines and lhp1 mutants. Cyt450 is a control. (I) qRT-PCR analysis of TEs in heterochromatic regions in the lhp1 mutant and TrAP transgenic lines.


A
Col-0
35S-TrAP
lhp1
B
Fold-change of Up-regulated transcripts
Col-0 qPCR Microarray
32
16
8
4
2
1
0.5
At5g55170
At5g10650
At3g13640
At1g14790 (RDR1)
At1g69880
At5g20420 (CHR2)
At4g36930
C
Fold-change of Down-regulated transcripts
Col-0 qPCR Microarray
-16
-8
-4
-2
-1
-0.5
At1g21100
At4g04840
At4g69523
At5g39860 (PRE)
At4g17810 (C2H2)
D
16.1%
8.5%
3.4%
2.2%
6.2%
31.1%
17.9%
14.7%
Hormone Response
Stress response
Development
Transcription regulation
RNA processing
Post-translational modifications
Metabolism
Others
E
35S-TrAP
lhp1
174
152
120
137
3
137
108
F
Auxin response
Cell wall loosening
Cell growth
Regulation of transcription
Aging
Stress response
TrAF
lhp1
-10 -5 0 5 10
G
Fold change in transcripts
Col-0 lhp1 35S-TrAP
125
25
5
1
0.2
0.04
FT
AGL5
AGL9
TFL1
Cyt450
H
Col-0
35S-TrAP
35S-TrAP-CFP
lhp1
FT
TFL1
AGL5
AGL9
Cyt450
25S rRNA
I
Fold change in transcripts
25
5
1
0.2
Chr3 AtSN1
Chr5 AtSN1
TA3
TS1
Solo LTR
CACTA

Then, we compared the DEG profiles of the *TrAP* overexpression lines and loss-of-function *lhp1* mutant (Figure 19E, F; Supplementary file 2). Transcriptome analysis revealed that out of 295 genes significantly upregulated in the *35S-TrAP* transgenic lines, 120 (40.7%) were also upregulated in *lhp1* mutant. This is significantly greater than 1.28% expected by chance ($p < 2.2e{-16}$, Pearson's Chi-squared test). Interestingly, among the co-upregulated genes are a group of flowering-stimulated transcriptional factors including the key flowering-time integrator, *FT*, and 12 other genes such as *TFL1*, *AGL5*, and *AGL9* (Farrona et al., 2008) (Figure 19F-H; Supplementary file 2; Supplementary file 8) clustered in the transcriptional regulation category. Importantly, all these genes are regulated through epigenetic pathways and account for the early flowering phenotypes of *lhp1* mutant and possibly of *TrAP* transgenic plants as well (Gan et al., 2013). Other highly represented categories included 31 genes involved in aging and 116 genes engaged in stress responses (Supplementary file 2). Notably, the stress-responsive genes included genes specific to biotic stress such as *PR4*, *WRKY18*, *FLS2*, and *PDF1.2*; additionally, genes related to chemical stress, such as *PTR3* and *TAT3*, were also identified. Thus, constitutive expression of *TrAP* could trigger plant senescence and innate defense pathways, and this activation is potentially through interference with the *LHP1*-related epigenetic silencing.

Similarly, out of 291 genes significantly downregulated in the *35S-TrAP* transgenic lines, 137 (47.1%) were also repressed in *lhp1* mutant. This is significantly greater than 1.25% expected by chance ($p < 2.2e{-16}$, Pearson's Chi-squared test) (Supplementary file 2). Genes related to auxin response were of special interest. Of

the 32 DEGs involved in the auxin pathway, 29 genes were downregulated in the *35S-TrAP* transgenic plants, classified as small auxin upregulated mRNAs (*SAUR*s). These results suggested a possible hyposensitivity to auxin in *lhp1* mutants and *TrAP* transgenic plants, which could explain the smaller statues of both genotypes. Consistent with this hypothesis, numerous auxin-repressed loci including *PS2* and aging genes like *TET9*, *SAG13*, and *SRG1* were upregulated in both lines. Concomitantly, six genes related to cell growth and five genes engaged in cell wall loosening were also repressed. Significant genome-wide overlap of*35S-TrAP* and *lhp1* loss-of-function-responsive genes suggested that *TrAP* might be genetically involved in the *LHP1*-related TGS pathway.

Since LHP1 is believed to reside in euchromatic regions, we wondered whether *TrAP* also deregulates expression of heterochromatic loci. To this end, we selected numerous transposable elements (TEs) that were not recovered from the microarray assays and assessed them directly by qRT-PCR. Excitingly, most of the tested transposons were transcriptionally active (Figure 19I), further suggesting that *TrAP* indeed inhibits the TGS pathway, in both euchromatic and heterochromatic regions.

2.3.3 KYP is a bona fide target of TrAP

Given that *TrAP* transgenic plants phenocopied several TGS mutants and displayed transcriptional activation of heterochromatic loci, we hypothesized that TrAP epistatically regulates a TGS integrator (s), indirectly leading to deregulation of the epigenetic marks. Analysis of the microarray data challenged this possibility as no significant changes in the transcripts of any canonical TGS components were revealed

(Supplementary file 3). An alternative hypothesis was that TrAP directly interferes with the function of a TGS component (s). To test this in an unbiased manner, we used luciferase complementation imaging (LCI) assay to screen 34 TGS-related proteins and some other cellular factors for TrAP interaction (Zhang et al., 2011c) (Supplementary file 4). In the LCI experiments (Figure 20A), the N- and C-terminal parts of firefly luciferase (NLuc and CLuc) are fused to different test proteins to be transiently expressed in *Nicotiana benthamiana* (*N. benthamiana*). When NLuc and CLuc are brought together through interaction of the test proteins, catalytic activity is restored and recorded through CCD camera. In our LCI screening, we recovered LHP1 and KYP, a SUVH-type H3K9me2 methyltransferase, suggesting that TrAP is physically close to LHP1 and/or KYP proteins *in vivo* (Figure 20B). Next, we carried out confocal microscopy imaging assays. TrAP co-localized with both LHP1 and KYP in scattered but not yet clearly defined nuclear foci, whereas co-expression of TrAP-CFP and other YFP-tagged proteins in *N. benthamiana* cells did not show such patterns (Figure 20C). These observations further suggested that TrAP was in the same complexes as LHP1 or KYP. To further examine if TrAP interacted with these proteins, we conducted co-immunoprecipitation (Co-IP) assays (Figure 20D). Interestingly, we validated the TrAP-KYP interaction (Figure 20D) but not TrAP-LHP1 (data not shown), indicating that the LCI signal resulting from the TrAP-LHP1 combination likely involved additional cofactors between TrAP and LHP1 *in vivo* (Figure 20B). Alternatively, TrAP-LHP1 interaction might be transient or unstable in our stringent co-IP condition. We observed that expression of CLuc-HA3-KYP in *N. benthamiana* yielded truncated proteins of various lengths that accumulated to

comparable levels as the full-length protein; only the full-length KYP showed specific interaction with TrAP, implying the KYP C-terminal domain as the interaction interface with TrAP (Figure 20D). We further confirmed the *in vivo* TrAP-KYP interaction by Förster resonance energy transfer-acceptor bleaching (FRET-AB), using TrAP-CFP as a donor and YFP-KYP as an acceptor. Shortly, FRET involves the energy transfer from an excited donor to an adjacent acceptor when the fluorophores are less than 10 nm apart. If the fluorophores are coupled, the excited donor leads to acceptor emission; bleaching the acceptor allows the emission of the donor to be measured (Figure 20E). Consecutive cycles of YFP-KYP bleach/recovery correlated with release/quench on the TrAP-CFP signal (Figure 20E-F); the positive FRET-AB (CaLCuV-TrAP 2.699 ± 0.23, TGMV-TrAP 3.8228 ± 0.58, $p < 0.05$) corroborated TrAP-CFP/YFP-KYP interaction (Figure 20G).

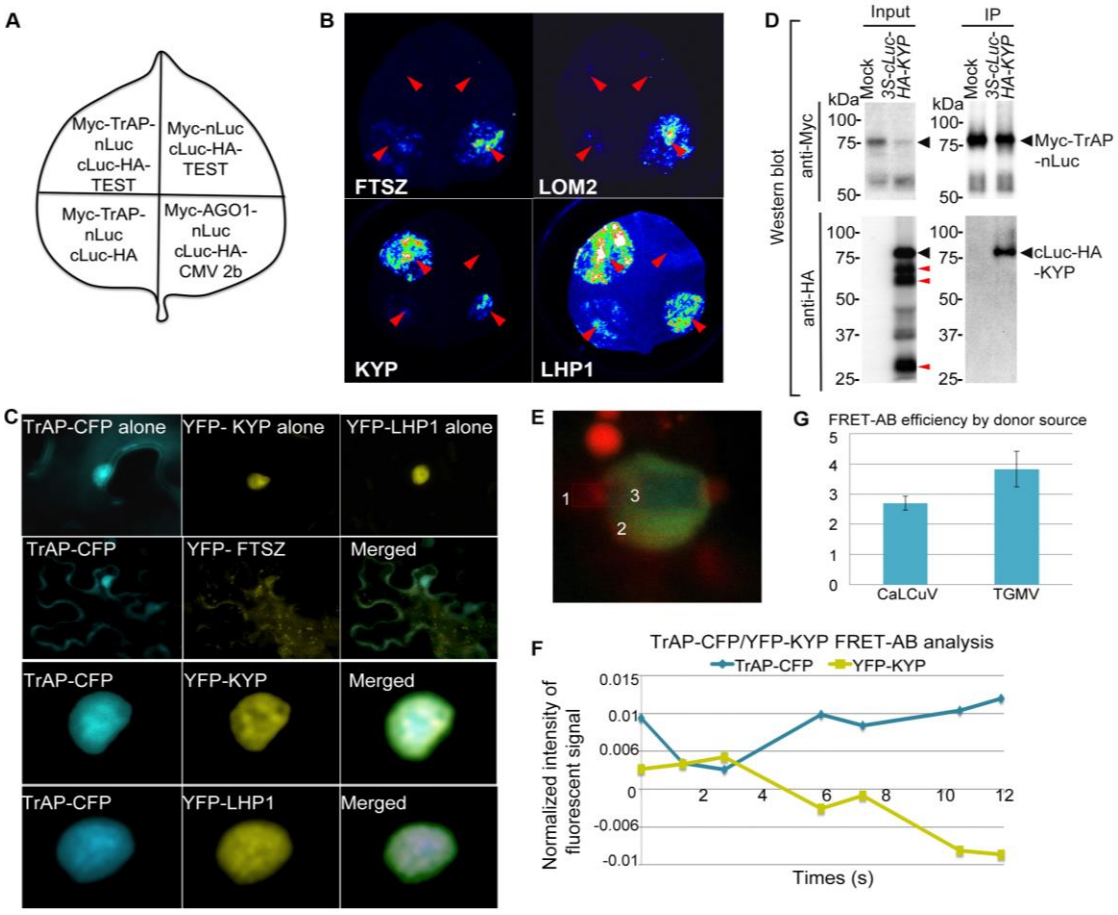


Figure 20. TrAP interacts with KYP *in vivo*.
(A) Schematic representation of the luciferase complementation imaging assay. The bottom panel shows different combinations of infiltrated constructs that were fused either to N-terminal (NLuc) and C-terminal (CLuc) regions of luciferase. (B) Screening of host factors targeted by TrAP. The infiltration positions of the constructs (red arrows) and luminescence signal resulting from the protein-protein interaction in a leaf are shown. FTSZ and LOM2 serve as negative controls. (C) Confocal imaging assays show the co-localization of TrAP-CFP with YFP-KYP in the nuclei in *N. benthamiana*. FTSZ serves as a negative control. (D) Specific interaction between KYP and TrAP was confirmed in *N. benthamiana* by co-immunoprecipitation (Co-IP). Constructs harboring *35S-Myc-TrAP-nLuc* and *cLuc-HA-KYP* were co-infiltrated in *N. benthamiana* leaves. IP was conducted using anti-Myc antibody. Western blot analyses were done with the crude extract (input) and the IP products using anti-Myc, or -HA antibodies. Truncated versions (red arrows) serve as an internal control. (E) Exemplary imaging of FRET assays of TrAP-CFP and YFP-KYP co-expressed in a nucleus. The nucleus is irradiated with 458nm laser to excite the CFP fluorophore. Three regions were selected for the assay: #1, autofluorescence control, #2 fluorophore decay control, and #3, FRET-Acceptor Bleaching test. Regions #1 and #3 were treated with pulses of 514nm laser to bleach the YFP fluorophore. The CFP signal is then visible in region #3 when the emission of CFP is dequenched. (F) Quantification of the signals from each fluorophore observed during FRET-AB experiment in E. (G) FRET is positive for YFP-KYP paired with the CFP-tagged TrAPs from either CaLCuV or TGMV.

To examine whether TrAP interacted with KYP under physiological conditions, we isolated TrAP complexes through two-step immunoprecipitation (Zhu et al., 2013) from stable transgenic plants expressing *FM-TrAP* under the inducible promoter (*XVE*) (Zuo et al., 2000) followed by mass spectrometry analysis. A total of 624 peptides representing 288 unique sequences were recovered from the TrAP sample; of those, 31 unique peptides matched specifically to KYP/SUVH4 and were not found in control immunoprecipitates using Col-0 plants (Figure 21). Together, all these assays clearly indicated that TrAP and KYP interact *in vivo*.

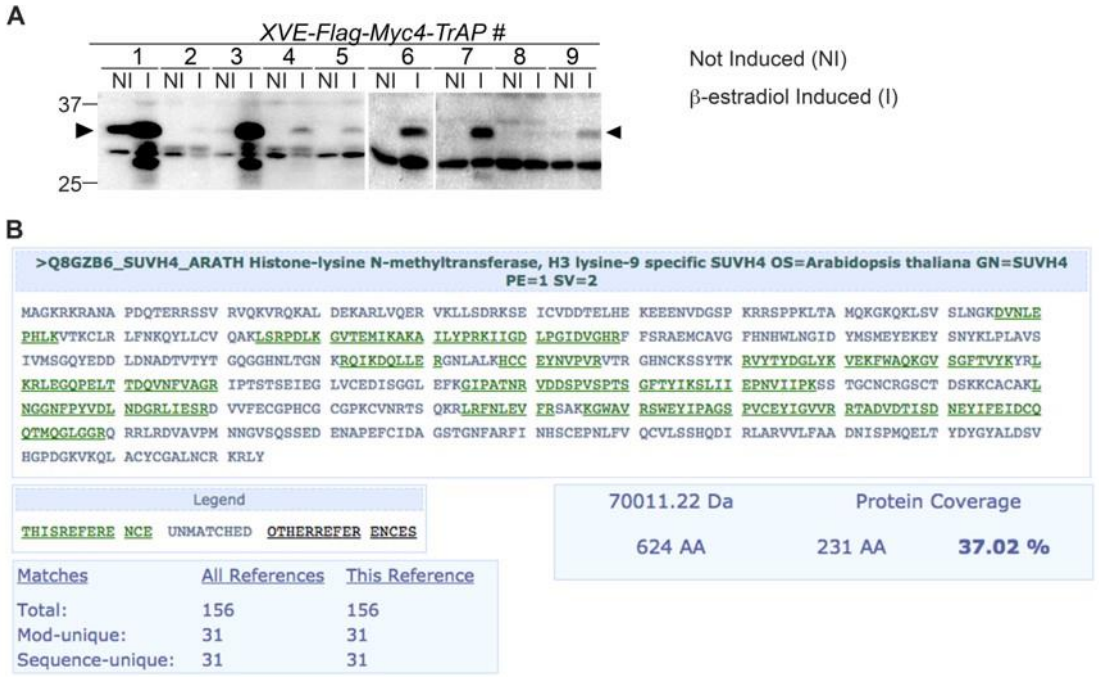


Figure 21. Mass spectrometry analyses confirmed endogenous KYP as a *bona fide* TrAP interacting partner.
(A) Western blot indicating expression of FM-TrAP in seedlings upon induction with bet-estradiol. Anti-Myc antibody as used for detection. (B) 31 peptides (green) uniquely match to the KYP sequence.

2.3.4 TrAP binds to the catalytic domain of KYP

We performed *in vitro* pull-down assays to examine whether TrAP interacts directly with KYP. We found that maltose-binding protein (MBP)-KYP, but not MBP and other MBP-tagged control proteins, was able to pull down GST-tagged TGMV- and CaLCuV-encoded TrAP proteins (Figure 23A, B). This interaction was specific as MBP-KYP was unable to pull down GST protein alone.

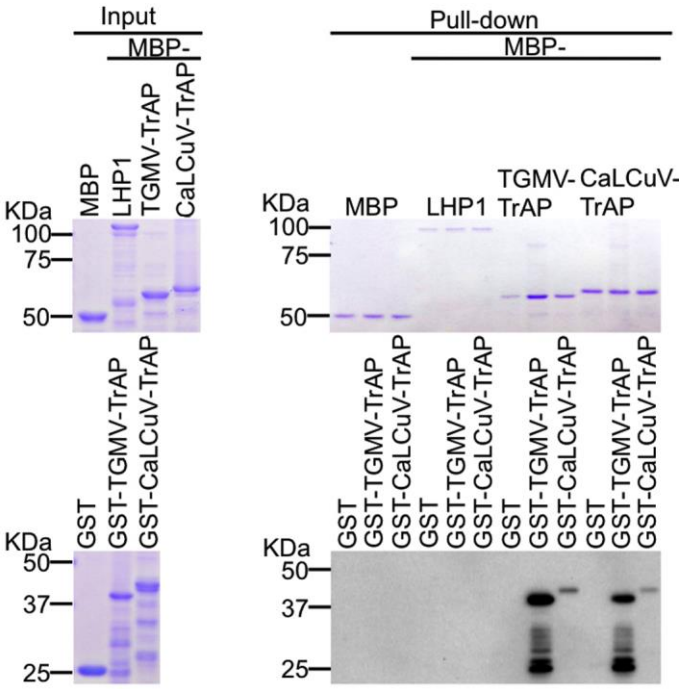
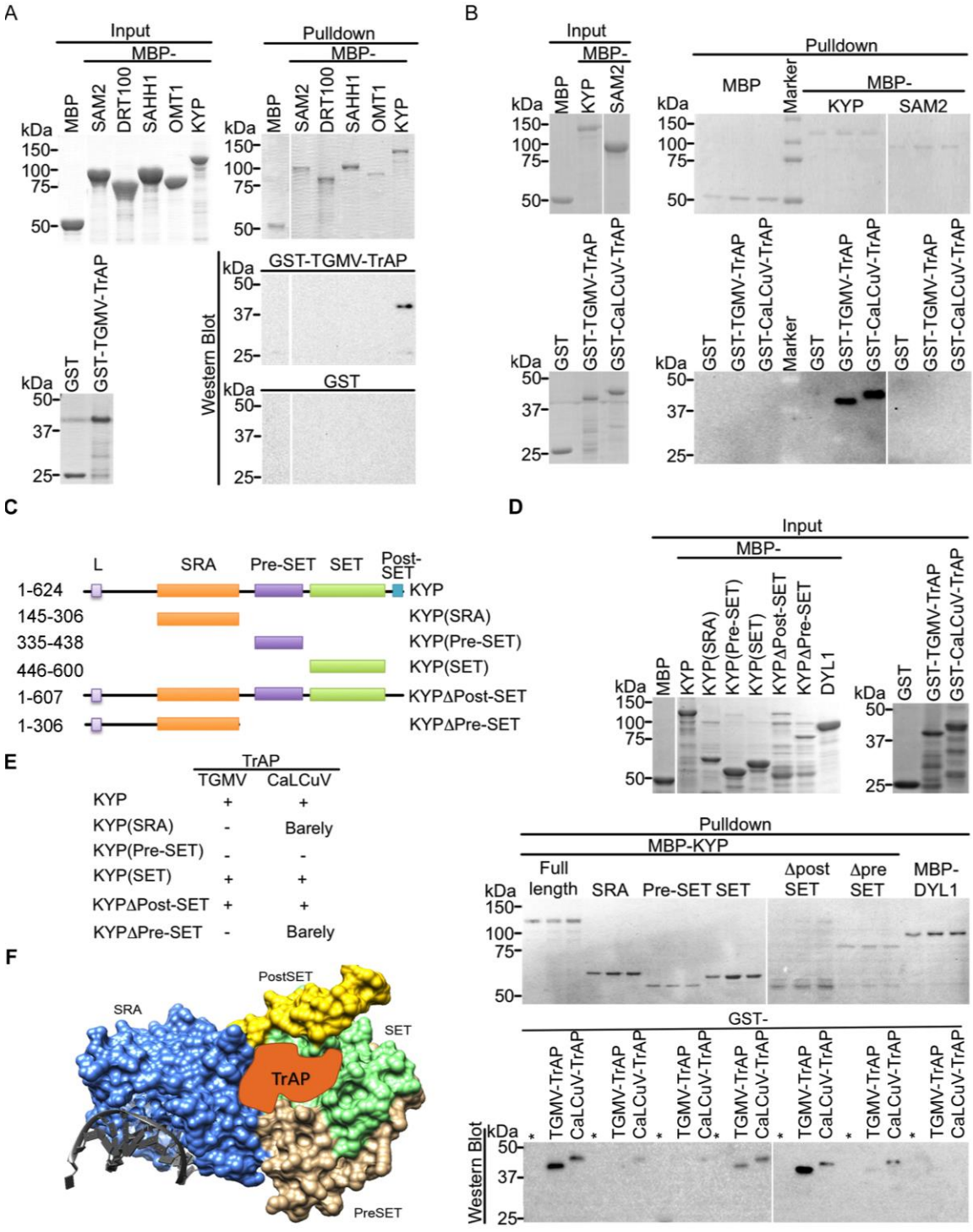


Figure 22. TrAP does not interact with LHP1 *in vitro*.
In vitro pull down assays of GST-TrAP by MBP-LHP1 is shown. Left panel, Coomassie brilliant blue R250 staining of the proteins shows their mobility. Right panel, output of *in vitro* pull-down assays. The recovered MBP-tagged bait proteins were monitored by Coomassie brilliant blue R250 staining. The output of the GST-tagged prey proteins was analyzed by western blot using a monoclonal anti-GST antibody. TrAP dimerization is shown as positive control. 2.5µg of prey proteins were pulled down with the indicated bait proteins (2.5µg each).

Notably, we did not observe direct LHP1-TrAP interaction in the parallel pull-down experiments (Figure 22); and this result was consistent with the *in vivo* Co-IP experiments. From N- to C- termini, KYP contains SRA, PreSET, SET, and PostSET domains. To further define the protein domain (s) responsible for the specific interaction, we generated five truncations of KYP (Figure 23C). Pull-down assays showed that the SET domain interacted with TrAP proteins from both TGMV and CaLCuV at a higher affinity, whereas SRA domain might interact with CaLCuV TrAP at a reduced affinity *in vitro* (Figure 23E, F). These results were consistent with the results of the *in vivo* Co-IP experiments in which only full-length KYP, but not C-terminal-truncated versions, could recover TrAP (Figure 23D). Recent structural analysis on KYP revealed that SET and pre-SET domains constitute two modules: one forms a narrow pocket harboring the H3 tail (1–15aa), whereas the other binds the SAM cofactor, together with the post-SET domain (Du et al., 2014). Because the post-SET domain does not seem to contribute to the KYP-TrAP interaction, TrAP could potentially occupy the histone-binding cavity (Figure 23F).

Figure 23. TrAP interacts directly with KYP through the SET domain.
(A and B) *In vitro* pull-down assays showed that KYP specifically interacted with TGMV TrAP (A) and CaLCuV TrAP (B). Left panel, Coomassie brilliant blue R250 staining of the proteins shows their mobility. Right panel, Output of *in vitro* pull-down assays. All His-MBP-tagged bait proteins and His-GST-tagged prey proteins were purified from E. coli using Ni-NTA columns. In all assays, 2.5μg of prey proteins were pulled down with the indicated bait proteins (2.5μg each) using amylose resins. The recovered MBP-tagged bait proteins were monitored by Coomassie brilliant blue R250 staining. All the experiments were done at the same time and the samples were run in the same gels. The spacers in the images indicate digital rearrangements of the pictures. The output of the GST-tagged prey proteins was analyzed by western blot using a monoclonal anti-GST antibody. (C) Schematic diagram of full-length and truncated forms of KYP. The numbers on the left refer to the amino acid residues in KYP protein. Locations of SRA, Pre-SET, SET, and Post-SET domains are shown. (D) *In vitro* pull-down assays of truncated KYP proteins and TrAPs. The experiments were done as in (A and B). The GST negative control was loaded in the lanes marked with (*). (E) Summary of interaction between the truncated KYP proteins and TrAP encoded by TGMV and CaLCuV. (F) Model of possible KYP-TrAP interaction based on the experimental results from panels D and E. KYP structure was generated in Chimera from PDB: 4QEN dataset, domains are color coded and indicated.


A
Input
MBP-
MBP
SAM2
DRT100
SAHH1
OMT1
KYP
kDa
150
100
75
50
Pulldown
MBP-
GST
GST-TGMV-TrAP
Western Blot
GST-TGMV-TrAP
GST
37
25
B
Input
MBP-
Pulldown
MBP
Marker
KYP
SAM2
GST-CaLCuV-TrAP
C
L
SRA
Pre-SET
SET
Post-SET
1-624
KYP
145-306
KYP(SRA)
335-438
KYP(Pre-SET)
446-600
KYP(SET)
1-607
KYPΔPost-SET
1-306
KYPΔPre-SET
D
KYP(SRA)
KYP(Pre-SET)
KYP(SET)
KYPΔPost-SET
KYPΔPre-SET
DYL1
MBP-KYP
Full length
Δpost SET
Δpre SET
MBP-DYL1
GST-
TGMV-TrAP
CaLCuV-TrAP
E
TrAP
TGMV
CaLCuV
KYP + +
KYP(SRA) - Barely
KYP(Pre-SET) - -
KYP(SET) + +
KYPΔPost-SET + +
KYPΔPre-SET - Barely
F
PostSET
SRA
SET
TrAP
PreSET

Since the SET domain is well conserved among histone methyltransferases (HMTase) (Liu et al., 2010), we next investigated whether TrAP interacts with KYP paralogs. *In vitro* pull-down assays showed that TrAP indeed interacted with numerous tested HMTases (SUVH2, 5, and 6) (Figure 24). Notably, loss-of-function mutants of *SUVH2*display early-flowering phenotype, suggesting that TrAP might target this protein *in vivo* as well.

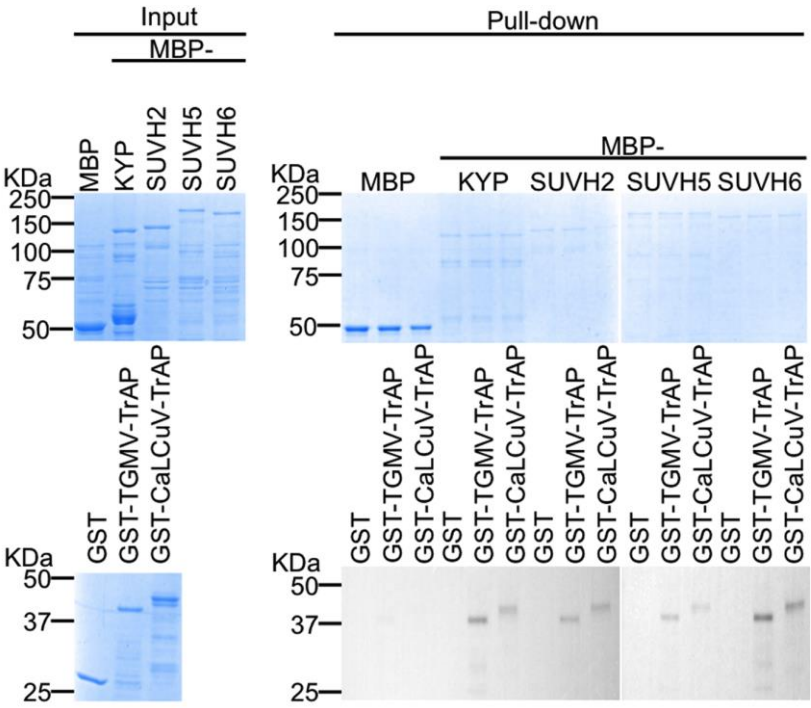


Figure 24. TrAP directly interacts with KYP paralogs SUVH2, 5, and 6.
In vitro pull down assays. Left panel, Coomassie brilliant blue R250 staining of the proteins shows their mobility. Right panel, output of *in vitro* pull-down assays. The recovered MBP-tagged bait proteins were monitored by Coomassie brilliant blue R250 staining. The output of the GST-tagged prey proteins was analyzed by western blot using a monoclonal anti-GST antibody. SUVH4/KYP was used as a positive control. 2.5 μg of prey proteins were pulled down with the indicated bait proteins (2.5 μg each). All the experiments were performed simultaneously and run in two separate gels.

2.3.5 TrAP inhibits the catalytic activity of KYP in vitro

The specific interaction of TrAP with the catalytic SET domain of the SUVHs prompted the question of whether TrAP inhibits KYP activity. To test this, we set up an *in vitro* reconstitution of H3K9 methylation using His-GST-KYP purified from *Escherichia coli* as the enzyme source, recombinant histone 3 as the substrate, and methyl-^{14}C-SAM as the methyl donor (Figure 25A). Under our experimental conditions, 1 µM GST-KYP methylated 3 µM of histone 3 in less than 5 min at 37°C, as detected by saturation of the radioactive signal (Figure 25B). Excitingly, incubation of His-GST-KYP with His-MBP-TrAP from either TGMV or CaLCuV reduced the initial velocity of KYP transmethylation activity in a dose-dependent fashion, whereas His-MBP alone did not affect KYP catalysis (Figure 25B). Quantification of signal intensity revealed that the TrAP-KYP molar ratio of 2 was enough to cause approximately 50% inhibition of KYP activity (half maximal inhibitory concentration [IC_{50}]) (Figure 25C). Thus, our results indicated that TrAP potently inhibited the catalytic function of the HMTase.

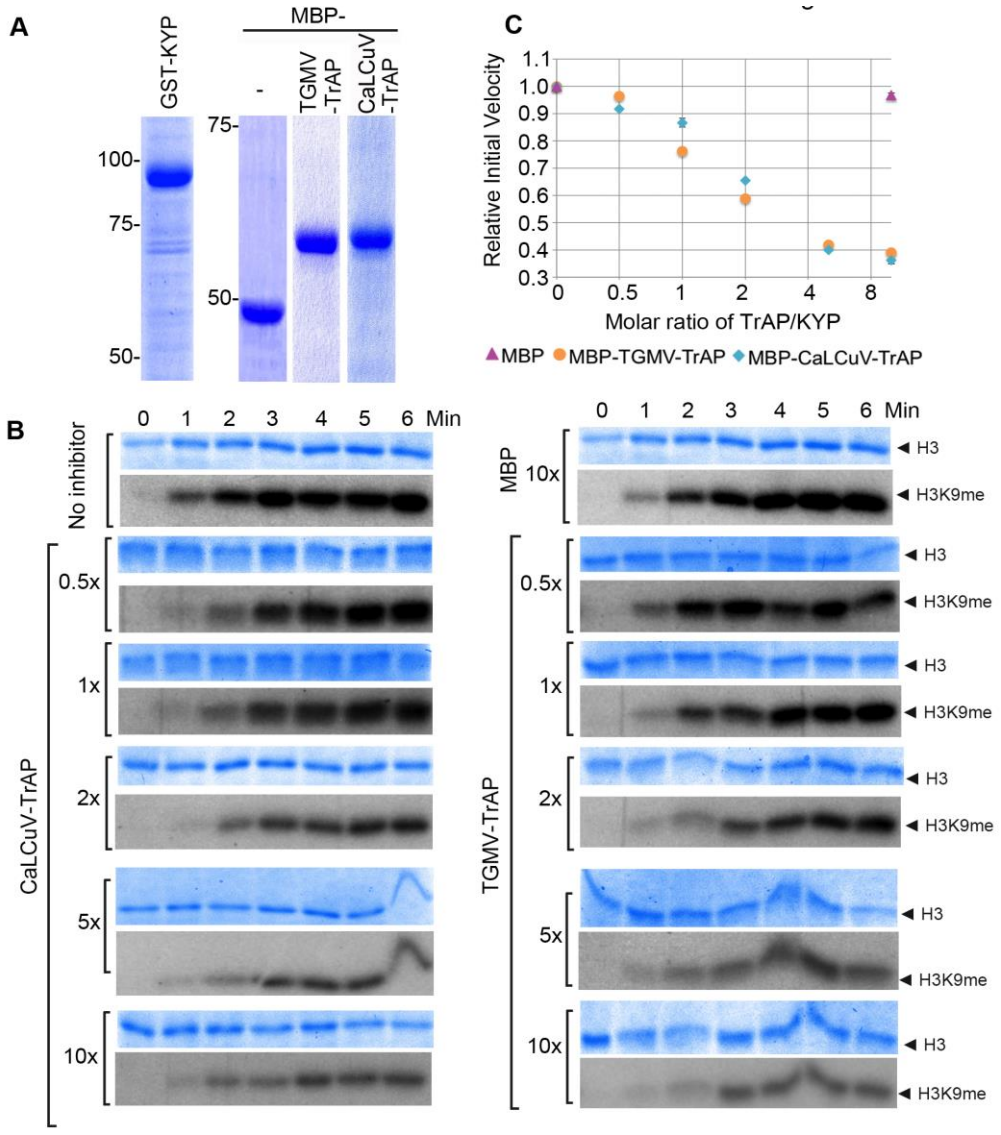


Figure 25. TrAP inhibited HMTase activity of KYP *in vitro*.
(A) Coomassie staining of purified proteins uses for the assays. (B) *In vitro* HMTase reconstitution assays with different molar ratio of MBP and MBP-TrAP proteins (0-10x) relative to GST-KYP. The recombinant KYP was incubated without (buffer only) or with the indicated proteins before addition of Histone3 and C14-SAM. The reactions were done in a 6-minutes time course; aliquots were resolved in 18% SDS PAGE, and stained with Coomassie blue R250 to show Histone3 input (top panels). The dried gels were auto-radiographed to detect 14C-methylated Histone3. (C). Plotting of KYP initial velocity vs TrAP/KYP molar ratio. The initial velocity was calculated from the slope of the linear range in the signal vs. time plot for each assay, and then the values were normalized using the non-inhibitor control as a standard of 1 to obtain the relative initial velocity with standard deviation (SD) from at least three biological repeats. The relative initial velocity is plotted as a function of inhibitor:enzyme molar ratio in a logarithmic scale of base 2.

2.3.6 TrAP reduces H3K9me2-repressive marks in vivo

Given that *in vivo* TrAP genetically interferes with the TGS pathway, and *in vitro* it physically interacts with HMTases (KYP, SUVH2, 5, 6) and inhibits the activity of KYP, we wondered whether TrAP alters KYP function *in vivo*. To address this question, we conducted chromatin immunoprecipitation (ChIP) analyses of H3K4me3, H3K9me2, and K3K27me3 marks (Figure 26, Figure 27, Figure 28) on numerous KYP-regulated TEs (Figure 27B, Figure 28A) in TrAP overexpression plants as compared to Col-0, *ddm1*, *kyp*, and *lhp1* control plants. Remarkably, all tested loci in *TrAP* transgenic lines displayed consistent reduction of H3K9me2 and H3K27me3, whereas changes of H3K4me3 were variable. This molecular phenotype mimicked *kyp* mutant at all tested loci, but not *lhp1* mutant. This result indicated inhibitory effect of TrAP on KYP activity on the TEs *in vivo*.

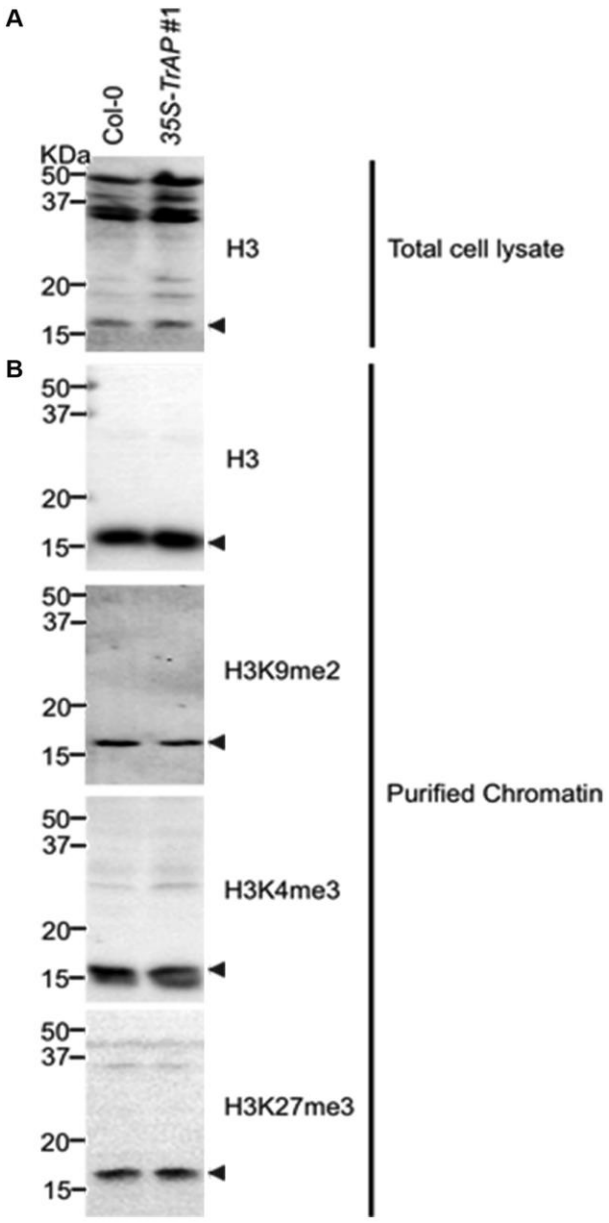


Figure 26. Western blot analysis to show specificity of antibodies used for ChIP assays in the study.

Crude extract (A) and isolated nuclei (B) were probed with antibodies against histone 3, H3K9me2 (Abcam Cat# ab1220), H3K27me3 (Millipore Cat# 07-449) and H3K4me3 (Millipore Cat# 04-745), respectively.

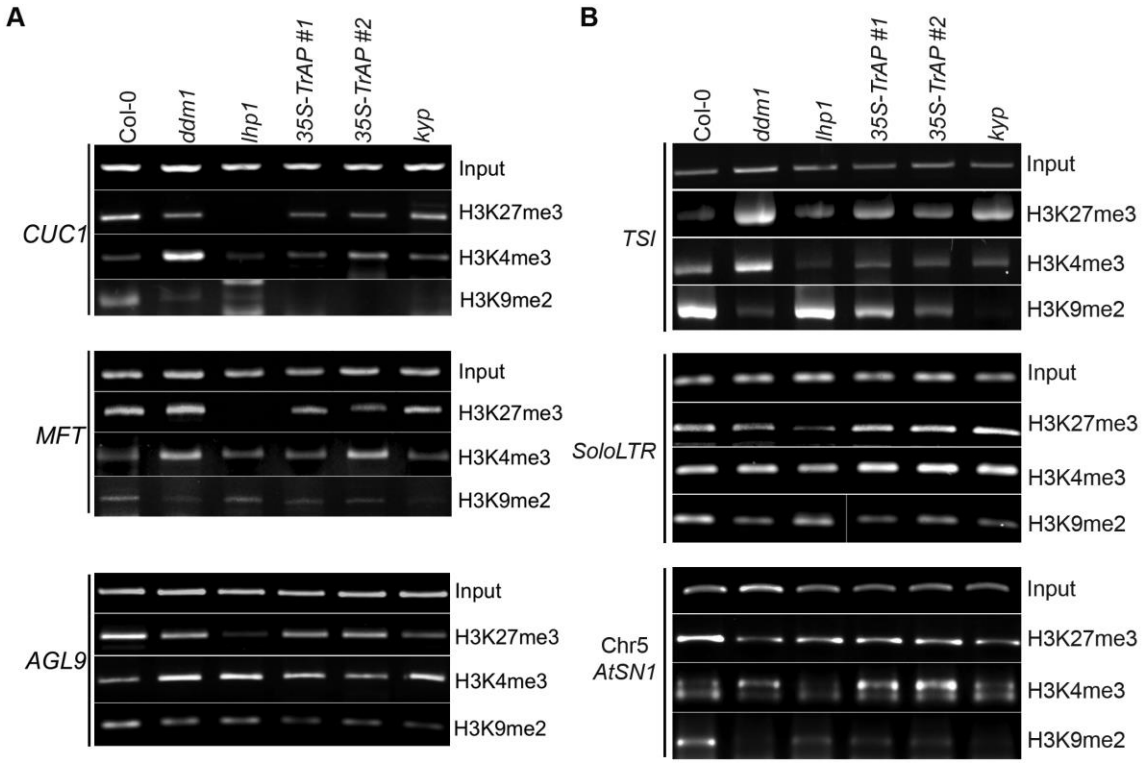


Figure 27 ChIP-PCR assays for selected flowering genes and heterochromatic loci confirm ChIP-qPCR.
(A) ChIP-PCR analysis of various histone 3 modifications in flowering genes in different genetic backgrounds. (B) ChIP-PCR analysis of various histone3 modifications in TEs in different genetic backgrounds. ChIP assays were conducted on 9-day old seedlings using antibodies specific for H3K9me2 (Abcam Cat# ab1220), H3K27me3 (Millipore Cat# 07-449) and H3K4me3 (Millipore Cat# 04-745). The PCRs were done with 22 cycles for the input samples and with 30 cycles after ChIP.

Since TrAP is a transcriptional activator protein, we next asked whether the increased transcription of the *TrAP*-responsive protein-coding genes is associated with changes in the histone methylation status. First, we screened numerous loci that showed transcriptional deregulation in the *TrAP* overexpression lines (Figure 19B-D,G,H) for the presence of various histone modifications. We identified a dozen loci in which H3K9me2

marks are easily detected in wild-type plants. We then conducted ChIP-qPCR for these loci in the *TrAP* transgenic plants.

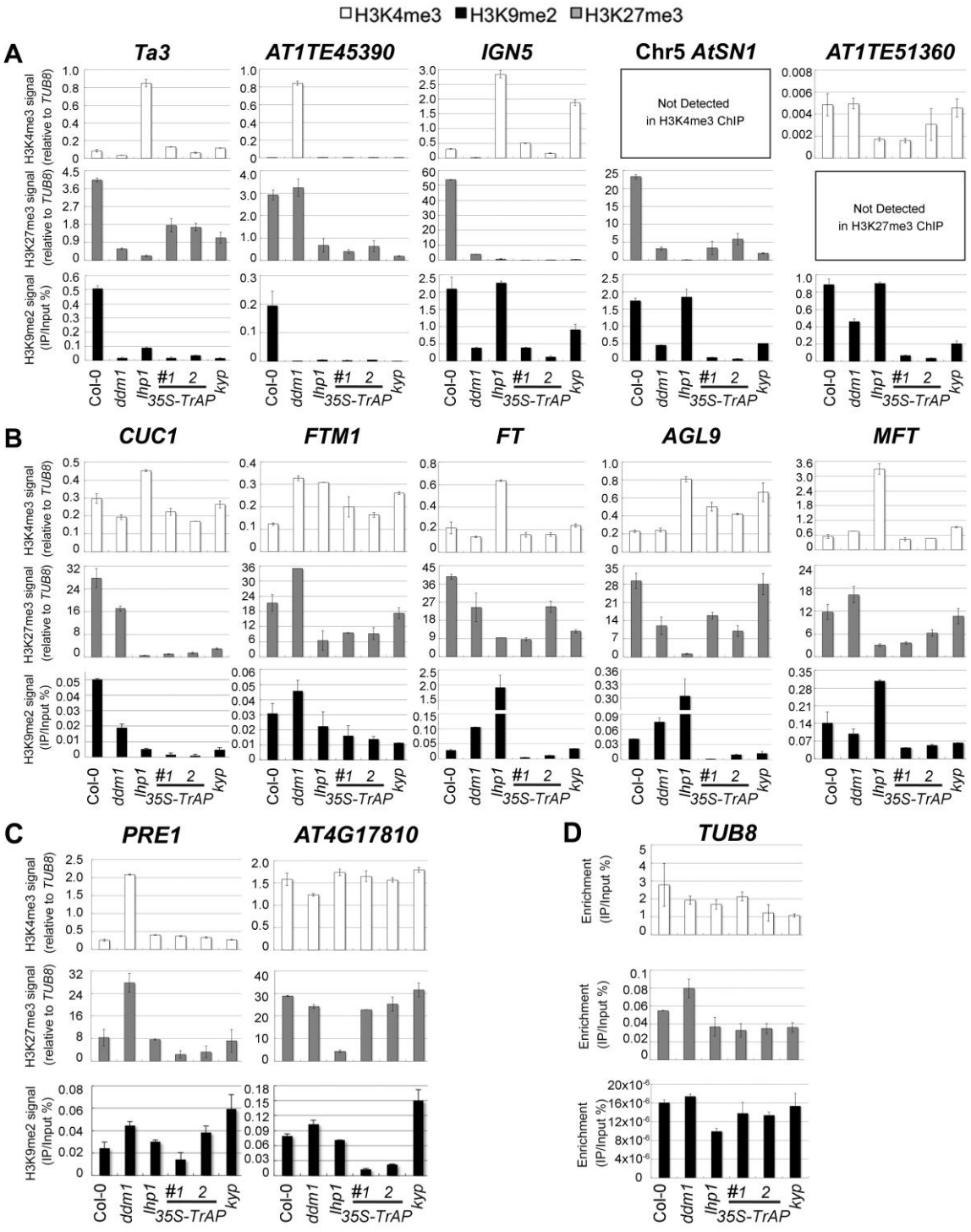
As expected, H3K9me2 and H3K27me3 were reduced in most of the tested loci in *ddm1*, *kyp* and *lhp1* mutants, whereas H3K4me3 was enriched. These results are consistent with the generally antagonistic roles of H3K4me3 and H3K9me2 modifications. Specifically, six out of seven TrAP deregulated loci including the flowering-promoting genes displayed greater than twofold reduction in H3K9me2 and H3K27me3 in the *TrAP* overexpression plants compared to wild-type Col-0, while changes in H3K4me3 were inconsistent. This scenario was similar to that observed for TEs (Figure 28B,C). Collectively, the ChIP assays indicated that TrAP interferes with the epigenetic pathways through reducing repressive H3K9me2 marks *in vivo*.

2.3.7 TrAP decreases CHH DNA methylation

KYP, SUVH5, and SUVH6 are required for maintenance of non-CG (CHG and CHH) methylation in *Arabidopsis* (Stroud et al., 2013, 2014). We predicted that inhibition of KYP function by TrAP might indirectly cause reduction in non-CG DNA methylation. To test this, we conducted genome-wide bisulfite sequencing with 11-day-old seedlings of TrAP transgenic plants, *kyp* mutant, and Col-0. Consistent with previous studies, *kyp* mutant showed genome wide loss of methylation in CHG (~42.1%) and CHH (~21.7%), but not in CG (~2%) contexts when compared to Col-0. To our surprise, *TrAP* transgenic plants only exhibited decrease in methylation of CHH (~11%) but not CHG (Figure 29A).

Figure 28. ChIP-qPCR analyses of H3K4me3, H3K9me2 and H3K27me3 in TrAP-regulated loci *in vivo*.

(A) TrAP-activated transposons in heterochromatic regions contained reduced H3K9me2 and H3K27me3 but did not show consistent variation in H3K4me3 marks. (B) TrAP-upregulated flowering genes showed consistently reduced H3K9me2 and H3K27me3 marks compared to wild-type Col-0. (C) TrAP-downregulated genes displayed variable changes of H3K9me2 and H3K27me3 marks and no obvious changes of H3K4me3 mark. (D) Tubulin (TUB8) was used as internal control for all the ChIP experiments; the percentage enrichment versus input is shown. ChIP assays were conducted on 11-day old seedling using antibodies specific for H3K9me2 (Abcam Cat# ab1220), H3K27me3 (Millipore, cat #,07-449), and H3K4me3 (Millipore, cat #04-745). Enrichment of H3K4me3 and H3K27me3 in each locus is normalized to that of TUB8; H3K9me2 enrichment is plotted as percentage of input. The standard deviation (SD) was calculated from at least three biological repeats.


□H3K4me3 ■H3K9me2 ▒H3K27me3
A
Ta3
AT1TE45390
IGN5
Chr5 AtSN1
AT1TE51360
Not Detected in H3K4me3 ChIP
Not Detected in H3K27me3 ChIP
H3K4me3 signal (relative to TUB8)
H3K27me3 signal (relative to TUB8)
H3K9me2 signal (IP/Input %)
Col-0
ddm1
lhp1
#1
2
35S-TrAP
kyp
B
CUC1
FTM1
FT
AGL9
MFT
C
PRE1
AT4G17810
D
TUB8
Enrichment (IP/Input %)

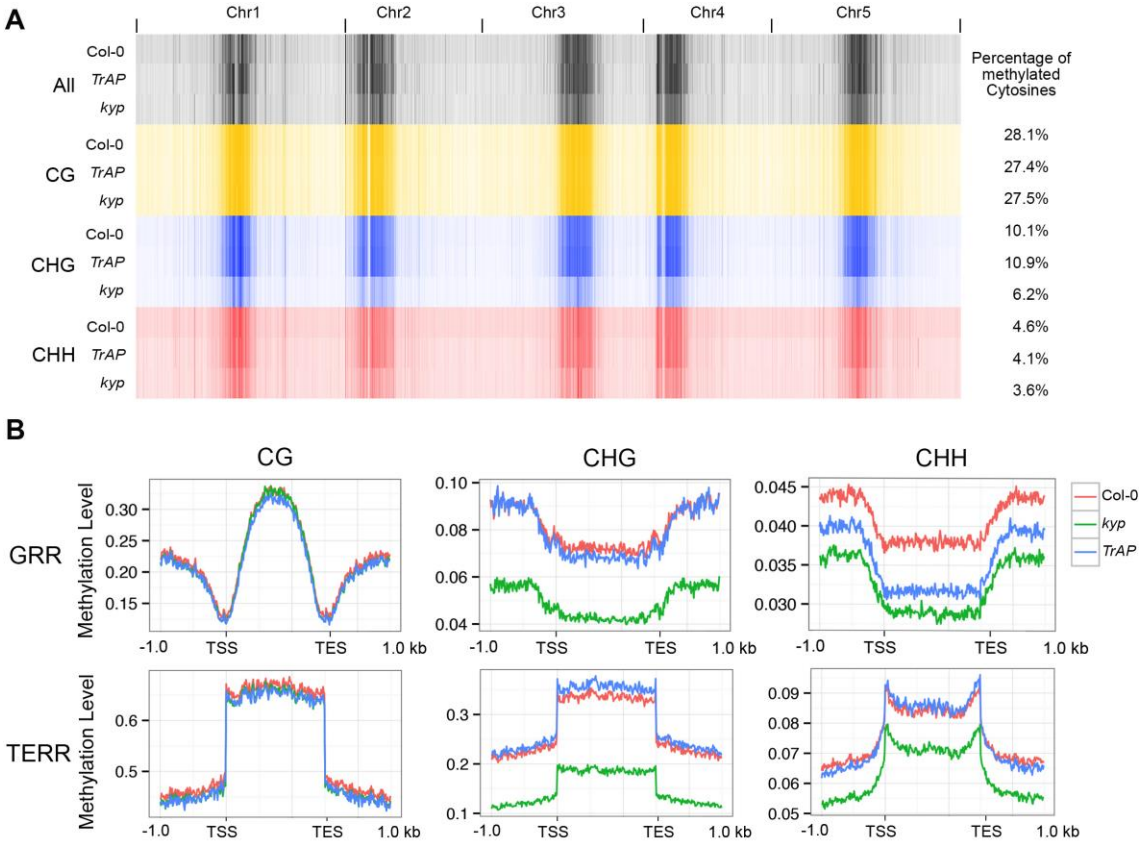


Figure 29. TrAP reduces CHH DNA methylation *in vivo*.
(A) Genome-wide heat map of DNA methylation levels in Col-0, kyp mutant, and TrAP transgenic plants (Left). Sequence context of all cytosine, CG, CHG and CHH methylation was depicted as black, yellow, blue, and red, respectively. The percentage of cytosine methylation is shown for each genotype (Right). (B) Average distribution of context specific DNA methylation in Gene- and TE- Rich Regions in Col-0 (red), *kyp* mutants (green) and TrAP transgenic plants (blue). (C). Overlap of CHH DMRs between TrAP transgenic plants and *kyp* mutant.

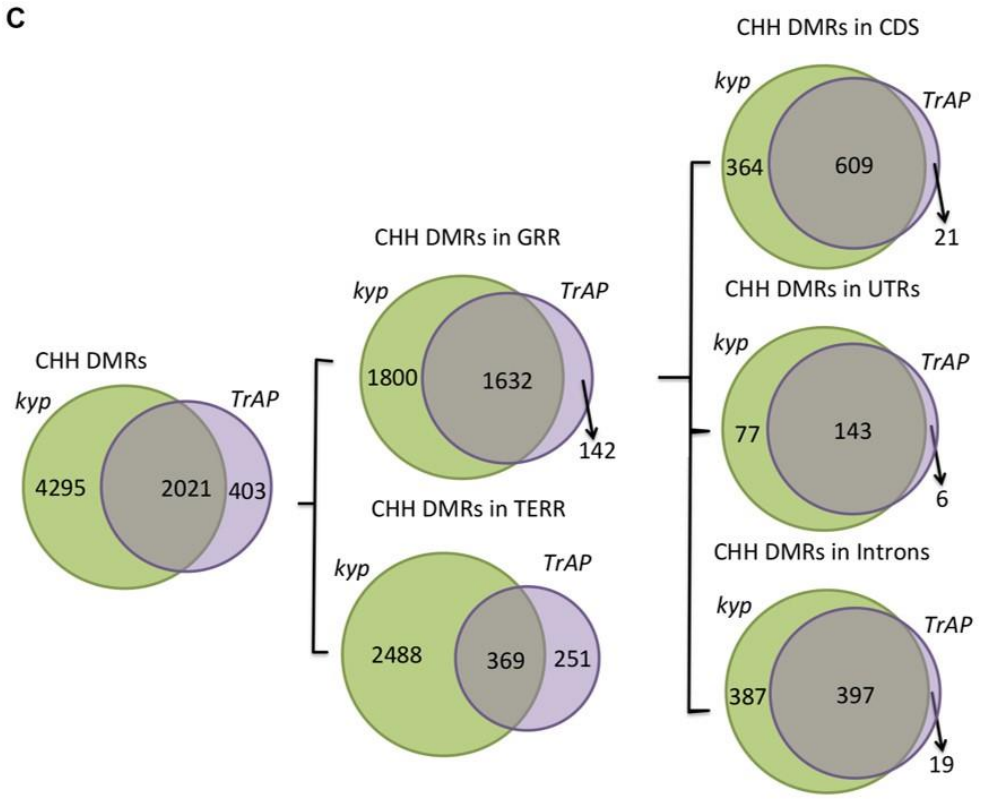


Figure 29 Continued.

To further analyze the effect of TrAP on DNA methylation, we identified differentially methylated regions (DMRs) by scanning the genomes in 200 bp tiles and comparing the level of methylation among *kyp* and *TrAP* overexpression plants with Col-0 (See Section 2.5 "Materials and Methods") (Supplementary files 5, 6, 7). Given that non-CG methylation is highly co-localized and predominant in TE-rich heterochromatic regions (Dubin et al., 2015; Shen et al., 2014a; Stroud et al., 2013, 2014; Yang et al., 2015), we separated the DMRs into gene- and TE-rich regions (GRR and TERR respectively); When compared to Col-0, *TrAP* transgenic plants displayed loss of CHH

methylation at GRR but not in TERR, nor in CG or CHG contexts (Figure 29B). Remarkably, of the 3442 and 1784 GRR hypomethylated DMRs identified in *kyp* mutant and TrAP transgenic plants, 1642 were shared. To better understand the effect of *TrAP* expression on CHH DNA methylation, the GRR DRMs were further separated into promoter, terminator, UTR, intronic, and coding regions. We found that *TrAP* hypomethylated DMRs in coding sequences, UTRs, and introns overlapped almost completely with the *kyp* mutant (96.7%, 96%, and 95.4% respectively) (Figure 29C). Notably, gene ontology analysis of the overlapped genes pointed to response genes, specifically in protein kinase categories (Figure 30; Supplementary file 7). This substantial overlap explains about one-third of the total CHH DMRs in *kyp* mutant, and one half of *kyp* DMRs in genic regions (Figure 29C). Together, these results suggest an inhibitory effect of TrAP in KYP-dependent CHH DNA methylation.

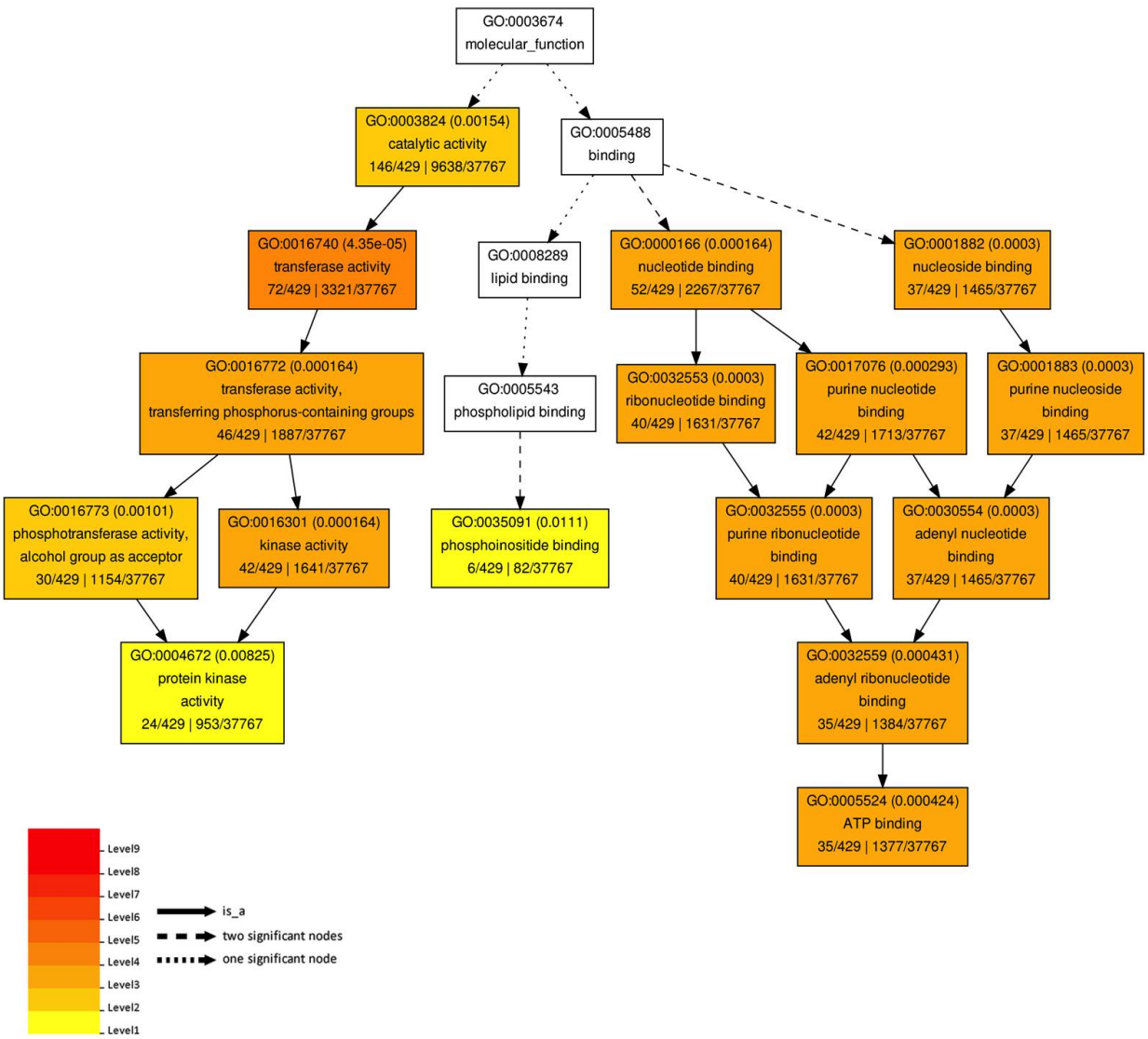


Figure 30. Gene ontology of CHH hypomethylated genes in *TrAP* transgenic plants and *kyp* mutant.

The genes associated to the CHH hypomethylated DMRs in both *TrAP* transgenic and *kyp* mutant plants underwent Gene Ontology analysis using AgriGo tool with TAIR10 as reference

2.3.8 TrAP inhibits KYP activity to counter host defense

In animals, H3K9 methylation promotes TGS and latency of integrated viruses. In plants, Geminivirus constitutes into a minichromosome that also undergoes epigenetic regulation. The specific TrAP-KYP interaction, inhibition of KYP activity *in vitro*, and reduction of H3K9me2 and CHH methylation *in vivo* suggest that KYP is a major factor in combating viruses. Previous studies showed that *kyp* mutants are hypersusceptible to Geminivirus infection (Hanley-Bowdoin et al., 2013); these experiments were reproducible in our hands (Figure 31A-D). Moreover, the hypersusceptibility of *kyp* mutant to the virus could be rescued by the wild-type *KYP* transgene under the control of both the native and the constitutive 35S promoters (Figure 31A-D); further validating a role for KYP in regulation of viral infection. In line with the phenotypic complementation, wild-type *KYP* almost completely rescued H3K9me2 defects of the TEs in the *kyp* mutant (Figure 31E, lanes 1–4). Interestingly, although the accumulation of *KYP* transcripts was substantially increased when transcribed from the *35S* promoter when compared to the native promoter, the steady-state protein level was only twofold to threefold higher in the *35S-Flag-4Myc-KYP* than in the P_{KYP}*-Flag-4Myc-KYP* transgenic plants (data not shown). These observations suggest a possible homoeostatic regulation of this critical TGS component.

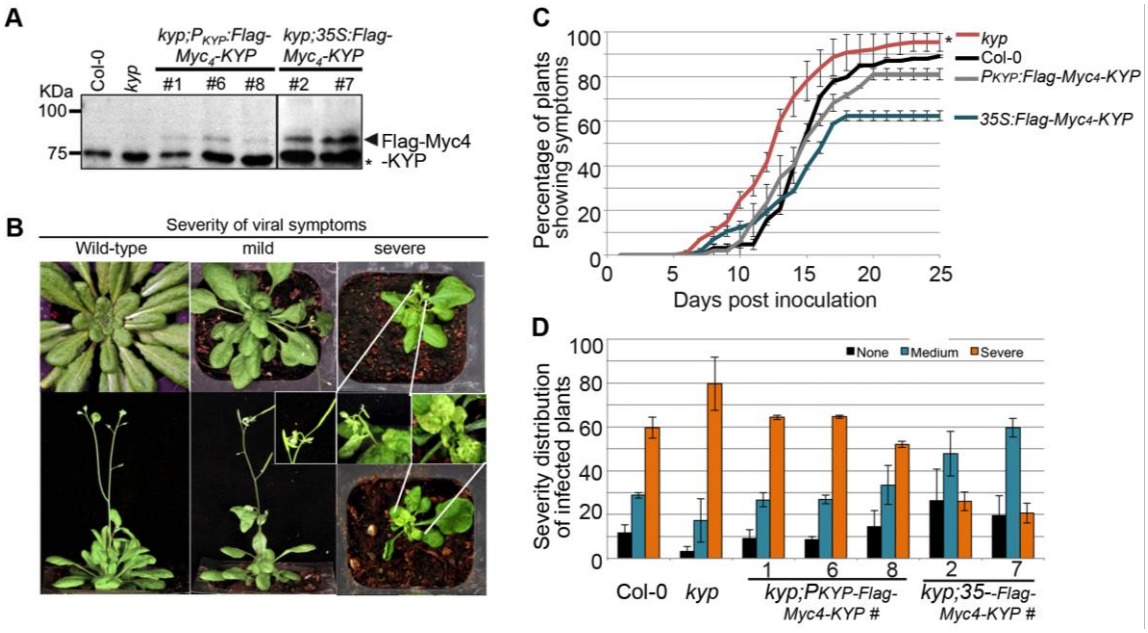


Figure 31. KYP methylates geminivirus chromatin as a host defense.
(A). Western blot analysis of kyp complementation lines expressing PKYP- or 35S-Flag-4Myc-KYP using anti-myc antibody. *, a cross-reaction band serves as a loading control. (B) Representative CaLCuV symptoms with different severities. (C) Time course of CaLCuV symptom development in kyp mutant and the complementation lines. The mean values were calculated with SD from at least three experiments (>30 plants/line). (D). CaLCuV symptom severity in Col-0, kyp mutant, and the complementation lines. The mean values were calculated with SD from at least three experiments (>30 plants/line). (E) ChIP-PCR of H3K9me2 marks in heterochromatic loci in kyp mutant and complementation lines inoculated with mock or CaLCuV. Note: CaLCuV infection largely removed H3K9me2 marks from heterochromatic loci. (F) Schematic linearized representation of the regions of viral genome A selected for ChIP assays. (G) ChIP-qPCR assays showed KYP-dependent enrichment of H3K9me2 in several tested loci in the viral chromatin. The relative value of histone methylation in each sample was normalized to that of wild-type control where the signal was arbitrarily assigned a value of 1 with standard deviation (SD) from at least three biological repeats. Note: the region defined by # Primer 7 serves as a negative control. (H) Western blot analysis of to detect FM-KYP in the ChIP (IP) samples using anti-myc antibody. (I) ChIP-PCR assays showed that KYP binds to the viral minichromosome. The ChIP assays were done with a monoclonal anti-Flag antibody.

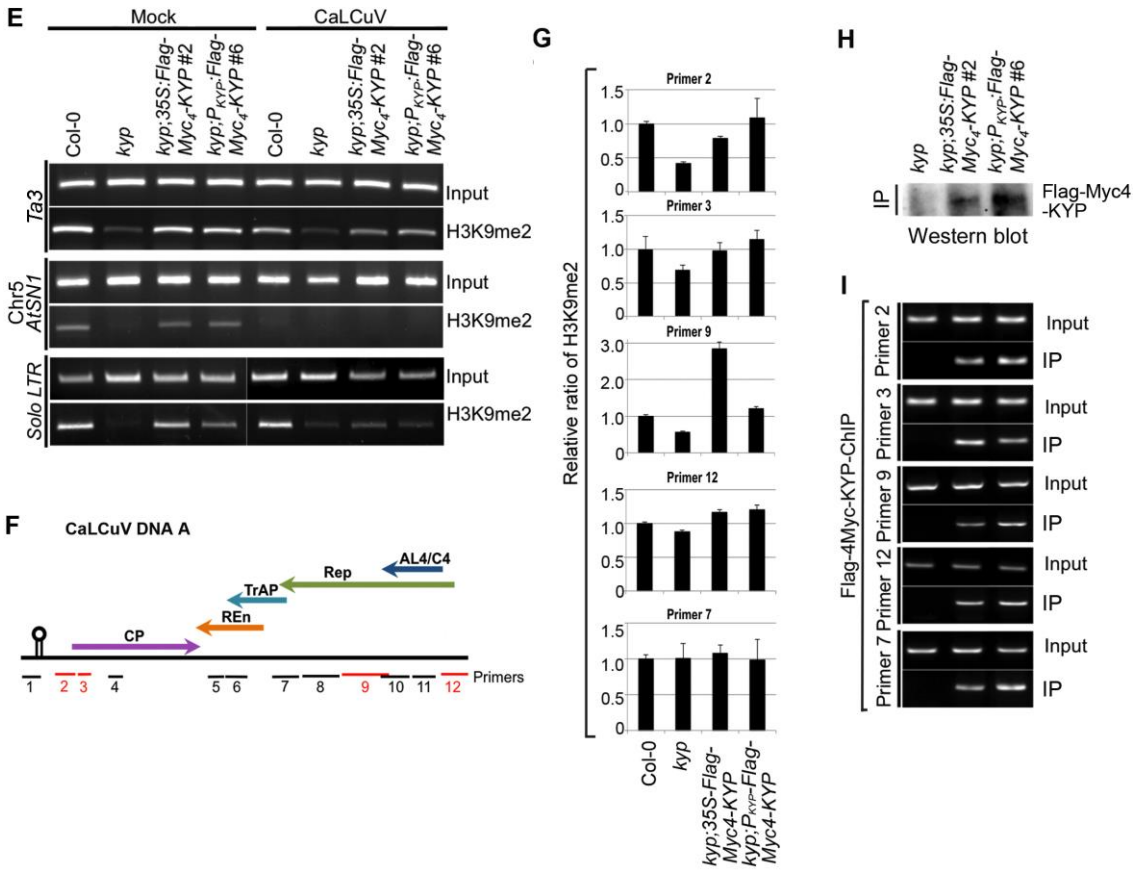


Figure 31 Continued.

Geminivirus minichromosome harbors H3K9me2 (Figure 31G, Figure 32) (Hanley-Bowdoin et al., 2013). Interestingly, H3K9me2 marks were dramatically decreased on several, but not all tested loci in the viral genome in *kyp* mutant, and the methylation marks were further restored by the *KYP* transgene, suggesting that viral chromatin H3K9me2 is catalyzed by KYP. Notably, the H3K9me2 amount was substantially enhanced in the locus defined by Primer #9 in the transgenic *KYP* overexpression plants (Figure 31G). It is noteworthy that this locus

harbors the promoter element for *TrAP* itself (Shung and Sunter, 2009). This result suggests that the tight control of *TrAP* expression likely determines the balance between the host and virus interactions. If so, the result might also explain the relatively milder viral symptoms of *35S-FM-KYP* lines compared to WT control or P_{KYP}*-FM-KYP* complementation lines (Figure 31A-D). To further test whether KYP methylated the viral chromatin, we performed ChIP assays using monoclonal anti-Flag antibody to pull down *FM-KYP* bound chromatin (Figure 31H). Excitingly, KYP was found at all tested loci in the viral chromatin (Figure 31I). Notably, regions delineated by the primer #7 showed KYP-independent H3K9me2 marks but were still immunoprecipitated in the KYP-chromatin complex. A possible explanation is incomplete chromatin shearing of the small viral mini-chromosome under the conditions used in this experiment, which were standardized for host chromatin ChIP. Alternatively, this might suggest the presence of additional epigenetic regulation that masks KYP activity on this locus. This notwithstanding, our results indicate that KYP directly deposits the H3K9me2 mark on the Geminivirus minichromosome to reinforce the silent status of the virus.

Since constitutive expression of *TrAP* reduced H3K9me2 *in vivo*, we wondered if Geminivirus infection could also decrease the repressive marks in the host. To this end, we tested the KYP-controlled endogenous transposons for H3K9me2 accumulation. CaLCuV-infected plants showed 60–100% H3K9me2 loss in the tested loci compared to the amount in the mock-inoculated plants (Figure 31E), this is reminiscent of the molecular phenotype of *TrAP* transgenic plants, which showed lower enrichment of H3K9me2 mark at the studied loci (Figure 28C). These results indicated that Geminivirus

infection largely removed the repressive H3K9me2 marks in the transposons, and that the removal resulted at least in part from TrAP function. Given that TrAP is known to activate the expression of viral genes in the minichromosome and that endogenous transposons can serve as a proxy for viral genomes, it is conceivable that virus-encoded TrAP acts to suppress KYP in order to prevent the deposition of H3K9me2-repressive marks in the epigenome, to activate the expression of viral genes.

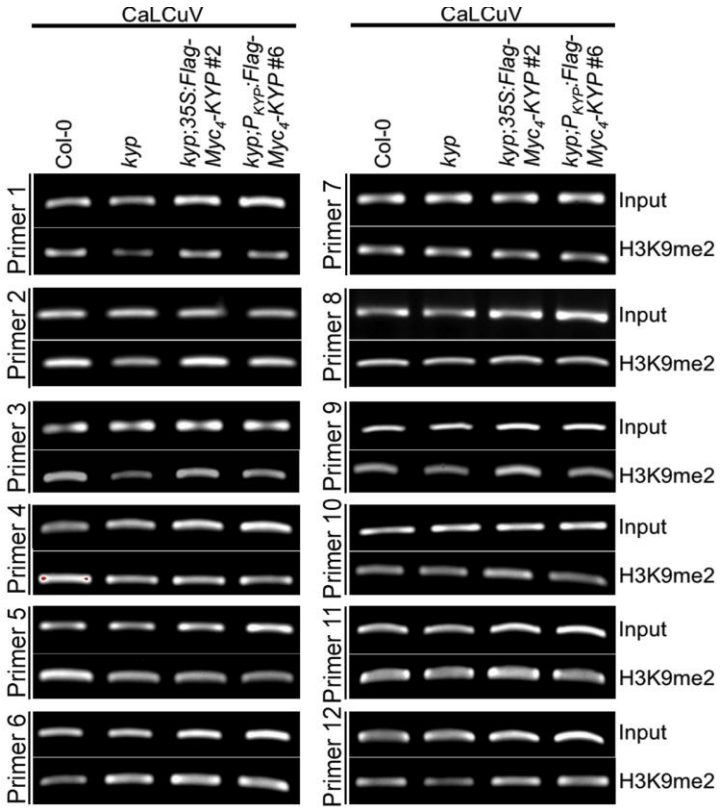


Figure 32. Virus chromatin contains H3K9me2 marks.
ChIP-PCR assays of H3K9me2 marks on the CaLCuV DNA A. ChIP assays were conducted on 9-day old seedlings using antibodies specific for H3K9me2 (Abcam Cat# ab1220), H3K27me3 (Millipore Cat# 07-449) and H3K4me3 (Millipore Cat# 04-745). The PCRs were done with 22 cycles for the input samples and with 28 cycles after ChIP.

2.3.9 kyp mutants sustain systemic infection of CaLCuV lacking a functional TrAP gene

Previous studies show that TrAP is required for the accumulation of the virus infective form, single-stranded (ss) DNA (Hayes and Buck, 1989). TrAP is indispensable for systemic infection of Begomoviruses because it activates the expression of the viral ssDNA binding proteins (nuclear shuttle protein and coat protein), which are essential for releasing the virus from the nucleus and for cell-to-cell spreading (Sunter and Bisaro, 1992). If TrAP-mediated transcriptional activation and accumulation of ssDNA result from inhibition of KYP and correspondingly heterochromatin formation, then infectivity of CaLCuV lacking TrAP should be impaired in wild-type but recovered in *kyp* plants. To test this hypothesis, we engineered a CaLCuV variant without functional TrAP protein (CaLCuV Δ*trap*) by changing a single nucleotide (T to A) that produces an amber mutation and introduces an XbaI restriction site after the sixth codon in the *TrAP* gene (Figure 33A). Then, we assessed systemic infection of CaLCuV and CaLCuV Δ*trap* on wild type and *kyp* plants (Figure 33B).

Consistent with the hypersusceptible phenotype of *kyp* mutants to CaLCuV (Figure 31C,D), these plants accumulated significantly higher titers of CaLCuV relative to wild-type control as both the replicative intermediate, open circle (OC), and the infective particle, SS DNA (Figure 33C). Consistent with previous reports, the wild-type plants did not show any symptoms of infection (Figure 33E) or systemic accumulation of CaLCuV Δ*trap*, as evidenced by Southern blot analyses and PCR (Figure 33C).

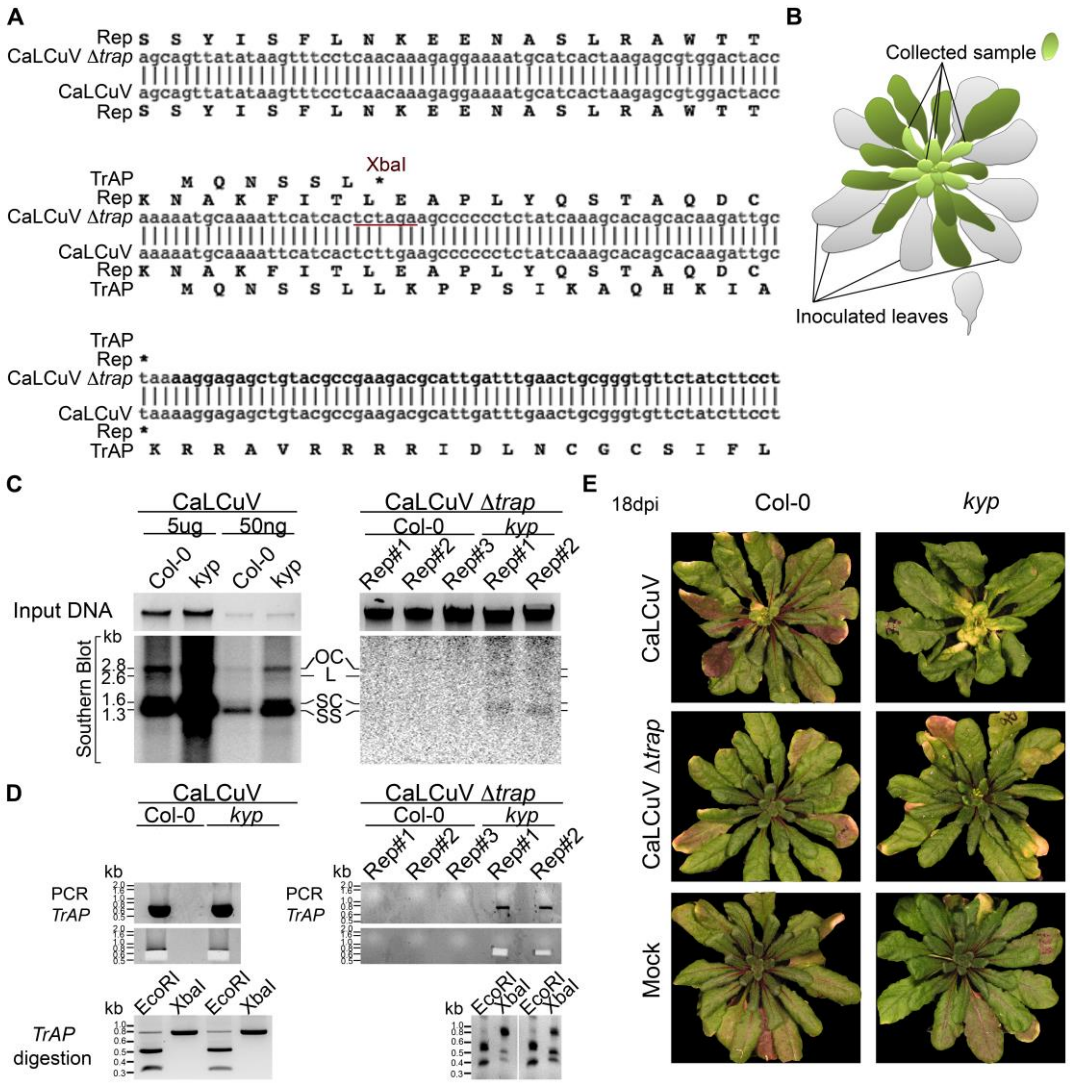


Figure 33. Infectivity of CaLCuV lacking functional TrAP protein.
(A) Sequence alignment of CaLCuV Δ*trap* and CaLCuV sequences. The translated amino acids are shown for each sequence, and the XbaI restriction site resulting from the T to A point mutation is highlighted. (B) Schematics of the systemic infection experiment. Plants with eight true leaves (depicted in grey) were inoculated with the begomovirus, and eighteen days post inoculation nine to 11 newly emerged, not inoculated, rosseta leaves (depicted in bright green) were collected to test for virus systemic infection. (C) Southern blot analysis of viruses in non-inoculated leaves of infected plants. Ethidium bromide staining of total genomic DNA serves as a loading control (top panels). Southern blots were probed against CaLCuV DNA A common region (CR) (bottom panels); the viral populations are indicated as the replicative intermediate open circle (OC), linear (L), super coiled (SC) and the infective particle ssDNA. (D) Genotypic confirmation of the systemically amplified viruses. Top panels show PCR amplification of a *TrAP*-containing region; bottom panels show EcoRI and XbaI digestions of PCR products to examine the presence of the amber mutation in the gene. (E) Exemplary phenotypes of wild-type and *kyp* mutant plants inoculated with mock, CaLCuV and CaLCuV Δ*trap*.

Particularly, PCR amplification followed by XbaI digestion of the viral *TrAP* region in wild-type plants infected with CaLCuV Δ*trap* showed loss of the XbaI restriction site (Figure 34). Excitingly, the viral DNA was detected, in low amount but reproducibly, in *kyp* mutants. Further PCR amplification of the *TrAP* gene followed by XbaI digestion, confirmed that the detected viral DNAs in *kyp* mutants indeed came from CaLCuV Δ*trap* and not from a reversion of the mutation in the virus genome. These results show that *kyp* plants sustained systemic infection of CaLCuVΔ*trap*. Noticeably, the titers of accumulated CaLCuV Δ*trap* were much lower than the ones of CaLCuV, suggesting either redundant activity of *KYP* paralogs in host defense or necessity of additional functions of TrAP besides inhibition of KYP activity for the virus to achieve efficient infection.

2.4 Discussion

TrAP was among the first viral suppressors identified to interfere with the TGS pathway. The prevailing model is that TrAP lowers the reservoir of transferable methyl groups by targeting ADK, a key enzyme in the SAM pathway. Here, we propose a novel model in which TrAP regulates TGS by directly targeting KYP (Figure 35). Several pieces of evidence supported our notion: (1) TrAP genetically interfered with the TGS pathway (Figure 19); (2) TrAP directly interacted with KYP *in vivo* and also with other HMTases *in vitro* (Figure 20, Figure 23); (3) TrAP inhibited the catalytic activity of KYP *in vitro* (Figure 25); (4) TrAP reduced the repressive H3K9me2 marks *in vivo* (Figure 28), and correspondingly, reactivated numerous loci that are otherwise repressed by KYP (Figure 19); (5) TrAP decreased CHH methylation in gene-rich regions that are also regulated by

KYP (Figure 29); (6) methylation of viral chromatin entailed KYP (Figure 31); (7) KYP bound the viral chromatin (Figure 31); and (8) *kyp* mutants but not wild-type plants sustain low systemic infection of CaLCuV lacking TrAP protein (Figure 33). To our best understanding, this is the first evidence that a viral protein directly suppresses HMTases in the host TGS machinery.

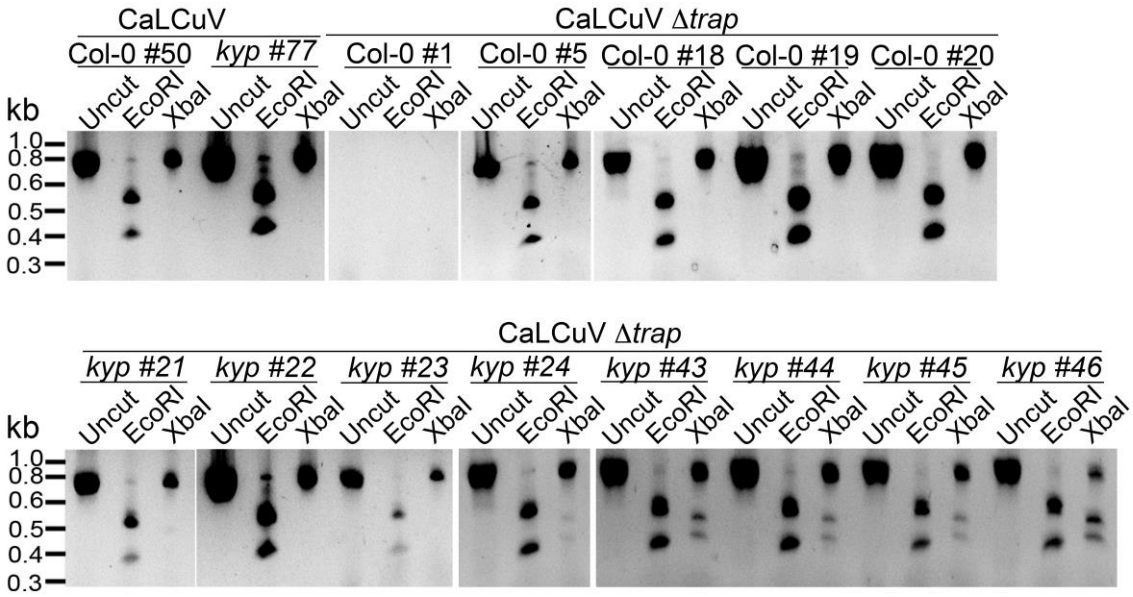


Figure 34. CaLCuV lacking functional TrAP protein cannot cause systemic infection in wild-type plants.
The region corresponding to TrAP gene was amplified by PCR using the primers AL3_cterm_F and AL1_cterm_R (Supplemental Table 8) from total DNA extracted from plants infected with CaLCuV or CaLCuV Δ*trap*. 5μL of the PCR products were run and 3μL were used for digestion with EcoRI or XbaI to examine the presence of the amber mutation in all the amplified products.

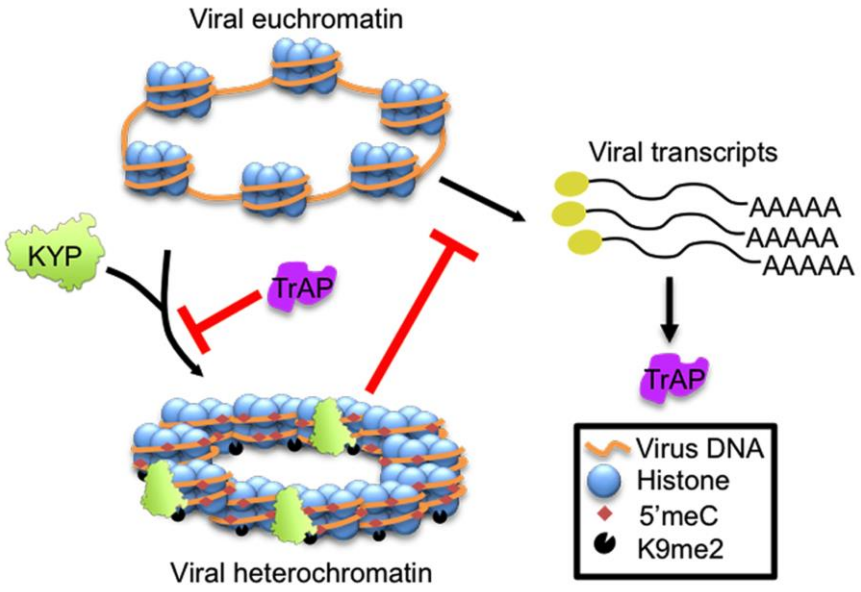


Figure 35. Model of TrAP suppression of KYP activity to prevent epigenetic silencing of the viral chromatin.
Geminivirus genome is packed on histone octamers to form a minichromosome. The minichromosome undergoes extensive H3K9me2 modification deposited by host-encoded KYP, and this modification could be further reinforced by DNA methylation, leading to formation of viral heterochromatin. As a counter-defense strategy, Geminivirus-encoded TrAP protein inhibits KYP activity to maintain the euchromatic status of the minichromosome to permit active replication and transcription of viral genes, and correspondingly to escape host surveillance.

Given that TrAP protein could concurrently limit the upstream supply of the methyl groups and directly inhibit downstream enzymatic activity of KYP, Geminiviruses appear to have evolved sophisticated strategies to cripple the host TGS pathway. What would be the biological advantages of blocking the TGS pathways? In eukaryotes, chromatin appears to be a critical battleground for virus–host interaction. Animals use histone modifications to reinforce the latency of integrated viruses (du Chéné et al., 2007; Narasipura et al., 2014; Park et al., 2014). Plant DNA viruses including Geminiviruses and pararetroviruses replicate as nuclear minichromosomes or episomes. Clearly, TGS

functions as an immune system to control virus replication and the expression of viral genes, in a similar fashion as the repression of endogenous TEs and transposon remnants. In previous studies and here in our experiments, *kyp* mutants displayed hypersusceptibility to Geminivirus infection, which could be rescued by exogenous wild-type *KYP* gene. KYP directly acts on viral chromatin and deposits the repressive H3K9me2 mark on the viral chromatin. Moreover, viruses lacking the essential TrAP protein can, although inefficiently, cause systemic infection in *KYP*-deficient hosts. All these facts point out the unambiguous role of KYP-mediated TGS in defense against viral infection. On the other hand, TrAP functions to inhibit KYP catalytic activity, reducing the repressive H3K9me2 mark, to activate transcription of viral genes. Our KYP reconstitution assays show that TrAP:KYP molar ratio of around 2 is enough to cause ~50% inhibition of KYP activity, indicating that TrAP is a potent inhibitor of HMTase activity. Thus, direct inhibition of KYP represents a novel counter-defense mechanism for virus survival in the hosts. This mechanism could account for the long-documented essential role of TrAP in expression of viral genes including the coat protein and the nuclear shuttle protein (Hanley-Bowdoin et al., 2013; Shen et al., 2009; Sunter and Bisaro, 1992; Yang et al., 2007).

In plants, H3K9me2 marks often correlate with non-CG DNA methylation, and in particular with CHG modification. The tight coordination results from a self-enforcing loop consisting of KYP and CMT3 (Du et al., 2014). Briefly, KYP methylates H3K9 to generate the binding sites for CMT3, which further methylates CHG DNA to create more binding sites for KYP. Consequently, the crosstalk between DNA and histone methylation ascertains the silent status of heterochromatin. We hypothesized that TrAP targets KYP

to reduce H3K9me2 marks, and this inhibition would further decrease DNA methylation both on the host and on the viral genome. Here, we found that ectopic expression of TrAP had no effect on CHG but only on CHH methylation, at sites that are up to 96.7% overlapped with the hypo-methylated regions in *kyp* mutant (Figure 29). In plants, sequence contexts of CHG and CHH methylation are largely overlapped throughout the genome and are maintained to a limited extent by all non-CG methyltransferases (Stroud et al., 2014). One outstanding question would be how could TrAP differentially alter methylation of CHH rather than CHG, given that TrAP targets KYP? Methylation in the CHH context is catalyzed by CMT2 and DRM1/2 in *Arabidopsis*, and the sites regulated through the two sets of enzymes are mostly non-overlapping. Thus, the functions of CMT2 and DRM1/2 are mostly non-redundant at CHH sites (Stroud et al., 2013, 2014). DRM1/2 largely catalyze CHH methylation at the TE-rich regions (Stroud et al., 2014). In our study, we did not observe any effect of TrAP on CHH methylation in the TEs, suggesting that TrAP might not (or not sufficiently) interrupt the RdDM pathway that entails DRM1/2 and 24-nt siRNAs (Stroud et al., 2014). Importantly, it has been recently discovered that bulk CHH methylation is maintained by CMT2 (Dubin et al., 2015; Shen et al., 2014b; Stroud et al., 2014) and that its activity is largely dependent on H3K9me2. In this scenario, CMT2 recognizes methylated H3K9 but preferentially binds to di-methylated over mono-methylated histone tails. This preference towards H3K9me2 is not observed in the CHG methyltransferase CMT3, which can equally bind to all forms of H3K9 methylation (Stroud et al., 2014). We might envision that methylation at CHH sites could be more sensitive to changes in H3K9me2 than CHG, since CMT3 could still maintain CHG DNA

methylation in the presence of H3K9me1 resulting from residual activity of HMTase, as would be the case when KYP is inhibited by TrAP. Alternatively, CHH methylation might play an important but yet unappreciated regulatory role in host defense genes (Figure 30). If so, preferentially targeting these loci by TrAP protein might represent a new counter-defense mechanism.

In the host, Geminivirus DNAs also undergo extensive methylation modification (Raja et al., 2008). In our study, TrAP reduces CHH methylation in the host and possibly in the viral genome, too. Remarkably, the Geminivirus replicase, AC1/C1 and the embedded protein AC4/C4, downregulate the expression of host DNA methyltransferases *MET1* and *CMT3* (Pumplin and Voinnet, 2013; Rodríguez-Negrete et al., 2013); hence, it interrupts the reinforcing loop of histone and DNA methylation. Consistent to this study, *cmt3* mutants exhibit hypersusceptibility to viral infection (Raja et al., 2008). The fact that Rep/C4 represses *CMT3* expression in the host is in perfect alignment with our ChIP assay results on CaLCuV-infected plants, where the loss of H3K9me2 is even more severe than the observed by TrAP transgenic plants. Thus, interruption of the compelling feedback loop of histone and DNA methylation represents an important strategy to sustain transcription of viral chromatin. Together, it seems that the synergistic inhibition of histone and DNA methyltransferases by Geminivirus proteins evolves as a powerful tactic to win the arms race between host and pathogen.

Our *in vitro* assays clearly demonstrated that TrAP predominantly binds to the catalytic domain of KYP and inhibited its enzymatic activity. Whether TrAP might alter KYP conformation or block the accessibility of substrates to the active sites upon

interaction awaits future structural analysis. We note, however, that *Arabidopsis* has 49 SET domain containing proteins of which 31 are considered to have HMTase activity (Liu et al., 2010). Arising from this fact is whether TrAP specifically targets KYP or promiscuously acts on additional HMTases. Although TrAP overexpression plants display molecular features in common with *kyp* mutant (i.e., reduced H3K9me2 levels), and an early flowering phenotype similar to loss-of-function mutant of the *kyp* paralog, *SUVH2*, the transgenic plants are phenotypically different from *kyp* single mutant. This could be due to functional redundancy between *KYP* and its paralogs *in vivo*; alas, the morphological phenotype of the higher-order mutants has not been fully documented.

We also noticed that *TrAP*-overexpression plants are morphologically similar to *lhp1*, which is characteristic of the H3K27me3 pathway. Moreover, the substantial overlapping of *TrAP*-responsive genes with *lhp1*-regulated genes strongly suggested that TrAP might target the*LHP1*-related H3K27me3 pathway (Zheng and Chen, 2011). Indeed, most of our tested loci in the host genome exhibited decreased H3K27me3 levels, consistent with the fact that ~50% genes were co-regulated by TrAP and *LHP1* in a genome-wide scale. Interestingly, recent ChIP–chip studies revealed that H3K27me3 and *LHP1*-bound sites are predominantly distributed in the euchromatic regions (Turck et al., 2007; Zhang et al., 2007a, 2007b). This distribution is not correlated with KYP-dependent H3K9me2 marks that are highly enriched at pericentromeric regions as large and uninterrupted heterochromatic blocks (Black et al., 2012; Du et al., 2014; Liu et al., 2010, 2014d). H3K9me2 can also occur in euchromatic regions but rather exist as small heterochromatin patches (Zheng and Chen, 2011). In our current study, we did not further

examine whether TrAP physically targets H3K27me3 HMTase in the *LHP1* pathway. But it is plausible that TrAP inhibits both KYP and *LHP1* pathways. Notably, we did not observe the repression of transcriptionally active histone methylation marks, such as H3K4me3 and H3K36me3. This observation suggests that TrAP does not target the HMTases that are writers for the active marks. How can TrAP distinguish KYP from those active writers remains unclear, but we propose that structural components other than the SET domain might contribute to the recognition and affinity of TrAP binding to HMTases (Figure 23). Alternatively, additional cellular factors might also contribute to the specificity.

The direct consequence of TrAP-dependent inhibition of KYP activity is to activate viral transcription and replication. Because KYP is a key effector of TGS in the host, and regulates a broad array of endogenous genes, interference with this core component would reprogram the expression profile of the host genome and thus trigger a series of downstream cascade signaling events that impact the balance of host/virus interaction. As an example, TrAP suppresses auxin and cell growth, whether this change might constitute a defense mechanism for the host benefit or create a favorable cellular niche for virus propagation remains for further investigation.

In conclusion, our results support the notion that Geminivirus-encoded TrAP protein interferes with the TGS pathway and abrogates epigenetic silencing by direct interaction with KYP and inhibition of its transmethylation activity. Thus, Geminivirus TrAP functions clearly different from most of previously characterized viral suppressors, which target various steps of the PTGS pathway (Ding and Voinnet, 2007). Together with

previous studies (Raja et al., 2008), we provide evidence that KYP evolves as a critical immune system to control invading nucleic acids in plants; this is reminiscent of the roles of human SUV39H1 in maintaining the latency of HIV (du Chéné et al., 2007), Epsteinn-Barr virus (Imai et al., 2014), and some other mammalian viruses. Thus, sequestering or interfering with this core component is an effective strategy for Geminivirus to block TGS and to subvert host defense; we expect this strategy to be used by other suppressors in plants and mammals. Given that many viral suppressors interrupt different steps in the PTGS pathways to efficiently combat host surveillance, it would not be surprising that additional key cellular factors in the TGS pathway, besides HMTase, were readily targeted by invading viruses in eukaryotes. Importantly, it is believed that compounds or drugs that alter chromatin methylation might ultimately be the most effective means of combating disease. Our discovery that TrAP inhibits a histone-modifying enzyme also offers a new natural strategy to develop epigenetic-targeted drugs to cure human diseases that arise from epigenetic dysfunction (Højfeldt et al., 2013) or to engineer new biotechnological products to improve agricultural productivity.

2.5 Materials and methods

2.5.1 DNA constructions

All the plant constructs were made using the Gateway system (Invitrogen) (Zhang et al., 2005). The destination vectors (containing the destination cassette–DC-) pHyg-DC-CFP, pBA-DC-CFP, pBA-DC-3HA, pBA-DC, pBA-Flag-4Myc-DC, and pER10-YFP-DC (Zhang et al., 2006a) were used for transient expression in *N. benthamiana* and for stable *Arabidopsis* transformation. The vector pER10 for beta-estradiol-inducible

expression under the XVE promoter (Zuo et al., 2000) was modified to obtain the destination vector pER10cLUC-3HA-DC. For this, the pER10 vector was linearized with XhoI, filled in with Klenow fragment; and further digested with PacI. In parallel, the cLuc-3HA-DC insert was obtained by linearizing the pCambia1300-cLuc-3HA-DC (Zhang et al., 2011c) plasmid treated with SacI/Klenow treatment and further digested with PacI. The vector and insert were ligated and transformed into *E. coli* DB3.1 and finally confirmed by sequencing. This vector was used together with pCambia Myc-DC-nLUC (Zhang et al., 2011c) for transient expression in *N. benthamiana*.

The cDNA or DNA fragments were cloned into pENTR/D vectors, sequencing confirmed; and then transferred to the appropriate destination vectors by recombination using the LR Clonase (Invitrogen). The primers for the cloning are listed in the Supplementary file 8. To drive KYP expression from its native promoter, we amplified a 2.7-Kb genomic region immediately upstream of the *KYP* start codon with the primers P_{KYP} EcoRV for and P_{KYP}BamHI rev. The binary vector pBA002a Flag-4Myc-KYP was obtained by the Gateway system using LR clonase. The resultant plasmid and PCR product harboring *KYP* promoter were digested with EcoRV/BamHI and ligated with T4 DNA ligase. The plasmid was confirmed by sequencing using the primer P_{KYP} seq For.

For expression of recombinant proteins, the cDNA or DNA fragments were cloned into pMCSG9 or pMCSG10 vectors to produce His-MBP- or His-GST-tagged proteins respectively, by ligation independent cloning (Eschenfeldt et al., 2009) using primers that include 18nt identical to the ends of the linearized pMCSG vector (Supplementary file 8). The pMCSG plasmids were linearized with the blunt end restriction enzyme SspI, and the

sticky ends were generated by T4 DNA Polymerase (NEB) supplemented with 2.5 mM dGTP. In parallel, the complementary sticky ends in the PCR products were generated supplementing the T4 DNA Polymerase (NEB) with 2.5 mM dCTP. The mixture of treated plasmid:PCR (3:4, respectively) was incubated on ice for 30 min to allow annealing of the complementary free strands and transformed in *E. coli* DH5α. The plasmids were further confirmed by sequencing and transformed into *E. coli* BL21 Rossetta DE3 for expression.

To engineer the CaLCuV Δ*trap* DNA A infective clone the pNSB1090 plasmid was subjected to site directed mutagenesis by amplification with the primers CaLCuV_AL2_null_XbaI_for and CaLCuV_AL2_null_XbaI_rev (Supplementary file 8), digested overnight with DpnI, cleaned with the QIAquick PCR Purification Kit (Qiagen), and amplified in *E. coli* DH5α. The plasmid was confirmed by sequencing and transformed in *Agrobacterium tumefasciens* ABI for the virus infection assays.

2.5.2 Transgenic plants

A. thaliana (Col-0) plants were transformed with binary vectors by the floral-dip method (Clough and Bent, 1998; Zhang et al., 2006b). The transgenic seeds were selected on standard MS medium (Murashige and Skoog, 1962) containing the appropriate selective agents: 10 mg/l glufosinate ammonium (Sigma) or 25 mg/l hygromycin (Sigma), together with 100 mg/l carbenicillin (Sigma). *kyp* mutant (SALK_130630C) was obtained from the Arabidopsis Biological Resource Center and confirmed by genotyping and qRT-PCR.

2.5.3 Expression and purification of recombinant proteins

For *in vitro* pull-down or HMTase reconstitution assays, the recombinant proteins were purified following either one or two-step affinity purification procedure. His-MBP-tagged proteins were first purified through Immobilized Metal ion Affinity Chromatography (IMAC) using the Ni-NTA resin (Qiagen), followed by amylose resin (NEB) according to manufacturers' protocols. His-GST-tagged proteins were initially purified through sepharose glutathione column (GE) followed by IMAC. Specifically for HMTase assays, His-MBP-CaLCuV_TrAP, His-MBP-TGMV_TrAP, and His-MBP were prepared in lysis buffer (50 mM Tris-HCl pH 9, 300 mM NaCl, 10 mM 2-mercaptoethanol, 2 mM PMSF), incubated with the Ni-NTA resin at 4°C for 1 hr, eluted with 300 mM imidazole and immediately incubated with the amylose resin (NEB) at 4°C for 1 hr. The proteins were eluted with 10 mM maltose and the elute was further separated by size exclusion chromatography (SEC) in column buffer (20 mM Tris-HCl pH 9, 100 mM NaCl); the fractions containing the target protein were pulled together, concentrated to 100 µM, aliquoted and stored at −80°C until usage. His-GST-KYP was prepared in PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2HPO_4, pH 7.3, 10 mM 2-mercaptoethanol) incubated with the sepharose glutathione 1 hr at 4°C, eluted with elution buffer (50 mM Tris-HCl pH 9, 200 mM NaCl, 10 mM reduced glutathione, 10 mM 2-mercaptoethanol). The elute was further purified through the Ni-NTA column and finally through SEC in column buffer (20 mM Tris-HCl pH 9, 100 mM NaCl); the fractions containing the target protein were pulled together, concentrated to 25 µM and aliquoted for usage.

For *in vitro* pull down assay, both the prey (His-GST-CaLCuV-TrAP, His-GST-TGMV-TrAP, and His-GST) and the bait (His-MBP-Bait) proteins were purified by IMAC using the lysis buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 2 mM PMSF), incubated with the Ni-NTA resin (at 4°C for 1 hr, eluted with 300 mM imidazole and immediately dialyzed in storage buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM 2-mercaptoethanol, 50% Glycerol) at 4°C overnight.

2.5.4 In vitro pull down and Co-IP assays

In vitro pull-down assays and *in vivo* Co-IP were done as described (Zhang et al., 2005). Briefly, 2.5 µg of 6His-GST-tagged prey proteins were pre-absorbed to 50 µl of the amylose resin (NEB) for 1 hr at 4°C in 1 ml of binding buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.2% glycerol, 0.6% Triton X-100, 0.5 mM 2-mercaptoethanol, 2 mM PMSF). The proteins were recovered by ultracentrifugation at 12,000×*g* for 2 min, transferred to a second tube containing 2.5 µg of the MBP-tagged bait protein, and incubated at room temperature for 2 hr. The protein complexes were harvested by adding 50 µl amylose resin beads, followed by 2 hr incubation at room temperature, and cleaned with six vigorous washes with buffer. The pulled-down proteins were resolved by SDS-PAGE and the preys were detected by western blot using anti-GST antibody.

For Co-IP experiments, *N. benthamiana* leaves were collected 2 days after agroinfiltration, ground in liquid nitrogen and stored at −80°C until use. For the assay, total proteins were extracted from 0.4 g of ground powder in 1.2 ml (3 volumes) of IP buffer (40 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM $MgCl_2$, 2 mM EDTA, 4 mM DTT, 0.5% Triton X-100, 1 mM PMSF, 5% glycerol, 2× Roche Complete EDTA-free protease

inhibitor); then, the soluble proteins were cleared twice by ultracentrifugation at 20,000 × rcf for 15 min at 4°C. The protein complexes were immunoprecipitated with 15 µl Anti-c-Myc-agarose affinity gel (Sigma–Aldrich #A7470) at 4°C for 2 hr, the unspecific-bound proteins were removed by four consecutive washes with the IP buffer with 10 min incubation each at 4°C. The protein complexes were eluted in 200 µl of elution buffer (5 mM EDTA, 200 mM NH_4OH) for 20 min. The supernatant was collected, frozen in liquid nitrogen and dried using the Savant SpeedVac concentrator; finally, the sample was solubilized in 50 µl of 2× SDS-loading buffer for western blot analyses.

2.5.5 Two-step immunoprecipitation for mass spectrometry analysis

9-day-old wild-type control and *Arabidopsis* transgenic plants expressing *XVE-FM-TrAP* were induced for 16 hr with 25 µM ß-estradiol in liquid MS media, ground in liquid nitrogen and stored at −80°C until use. For the assay, total proteins were extracted from 10 g of ground powder in 40 ml (4 volumes) of IP buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 4 mM $MgCl_2$, 50 µM $ZnCl_2$, 0.1% Triton X-100, 1 mM PMSF, 1% glycerol, 3× Roche Complete EDTA-free protease inhibitor, 15 µM MG132); then, the soluble proteins were cleared twice by ultracentrifugation at 20,000 × rcf for 15 min at 4°C. The protein complexes were first immunoprecipitated using 500 µl of Anti-FLAG M2 magnetic beads (Sigma, Cat# M8823) and incubated in slow rotation for 2 hr at 4°C, the nonspecific-bound proteins were removed by three consecutive washes with 15 ml of IP buffer for 10 min incubation each at 4°C. The protein complexes were then eluted by competition with 100 mg/ml FLAG peptide and subsequently immunoprecipitated with 100 µl Anti-c-Myc-agarose affinity gel (Sigma–Aldrich #A7470) at 4°C for 1.5 hr, the

nonspecific-bound proteins were removed by five consecutive washes with the IP buffer with 5 min incubation each at 4°C. The protein complexes were eluted in 200 µl of elution buffer (5 mM EDTA, 200 mM NH_4OH) for 20 min. The supernatant was collected, frozen in liquid nitrogen and dried using the Savant SpeedVac concentrator; finally, the sample was solubilized in 30 µl of 2× SDS-loading buffer and run in 10% SDS-PAGE. The samples were run to one-third of the gel, stained with Coomassie blue and collected by excising the whole lane for mass spectrometry analysis in the Taplin Mass Spectrometry Facility at Harvard Medical School.

2.5.6 Southern blot analyses

The plant material was lysed in CTAB buffer (100 mM Tris HCl pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl, 2% CTAB, 2% ß-mercaptoethanol); then total DNA was extracted with phenol:chlorophorm:isoamyl alcohol (25:24:1) and precipitated with 2-propanol. The DNA was treated with RNase A and further purified with phenol:chlorophorm:isoamyl alcohol (25:24:1) and precipitated with ethanol, then dissolved in ultrapure water. The specified amount of DNA was separated by electrophoresis in 0.8% agarose, transferred overnight by capillarity to a Hybond-N membrane (Amersham), and probed with ^{32}P-labeled probe targeting the CR region of CaLCuV DNA A. The probe was obtained by PCR using the primers CR_F and CR_R (Supplementary file 8) and labeled using the Rediprime II DNA Labeling System (Amersham) following the manufacturer's instructions.

2.5.7 RNA blot analyses

Total RNA was extracted using Trizol reagent from either adult rosette leaves or 2-week-old seedlings of independent transgenic lines, the RNA blots were then performed as previously described (Zhang et al., 2006a).

2.5.8 Immunoprecipitation of Flag-AGO1-associated small RNAs

Immunoprecipitation of Flag-AGO1-associated small RNAs were performed as described (Zhang et al., 2006a). RNA was recovered with Trizol reagent from the immunoprecipitates, separated in 8 M urea, 15% polyacrylamide gels and subjected to RNA blot analysis of low-molecular-weight RNAs.

2.5.9 Luciferase complementation imaging (LCI) assay

The LCI was performed on 4-week-old *N. benthamiana* leaves infiltrated with various combinations of *A. tumefaciens* GV3101 harboring pCambia Myc-TrAP-nLUC or pCambia Myc-nLUC and *A. tumefaciens* ABI carrying pER10cLUC-3HA or pER10cLUC-3HA-candidate proteins. The agrobacteria containing the pER10 plasmids were incubated with 25μM beta-estradiol for 3 hr prior infiltration, and all the cultures were adjusted to OD_{600} = 0.8. The transfected leaves were assayed 2 days after agroinfiltration by adding the substrate (10 mM luciferin). The sprayed leaves were incubated in total darkness for 5 min and photographed using an electron multiplying charge-coupled device (EMCCD) camera, Cascade II 512, from Photomerics (Roper Scientific). The images were processed with WinView32 Ver 2.5.19.7 (Roper Scientific).

2.5.10 Confocal microscopy and FRET assays

Leaves of 4-week-old tobacco plants (*N. benthamiana*) were agroinfiltrated with syringe without needle as previously described (Zhang et al., 2005) with *A. tumefasciens* ABI carrying pBA-TrAP-CFP and pER10-YFP-Test protein. The agrobacteria containing the pER10 plasmids were incubated with 25 µM beta-estradiol for 3 hr prior infiltration, and all the cultures were adjusted to OD_{600} = 0.8. The plants were maintained for 2 days at 24°C (16 hr light/8 hr dark). The co-localization was evaluated using a Nikon inverted microscope Eclipse T*i*-E. CFP signal was measured by excitation with Shutter 10-3 filter 3 (CFPHQ [Ex]), and emission was detected at 485 nm; YFP used Shutter 10-3 filter 4 (YFPHQ [Ex]) and emission was detected at 540 nm. The images were processed using NIS-Elements-AR 4.30.01 (Nikon) and Adobe Photoshop software.

FRET-AB experiments were performed on *N. benthamiana* epidermal cells of 4-week-old leaves agroinfiltrated with a 1:1 mixture of pBA-TrAP-CFP and pER10-YFP-KYP to a final OD600 = 0.8. YFP and CFP signals were captured with a Zeiss LSM 710 confocal Microscope. FRET was determined by the acceptor photobleaching method (Daelemans et al., 2004; Kenworthy, 2001). First, to define the base line, the signal intensities of a pre-photobleach CFP (donor) and YFP (acceptor) are acquired by exciting with the 458 and 514-nm laser lines, respectively. Then, three regions of interest in the cell were selected: #1, Autofluorescence control; #2, non-photobleaching control; and #3, FRET-AB region. The CFP donor was excited with the 458 nm laser line for all FRET experiments; the emission of both CFP and YFP was recorded at 485 nm and 540 nm. Regions #1 and #3 were rendered free of YFP by consecutive cycles of bleaching recovery

with the 514-nm laser line until no recovery of YFP was detected. The CFP and YFP signals were monitored throughout the experiment. After correction for background with control region #1 and for photobleaching of the donor because of imaging with control region #2, the FRET efficiencies (E) in the region #3 was calculated from the CFP signal using $FRET_{Eff} = 1 - \frac{D_{pre}}{D_{post}}$, where D is the mean intensity of the donor CFP in the area where the acceptor was bleached, before (Dpre) and after (Dpost) acceptor bleaching. The FRET efficiency is considered positive when Dpost > Dpre. The image and statistical analyses were performed with the FRET module for the ZEN software (Zeiss). The average FRET efficiency and its standard deviation were calculated from the FRET efficiencies of each individual cell in 27–30 cells per experiment. The standard Student's *t*-test was used to determine the statistical significance of the results.

2.5.11 HMTase reconstitution assay

In vitro HMTase reactions were modified from (Rea et al., 2000; Tachibana et al., 2001) as follows: 20 µl of reaction mixture containing 3.3 µM Histone 3.2 (NEB), 1 µM His-GST-KYP, and 50 nCi of S-adenosyl-[methyl-^{14}C]-L-methionine in HMTase buffer (50 mM Tris-HCl pH 9, 10 mM $MgCl_2$, 1 mM ß-mercaptoethanol, 250 mM sucrose) was incubated for 0–10 min at 37°C. The reaction products were separated by 18% SDS-polyacrylamide gel electrophoresis and visualized by Coomassie Brilliant Blue R-250 staining; then, the gels were fixed 1 hr in fixing solution (25% Ethanol, 2% Glycerol) and scintillated for 30 min in 1 M sodium salicylate. Gels were dried 2 hr at 80°C. The ^{14}C signal was detected by fluorography using in a preflashed Classic autoradiography film blue sensitive; Filters Kodak Wratten No. 22 and No. 96 were used together for

preflashing. Preflashed film was exposed 5–7 days at −80°C. The film was developed using a Kodak M35A X-OMAT Processor, and the results were digitalized in a Chemi-Doc XRS System and analyzed with the Image Lab Software (Bio-Rad).

To assess the impact of TrAP on KYP activity, His-MBP-TrAP or His-MBP were pre-incubated with His-GST-KYP in different molar ratios, ranging from 0 to 10, for 1 hr at room temperature, then the assays were proceeded as described above. The experiments were performed 3–5 times for statistical analysis.

2.5.12 ChIP assays

The analysis of histone modifications was performed as described (Saleh et al., 2008). Two grams of 9-day-old seedlings were crosslinked with 1% formaldehyde for 10 min by vacuum infiltration at 4°C; the reaction was stopped with 2 M Glycine to a final concentration of 100 mM at room temperature. Plants were rinsed 5 times with ice cold with water, flash-frozen in liquid nitrogen, and ground with mortar and pestle. The powder was suspended in 6 volumes (12 ml) of nuclei isolation buffer (15 mM PIPES-KOH pH 6.8, 0.25 M sucrose, 0.9% Triton X-200, 5 mM $MgCl_2$, 60 mM KCl, 15 mM NaCl, 1 mM $CaCl_2$, 1 mM PMSF, 1 pellet/50 ml Complete EDTA-free Protease inhibitor [Roche]), filtered through two layers of Miracloth and centrifuged at 11,000 × rcf for 10 min in 4°C. After discarding the supernatant, the pellet was resuspended in 1 ml of Nuclei lysis buffer (50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0, 150 mM NaCl, 1% SDS, 0.1% Sodium Deoxycholate, 1% Triton X-100, 1 pellet/50 ml Complete EDTA-free Protease inhibitor [Roche]); the samples were sonicated in 10 cycles 30 s ON and 90 s OFF, using the Bioruptor (Diagenode) at the highest power in 4°Cs. The sonicated samples were

centrifuged for 10 min at 21,000 × rcf in 4°C. 100 µl of the clarified chromatin was diluted 10-fold with Nuclei lysis buffer without SDS for each assay. The immunoprecipitation was accomplished by the addition of 40 µl Protein A Dynabeads (Invitrogen) and 3 µl of the pertinent antibody, followed by 6 hr incubation at 4°C on mild rotation. The beads-conjugated complexes were washed with 1 ml of Low salt buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 150 mM NaCl, 0.5% Triton X-100, 0.2% SDS), followed by 1 ml of high salt buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 500 mM NaCl, 0.5% Triton X-100, 0.2% SDS), then with 1 ml of LiCl buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 1% sodium deoxycholate, 1% NP-40), and finally twice with 1 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) by incubating 5 min at 4°C in between washed. The samples were eluted twice at room temperature with 250 µl of elution buffer (100 mM $NaHCO_3$, 0.5% SDS), for 15 and 30 min, respectively. The samples were decrosslinked with 100 mM NaCl at 65°C overnight, followed by Proteinase K treatment for 90 min at 45°C. The DNA was purified by Phenol:Chloroform:Isoamyl Alcohol 25:24:1, and precipitated in 100% Ethanol at −80°C. The antibodies used are monoclonal anti-H3K9me2 (Abcam, #Ab1220); monoclonal anti-H3K4me3 (Millipore, cat #04-745); monoclonal anti-H3K27me3 (Millipore, cat #07-449).

The immunoprecipitation of Flag-4Myc-KYP-Chromatin complexes was done as in (Wierzbicki et al., 2008), using Anti-FLAG M2 magnetic beads (Sigma, Cat# M8823). Two grams of rosette leaves 1–12 of mock or CaLCuV inoculated plants at 18 dpi were crosslinked with 1% formaldehyde for 25 min by vacuum infiltration at 4°C; the reaction

was stopped with 2 M Glycine to a final concentration of 100 mM. Plants were rinsed five times with ice cold with water, flash-frozen in liquid nitrogen, and ground with mortar and pestle. The powder was suspended in 5 volumes (10 ml) of Honda Buffer (20 mM HEPES-KOH pH 7.4, 0.44 M sucrose, 1.25% ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 0.5% Triton X-100, 5 mM DTT, 2 mM PMSF, 1 pellet/25 ml Complete EDTA-free Protease inhibitor [Roche]), filtered through two layers of Miracloth and centrifuged at 2000 × rcf for 15 min in 4°C. After discarding the supernatant, the nuclear precipitates were washed three times with 1 ml of Honda buffer; subsequently, the pellet was suspended in 300 µl of Nuclei lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS, 2 mM PMSF, 1 pellet/25 ml Complete EDTA-free Protease inhibitor [Roche]) and sonicated in ten cycles 30 s ON and 90 s OFF, using the Bioruptor (Diagenode) at the highest power in 4°Cs. The sonicated samples were centrifuged for 10 min at 21000 × rcf in 4°C. 100 µl of the clarified chromatin was diluted 10-fold with ChIP dilution buffer (16.7 mM Tris-HCl pH 8.0, 1.2 mM EDTA pH 8.0, 167 mM NaCl, 1.1% Triton X-100, 1 pellet/25 ml Complete EDTA-free Protease inhibitor [Roche]) per ChIP. The immunoprecipitation was accomplished by the addition of 40 µl of Anti-FLAG M2 magnetic beads (Sigma, Cat# M8823), followed by 2 hr incubation at 4°C on mild rotation. The beads were washed five times with 1 ml of Washing buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 150 mM NaCl, 1% Triton X-100, 1% SDS, 2 mM PMSF, 1 pellet/25 ml Complete EDTA-free Protease inhibitor [Roche]) incubating 5 min at 4°C in between; then, two more washes with 1 ml TE buffer incubating 5 min at 4°C. Finally, the samples were eluted twice at room temperature with 125 µl of Elution buffer (100 mM $NaHCO_3$, 0.5% SDS), for 15

and 30 min, respectively. The samples were decrosslinked and the DNA extracted as above.

2.5.13 Microarray analysis

Microarray analyses using the Affymetrix ATH1 platform were performed with three biological replicates using wild-type plants, *35S-TrAP* transgenic plants, and *lhp1* mutants. Seedlings were grown on MS medium with 1% sucrose for 7 days. One mg of total RNA was used for reverse transcription using MessageAmp II aRNA kits (Ambion) and 15 mg of labeled cRNA for hybridization. GeneChip hybridization and scanning were performed at the Genomics Resource Center, Rockefeller University, New York.

Statistical analysis of microarray data was performed using R software. Initially the microarray plates were tested for quality by an M plot and the data normalized by the RMA method from the Affy package. Subsequently, the distribution of the samples was assessed with scatter plots and the normalized data sets were evaluated with the Moderate t-test from R package limma for p-value computation. Then, the eBayes method was used to compute moderated t-statistics and log-odds of differential expression by empirical Bayes shrinkage of the standard errors towards a common value. The moderated t-statistic (t) is the logFC to its standard error. In our DEG results our thresholds are p-value < 0.05 and logFC >1 (upregulated) or logFC < −1 (downregulated). The False Discovery Rate was approximated from the eBayes adjusted p-value.

The significance of the overlapping data sets was calculated through Pearson's Chi-squared test with 1° of freedom.

2.5.14 Quantitative PCR and RT-PCR

Expression levels of the tested genes were examined by quantitative RT-PCR. Total RNAs were prepared from 9 days-old seedlings and treated with DNase before being subjected to cDNA synthesis using Superscript III reverse transcriptase (Invitrogen) primed by random primers. The EF1a gene (Williams et al., 2005) was included as an internal control for normalization. The enrichment levels of specific genes after ChIP assay were also tested by quantitative PCR. Primers are listed in Supplementary file 8. The quantitative PCRs were performed in 384-well plates with an ABI7900HT real-time PCR system using the SYBR Green I master mix (Applied Biosystems) in a volume of 10 μl. PCR conditions were as follows: 50°C for 2 min, 95°C for 10 min, 45 cycles of 96°C for 10 s followed by 60°C for 1 min. Three biological repeats were performed, and the reactions were performed in triplicate for each run. The comparative CT method was used to evaluate the relative quantities of each amplified product in the samples. The threshold cycle (CT) was automatically determined for each reaction by the system according to the default parameters. The specificity of the PCR was determined by dissociation curve analysis of the amplified products using the standard method installed in the system.

2.5.15 Whole genome bisulfite sequencing

Approximately 500 ng of genomic DNA were used to generate libraries as described (Feng et al., 2011) using premethylated adapters (NEXTFlex Bisulfite-Seq Adapters, Bioo Scientific #511911). The adaptor-ligated fragments were purified by QIAQuick column (Qiagen) and bisulfite converted using the EpiTect Kit (Qiagen)

following the manufacturer's instructions. The converted DNA was later enriched by 15 cycles of PCR using the Pfu Turbo Cx Polymerase (Agilent), using the specific primers provided by Bios Scientific for enrichment. The library was finally purified with Agencourt AMPure XP beads (Beckman Coulter) according to manufacturer's instruction. The libraries were single-end sequenced using HiSeq High Output with read length of 50 bp. Base calling and sequence cleaning was performed with the standard Illumina software, then the clean reads were mapped to the *Arabidopsis* genome (Version: TAIR10) using Bismark v0.14.3 (Krueger and Andrews, 2011) with default parameters; the PCR duplicates were removed, and the methylation information was obtained with Bismark with cutoff 3. The DMRs were obtained using swDMR (Wang et al., 2015b) with window 200, step size 100, (left 1000, right 1000), the samples were compared using the Kruskal–Wallis analysis of variance with p-value < 0.01. The DMRs were then annotated using BEDtools (Quinlan and Hall, 2010).

2.5.16 CaLCuV pathogenesis assays

3-week-old (~eight true leaves stage) Col-0 wild-type, *kyp* mutant and *Flag-4Myc-KYP* complementation lines were infected by agroinfiltration of the CaLCuV (Arguello-Astorga et al., 2007) or CaLCuVΔ*trap*; disease progression was evaluated daily in terms of symptom development and severity. The assays were replicated 3 times on 36 plants per genotype per assay, grown in short day conditions (8 hr light/16 hr dark). Systemic infection was assessed by Southern blot on samples harvested 18 days after inoculation, when nine to eleven new rosette leaves had emerged (Figure 33B).

3. GENOME-WIDE IDENTIFICATION OF TrAP-TARGETED HOST FACTORS THROUGH PROTEOMICS ANALYSIS

3.1 Overview

The begomovirus encodes a transcriptional activation protein, TrAP, which is essential to systemic infection. Multiple studies have shown that TrAP is a silencing suppressor and can induce genome wide transcriptional reprogramming in the host. Several pieces of evidence also show that TrAP is multi-functional protein but only a limited number of TrAP-interacting proteins (TrIPs) in the host cell have been reported, with most of the data resulting from TrAP expression in heterologous systems. Here we present a comprehensive proteomics study to identify TrIPs *in planta*. Immunoprecipitation of TrAP-protein complexes from *Arabidopsis thaliana* and *Nicotiana benthamiana* hosts, followed by mass-spectrometry, resulted in numerous TrIPs. Gene ontology analyses of the identified TrIPs implicate their functions in a broad array of biological processes such as nucleotide binding and hydrolase activities, amino acid metabolism, transport, and GTP catabolism. The TrIPs were further validated by *in vivo* and *in vitro* protein-interaction assays, whereas the functional relevance of the identified TrIPs in viral pathogenesis has been validated by infectivity assays. In addition to the proteomic approaches, a number of TrIPs were also recovered through genetic mining of Arabidopsis mutants from the ABRC collection that phenocopy TrAP transgenic plants. The TrIPs revealed through genetic approaches appear to be involved in mRNA and chromatin metabolism.

3.2 Introduction

The geminivirus are circular single-stranded DNA (ssDNA) plant pathogens of great economic importance because of their wide host range and symptom severity. Geminiviruses, unlike most plant viruses, do not encode a dedicated polymerase and rely entirely on the host for their replication (Hanley-Bowdoin et al., 2013). Their genomes encode 6-8 proteins, which are sufficient to overcome the host defenses and take control of the machineries that determine chromatin activity and replication. Hence, these viruses have been greatly studied not only to design strategies to prevent the dramatic financial losses, but also to understand fundamental biological processes such as DNA replication, transcriptional regulation, host defense, and gene silencing (Rojas et al., 2005). Within the family Geminiviridae, the genus Begomovirus is the most diverse and widely studied. Begomovirus genomes can comprise one (monopartite) or two (bipartite) ssDNA molecules, named DNA A and DNA B, each of those ranging 2.5 – 3.0 kb. In bipartite begomoviruses, the DNA A molecule codes for replication and regulatory proteins, while DNA B encodes the movement proteins (Fondong, 2013). Monopartite begomoviruses, on the other hand, code for both replication and movement proteins in their ssDNA molecule, but are often associated with smaller satellite DNAs encoding pathogenicity determinants that enhance symptom development and infection (Zhou, 2013).

Given their limited coding capacity, it is not surprising that geminivirus proteins fulfill multiple functions to enable virus infection (Hanley-Bowdoin et al., 2013). One of the most notable is the transcriptional activator TrAP, encoded by the ORF AL2/AC2 or L2/C2 in bipartite and monopartite begomovirus, respectively. The expression of a

functional TrAP is required for the accumulation of the infective form of the virus, and for systemic infection, although it is dispensable for replication. Once the viral ssDNA enters the nucleus, a dsDNA intermediate is synthesized to serve as a replication and transcription template (Pilartz and Jeske, 1992). The viral dsDNA is organized on nucleosomes to form viral minichromosomes in the host nucleus; those, are not mobile and are regulated by the host in a similar fashion as the endogenous chromatin (Paprotka et al., 2015). Viral minichromosomes can be subjected to DNA methylation, histone modification and nucleosome rearrangements (Hanley-Bowdoin et al., 2013; Paprotka et al., 2011; Raja et al., 2008). The transition from the dsDNA replicative form to the ssDNA infective form of the virus depends on the expression of TrAP protein (Hayes and Buck, 1989), as it activates the transcription of the viral late genes (i.e. coat and movement proteins) (Sunter and Bisaro, 1991, 1992). The viral coat and movement proteins bind to ssDNA and compete with the host machinery for the newly synthesized viral DNA; hence, enabling encapsidation into new virions that are shuttled to the cytosol, and transported to neighboring cells, ultimately producing systemic infection. Lack of TrAP function results in accumulation of viral dsDNA at the infection site without spreading of the disease, nor accumulation of virions. This phenotype is reminiscent of viruses lacking the coat protein (Sunter and Bisaro, 1991, 1997).

TrAP protein hijacks host proteins to exert its function during pathogenesis, but host proteins also target TrAP to fight the viral infection. Previous studies, mostly performed using yeast one and two hybrid systems, have identified multiple TrAP-interacting partners, and have contributed to the elucidation of TrAP mechanism.

Specially significant are its interaction with PEAPOD to mediate the transcriptional activation of CP and NSP (Lacatus and Sunter, 2009); inhibition of ADK2, hampering the methyl cycle and preventing trans-methylation of DNA and other substrates (Baliji et al., 2010; Brough et al., 1992; Buchmann et al., 2009; Mohannath et al., 2014; Raja et al., 2008; Wang et al., 2003, 2005); and, its interaction with SnRK1, which phosphorylates TrAP at S109, impairing its function and delaying the onset of viral infection (Hao et al., 2003; Mohannath et al., 2014; Shen and Hanley-Bowdoin, 2006; Shen et al., 2014a). Studies *in planta* have shown that TrAP directly inhibits KYP/SUVH4 to prevent the epigenetic silencing of the viral minichromosome (Castillo-González et al., 2015; Sun et al., 2015). Notably, with the exception of the *kyp* mutant, which barely compensates for the loss of TrAP function (Castillo-González et al., 2015), none of the mutants of the identified TrAP partners have been able to mimic overexpression of TrAP, or sustain systemic infection in the absence of TrAP. In all, the mechanisms and functions of TrAP during pathogenesis are largely unknown (Jackel et al., 2014).

The study of TrAP function is essential for the understanding of the systemic infection process and such study also offers the opportunity of looking into the underlying regulation of gene expression in plants. Here we present a comprehensive analysis of the TrAP-protein complexes *in planta*. Our approach involved candidate searching for mutants that resemble TrAP overexpression line, whereas the other applied an unbiased strategy based on immunoprecipitation (IP) followed by mass spectrometry (MS) of Flag-Myc4 (FM)-TrAP from tomato golden mosaic virus (Begomovirus) in both *Nicotiana benthamiana* (*N. benthamiana*) and *Arabidopsis thaliana* ecotype Columbia (Col-0),

further combined with pathogenicity assays to show the relevance of the identified TrIPs during infection. Genetic mining through phenotypic comparison of TrAP transgenic lines with the ABRC mutant collection recovered three candidate TrIPs: SERRATE (SE) and ENHANCED SILENCING PHENOTYPE 3 (ESP3), which are involved in RNA metabolism; and RELATIVE OF EARLY FLOWERING 6 (REF6) which is engaged in chromatin regulation. In comparison, IP-MS recovered 51 and 30 potential TrIPs from *A. thaliana* and *N. benthamiana*, respectively. Gene ontology analyses showed high enrichment of TrIPs with nucleotide binding and hydrolase activities, mainly involved in amino acid metabolism, transport, and GTP catabolism. Consistent with previous reports, we found that TrAP interacts with proteins involved in the methyl metabolism; specifically, our data suggest that S-adenosyl homocysteine hydrolase (SAHH1) is a direct TrIP that can mediate not only the interference of the methyl cycle but also multiple trans-methylation reactions in the cell. Further analyses of some specific candidates let us to conclude that TrAP likely moves across the nuclear pore complex (NPC) as a large protein complex that requires active nuclear transport and specific nucleoporins. Together, our findings provide new insight into the molecular functions of TrAP, and offer a comprehensive resource for the community of plant virologists to further geminivirus pathogenesis.

3.3 Results

3.3.1 Identification of TrAP targeted host factors in planta

To mimic the temporal regulation of TrAP expression during infection, we isolated TrAP complexes from leaves of the tomato golden mosaic virus host plant, *N. benthamiana*, which was inoculated with *Agrobacterium tumefasciens* to induce transient expression of FM-TrAP. Given that geminivirus infection requires that the host cells are in a DNA replication competent state, we also generated *Arabidopsis thaliana* (Col-0) stable transgenic lines, expressing the FM-TrAP protein from the β-estradiol inducible promoter XVE (Zuo et al., 2000). In both cases, we opted for a small N-terminal fusion tag to liberate the transcriptional activation domain located at the C-terminus, potentially enabling the identification of proteins involved in this function.

We isolated the TrAP protein complexes following a stringent two-step IP protocol, optimized by monitoring the enrichment of FM-TrAP through western blot (Figure 36A). Briefly, proteins of 10g of ground tissue (*N. benthamiana* leaves or 9-day-old *A. thaliana* seedlings) were extracted, immunoprecipitated with anti-FLAG antibody, and washed to minimize non-specific binding. The FLAG-purified TrAP complexes were eluted by competition with FLAG peptide and successively isolated using anti-c-Myc antibody, thoroughly washed, and fully eluted using ammonium hydroxide. The final elutes were run in a 4-20% SDS-PAGE, and visualized by Coomassie staining. We excised the TrAP-specific and the control bands for LC/MS/MS analysis to the Taplin Lab in Harvard Medical School. We did this in parallel with mock-infected *N. benthamiana* plants (Figure 36B), and with FM-βC1 (data not shown) as controls. Due to the fact that

TrAP expression in *N. benthamiana* was low, and the fact that there is plant-to-plant variation, TrAP-complexes from ten independent experiments (~100g of tissue) were pooled together prior to the MS analysis. An identical procedure on *N. benthamiana* leaves agroinoculated with the empty vector was used as a negative control. Conversely, FM-TrAP expression from induced *A. thaliana* seedlings was highly efficient and accounted for higher TrAP-complex yields.

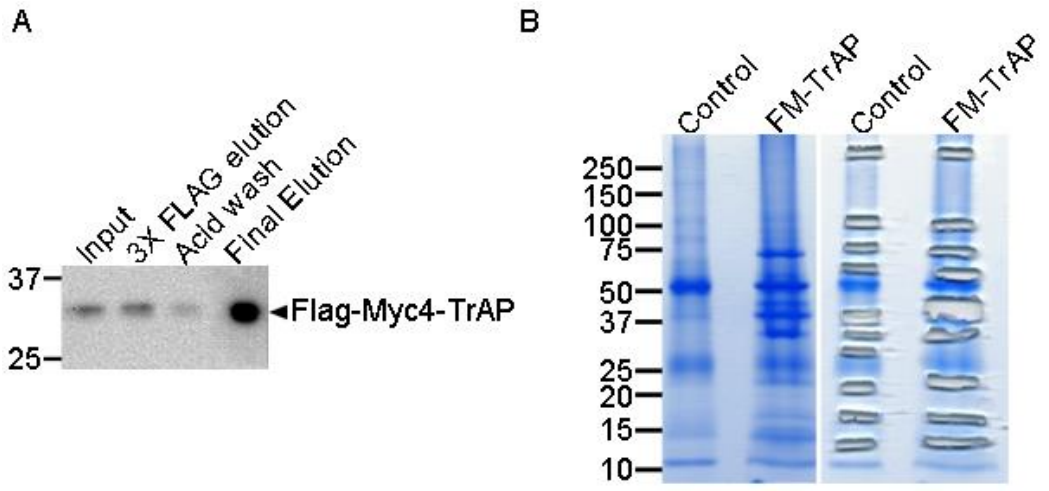


Figure 36. Isolation of TrAP-protein complexes from *Nicotiana benthamiana*.
(A) Enrichment of the TrAP signal through successive purification steps was monitored by western blot. Briefly, proteins from ground tissue are extracted in IP buffer (input) and incubated with anti-FLAG magnetic beads, the beads are washed and the TrAP-complexes are eluted by competition with 3X FLAG peptide (3X FLAG elution). The eluted TrAP complexes are subsequently incubated with anti-Myc agarose resin, while we strip the anti-FLAG magnetic beads to ensure that we recovered most of the TrAP complexes (acid wash). Finally the TrAP-conjugated anti-Myc resin is washed and eluted with ammonium hydroxide (Final Elution). 10ul samples of each described step were run in a 12% SDS-PAGE and enrichment of TrAP was assessed by western blot. A single experiment from 10g of ground *N. benthamiana* leaves is shown. (B) Successful isolation of TrAP protein complexes *in planta*. Final elutes of ten experiments from Control and TrAP-transient expression plants were pooled together and lyophilized prior to running a 4-20% SDS-PAGE. The proteins were visualized with Coomassie staining and the unique bands were excised.

In order to minimize the noise from the IP-MS analysis, we included in our study only proteins that were recovered by at least two unique peptides but not in the control samples. As such, we identified two sets of 51 (APPENDIX B, Table 2) and 30 (APPENDIX B, Table 3) potential TrIPs from *N. benthamiana* and Arabidopsis seedlings, respectively. Notably, the data set recovered from the *N. benthamiana* tissue provided more unique peptides that allowed higher coverage and confidence for the identification of TrAP-co-immunoprecipitated proteins than the dataset obtained from Arabidopsis seedlings, this is likely because of the larger sample size employed for the *N. benthamiana* (~10X) MS experiment.

To validate the results from the proteomics analyses, we randomly selected twelve candidates from both data sets, cloned the genes from *A. thaliana* Col-0 cDNA and transiently expressed them in *N. benthamiana* leaves as fluorescently tagged proteins. Specifically, we co-expressed TrAP-CFP and the YFP-target for assessment of co-localization by confocal imaging. Notably, eight of the twelve protein pairs were positive for the co-localization assays, two were negative, and in two cases we could not detect the expression of the YFP-target protein (Table 1).

Table 1. Summary of proteins tested for co-localization with TrAP-CFP fusion protein as validation of MS data

IP-MS Source	*N. bentha*	*A. thaliana* gene ID	Gene Name	Colocalization
A. thaliana	Not detected	AT5G13960	KYP	Positive with nuclear speckles
N. benthamiana	AGO4-2	AT2G27040	AGO4	Positive, general nuclear and cytoplasmic colocalization
A. thaliana *N. benthamiana*	S-adenosyl-L-homocysteine hydrolase	AT4G13940	Hog1	Positive, apparent quenching of CFP fluorophore upon expression of YFP, indicative of FRET
N. benthamiana	Leucine-rich repeat protein	AT3G12610	DRT-100	Positive, relocalization of YFP-DRT100 signal when co-expressed with TrAP-CFP
N. benthamiana	O-methyltransferase	AT5G54160	OMT1	Positive, general nuclear and cytoplasmic co-localization (filaments)
N. benthamiana	AGO1-2	AT2G32940	AGO6	Positive, apparent quenching of CFP fluorophore upon expression of YFP, indicative of FRET
N. benthamiana	S-adenosyl-L-methionine synthase	AT4G01850	SAM2	Positive, nuclear and cytoplasmic co-localization
A. thaliana *N. benthamiana*	GTP-binding nuclear protein Ran-B1	AT5G20010	RAN1	Undetectable expression of YFP-RAN1
N. benthamiana	MAP kinase kinase	AT4G29810	MKK2	Positive nuclear and cytoplasmic co-localization. TrAP-CFP signal is excluded from nucleolus when co-expressed with YFP-MKK2
N. benthamiana	FtsZ-like protein 2	AT3G52750	FTSZ2-2	Negative
N. benthamiana	Argonaute 1	AT2G27880	AGO5	Undetectable expression of YFP-AGO5
N. benthamiana	Cell division control protein	AT3G09840	CDC48	Negative

3.3.2 Functional categorization of the potential TrAP-interacting proteins

As an initial step to characterize the TrAP protein complexes we performed Gene Ontology (GO) analyses of the *A. thaliana* and *N. benthamiana* data sets using the GO descriptions available from TAIR10 and from *N. benthamiana* genome V1.0.1 (Sol Genomics), respectively. Possibly due to the more complete annotation of the Arabidopsis genome, many more GO-enriched categories were identified in the Arabidopsis-derived data set (Figure 37A) than from the *N. benthamiana* (Figure 37B). To overcome the incipient annotation of the *N. benthamiana* genome and to fully explore the information from this data set, we searched for homologs in the *A. thaliana* genome and looked for

enrichment of GO terms (Figure 37C). We identified homologs for 49 out of the 51 recovered proteins. Notably, most of the enriched GO terms were commonly identified in both datasets (Figure 37A and C), further validating the results obtained by this unbiased method, and pointing to a universal mechanism for TrAP protein in different hosts.

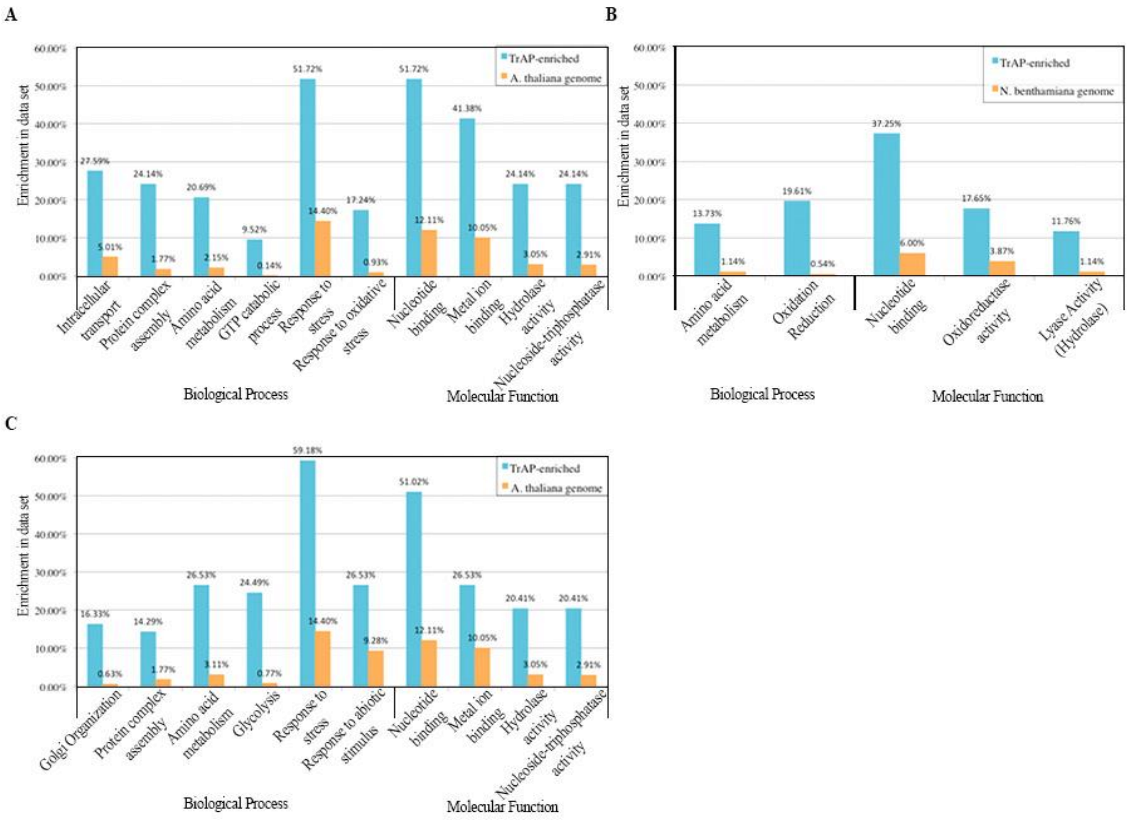


Figure 37. Gene ontology analysis of the tomato golden mosaic virus TrAP co-immunoprecipitated proteins from two host plants.
(A) Enrichment of GO terms among TrAP co-immunoprecipitated proteins in the *A. thaliana* data set. (B) Enrichment of GO terms among TrAP co-immunoprecipitated proteins in the *N. benthamiana* data set. (C) Enrichment of GO terms among *A. thaliana* homologs of TrAP co-immunoprecipitated proteins in the *N. benthamiana* data set.

Specifically, according to the enrichment of GO terms, the proteins that are co-immunoprecipitated with TrAP are related to several biological processes such as assembly of protein complexes, amino acid metabolism, GTP catabolism and response to stress. Regarding the molecular function, the most enriched GO terms among the recovered proteins were nucleotide binding, followed by metal ion binding, hydrolase and nucleoside-triphosphatase activities. Interestingly, the enriched molecular function GO terms were identical between the two datasets, suggesting that TrAP might specifically interfere with these functions in the cells. The differentially enriched GO terms were uniquely identified in the *N. benthamiana* data set and include response to abiotic stimulus (p-value = 6.69×10^{-10}), glycolysis (p-value = 6.70×10^{-13}), and, to a lesser extent, defense response to bacterium (p-value = 8.09×10^{-4}). These divergences might be the result of a larger sample employed for the MS analyses, or a consequence of the infiltration of *A. tumefasciens* on the *N. benthamiana* leaves to transiently express FM-TrAP. Furthermore, unlike the stable transgenic plants, in which all cells are capable of expressing FM-TrAP upon induction with β-estradiol, not all the cells in the transient expression system express the protein; this was evidenced in the lower accumulation of FM-TrAP per mg of tissue in *N. benthamiana* as compared to the transgenic *A. thaliana* seedlings. Hence, we expect some dilution of the TrAP-specific signal.

3.3.3 TrAP interferes with the methyl cycle through interaction with SAHH1

One of the first reported TrAP-interacting proteins was ADK2. The protein was identified from yeast-two-hybrid experiments, and shown to be inhibited upon interaction with TrAP (Wang et al., 2003, 2005), causing a defect in the methyl cycle in the cells and

ultimately impairing DNA methylation and gene silencing (Baliji et al., 2010; Buchmann et al., 2009; Raja et al., 2008). ADK2 appeared to be absent in our IP-MS results both from *A. thaliana* and *N. benthamiana,* likely due to the incomplete recovery of TrIPs from the SDS-PAGE of the immunoprecipitated proteins from *N. benthamiana* and /or due to the relative low sequencing depth for the TrAP-cofactors in *Arabidopsis*. Notably, two other enzymes active in the same pathway were co-immunoprecipitated with TrAP, the S-adenosyl homocysteine (SAH) hydrolase (SAHH1, Figure 38A-B) and S-adenosyl methionine (SAMe) synthase (SAM2, Figure 39A).

SAHH1, also known as *HOMOLOGY-DEPENDENT GENE SILENCING 1* (*HOG1*), was initially discovered in a genetic screening for revertants of silencing at the chalcone synthase locus, TT4, in the C-line (Furner et al., 1998; Loach et al., 2005).

Particularly, SAM2 was also recovered from the MS results and further confirmed by IP (Figure 39A) and LIC (data not shown), however it was not recovered from the in vitro pull-down, suggesting the interaction with TrAP is mediated by other cellular factors. Alternatively, this interaction might entail specific conditions that we have not yet optimized *in vitro*.

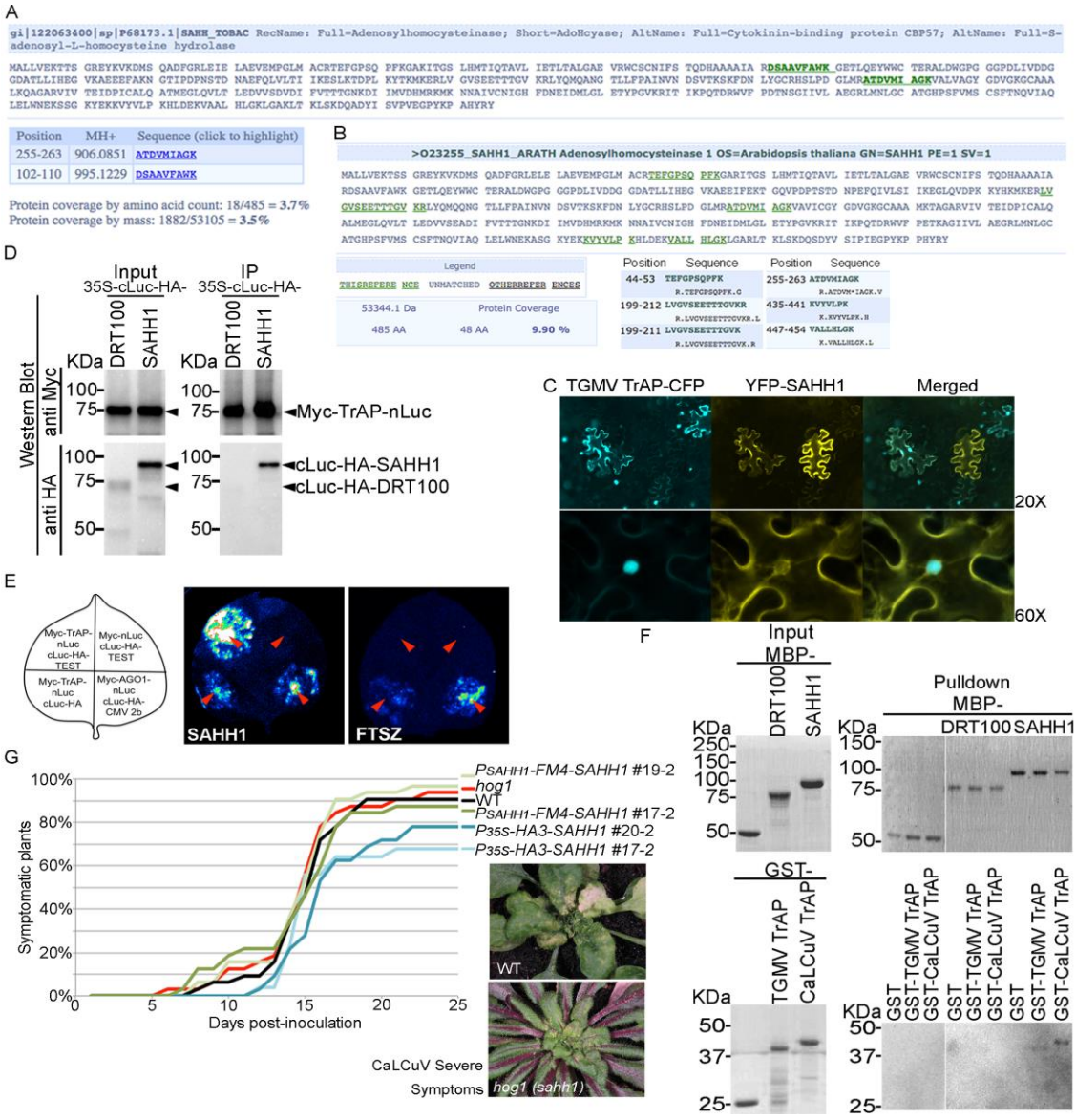


Figure 38. TrAP interacts with SAHH1

(A-B) IP-MS of TrAP complexes from *N. benthamiana* and *A. thaliana* recovered 2 and 6 SAHH1-specific peptides (green), respectively. (C) Localization of YFP-SAHH1 signal is cytoplasmic, and overlaps with the cytoplasmic signal from TrAP-CFP. (D) Co-immunoprecipitation of SAHH1 and TrAP, left panel shows input proteins while right panel shows IP. (E) Luciferase Complementation Imaging Assay. (F) *In vitro* pull-down assay; left panel shows the input proteins, while right panel shows the proteins pulled down using amylose resin. (G) Time course of CaLCuV symptom development in different genetic backgrounds; overexpression of SAHH1 (P35S-HA3-SAHH1) results in a moderate decrease of symptomatic plants, without obvious difference in the severity of CaLCuV infection as compared to the wild type.

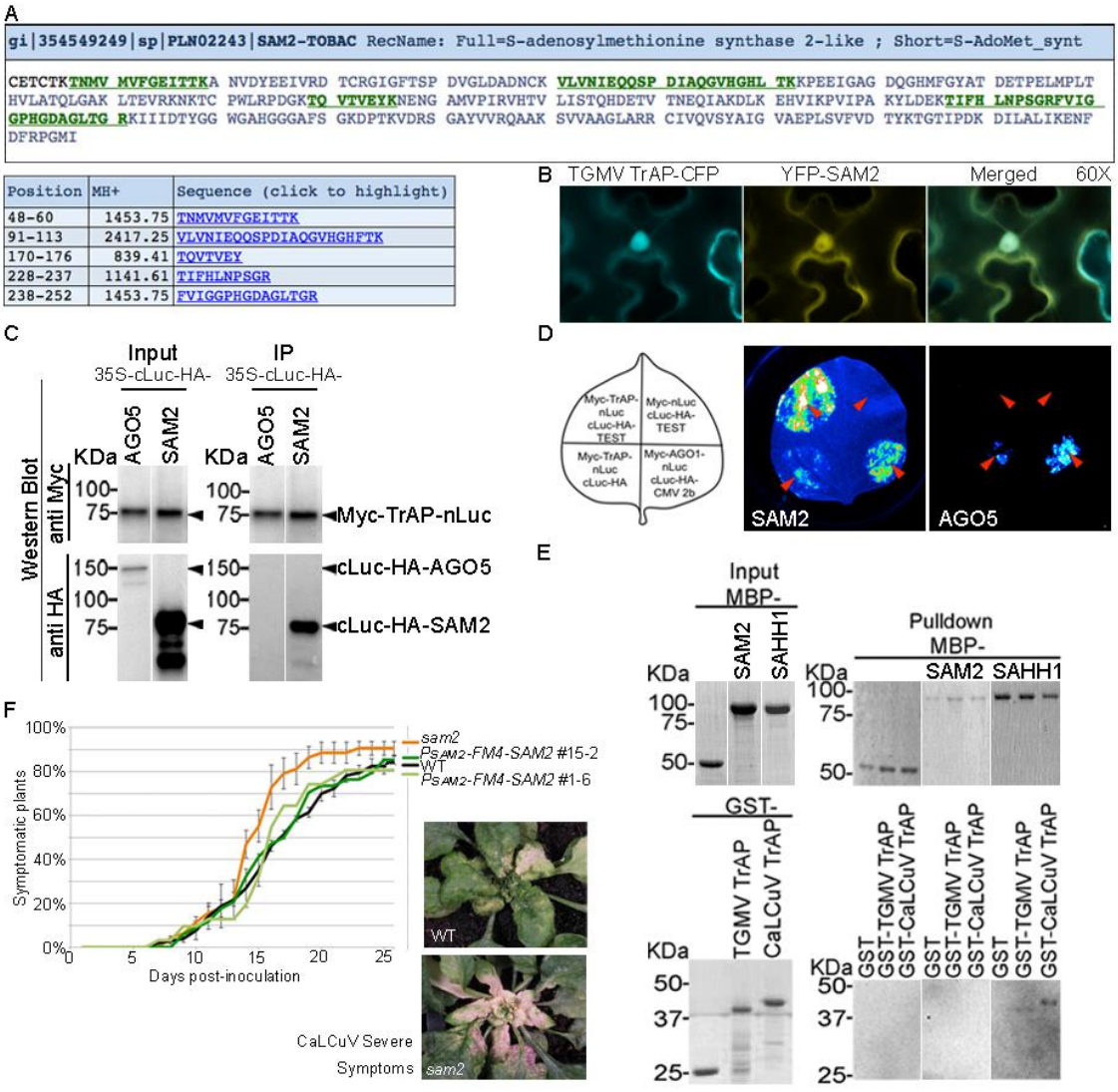


Figure 39. TrAP interacts with SAM2

(A) IP-MS of TrAP complexes from *N. benthamiana* recovered 5 SAM2-specific peptides (green). (B) Localization of YFP-SAM2 signal is nuclear and cytoplasmic, and overlaps with the signal from TrAP-CFP. (C) Co-immunoprecipitation of SAM2 and TrAP, left panel shows input proteins while right panel shows IP. (D) Luciferase Complementation Imaging Assay. (E) *In vitro* pull-down assay; left panel shows the input proteins, while right panel shows the proteins pulled down using amylose resin. MBP-SAM2 cannot pull-down GST-TrAP under the tested conditions (F) Time course of CaLCuV symptom development in different genetic backgrounds; *sam2* mutant plants are hypersusceptible to CaLCuV infection as compared to the wild type.

3.3.4 TrAP contributes to Geminivirus pathogenesis through its interaction with the nuclear transport factor RAN1

The Ras-Related Nuclear Protein (RAN) proteins are required for nuclear transport and they have been proposed as regulators of cell cycle progression in mammals and yeast (Haizel et al., 1997). Two peptides HLTGEFEK and SNYNFEKPFLYLAR were mapped exactly to the *N. benthamiana RAN1* homolog, and to the RAN1, 2 and 3 proteins in Arabidopsis (Figure 40A). As RAN1 protein is the model protein from RAN family in Arabidopsis, we chose it for our experiments.

In order to confirm the interaction *in vivo*, we performed LCI and CoIP with nLUC-TrAP-Myc and cLUC-HA3-RAN1 fusion proteins. Co-expression of the constructs successfully complemented luciferase activity (Figure 40B) and nLUC-TrAP-Myc was able to pull down cLUC-HA3-RAN1 under our experimental conditions (Figure 40C). Next we used the YFP-RAN1 fusion protein to analyze the cellular distribution of RAN1 (Figure 40D), as expected for a nuclear transporter, we found RAN1 both in nucleus and cytoplasm. Particularly, we evidenced the high mobility of the YFP-RAN1 across the cells, to and from the nucleus as little packages or spots. We then asked whether YFP-RAN1 co-localized with TrAP-CFP by co-expressing the proteins in our *N. benthamiana* system. However, the cell size and shape were dramatically distorted and we could only evidence the ubiquitous distribution of both signals (CFP and YFP) in the cells (data not shown). Together, these results indicate that RAN1 and TrAP interact *in vivo*.

To further investigate whether TrAP and RAN1 interacted with each other directly, we purified recombinant MBP-RAN1 protein, and performed an *in vitro* pull-down assay.

We tested the interaction of RAN1 with TrAP proteins from TGMV and CaLCuV, given that TGMV was used to identify RAN1, and that *A. thaliana* is a host of CaLCuV but not of TGMV. KYP was used as a positive control, while SAM2 served as negative control. We found that RAN1 interacts directly with both TrAP proteins, although the affinity might be lower than that of TrAP-KYP interaction (Figure 40E).

Next we asked whether TrAP-RAN1 interaction was relevant to the virus infection. To this end, we obtained *ran1* homozygous mutant (SALK_138680C). The mutant plants did not show any obvious phenotype either during development or flower maturation, likely due to functional redundancy from three additional *RAN* genes in Arabidopsis. We proceeded with the infection by CaLCuV and the assessment of the disease progression in terms of time and severity. In four independent experiments, each with >30 plants, we have evidenced a much slower disease development in the *ran1* mutants (Figure 41A), which has not been accompanied by changes in the severity of the symptoms exhibited in the infected plants (Figure 41B-D). Interestingly, the strong infection phenotype observed in the *ran1* mutants is not consistent with complete functional redundancy of the *Arabidopsis* RAN proteins.

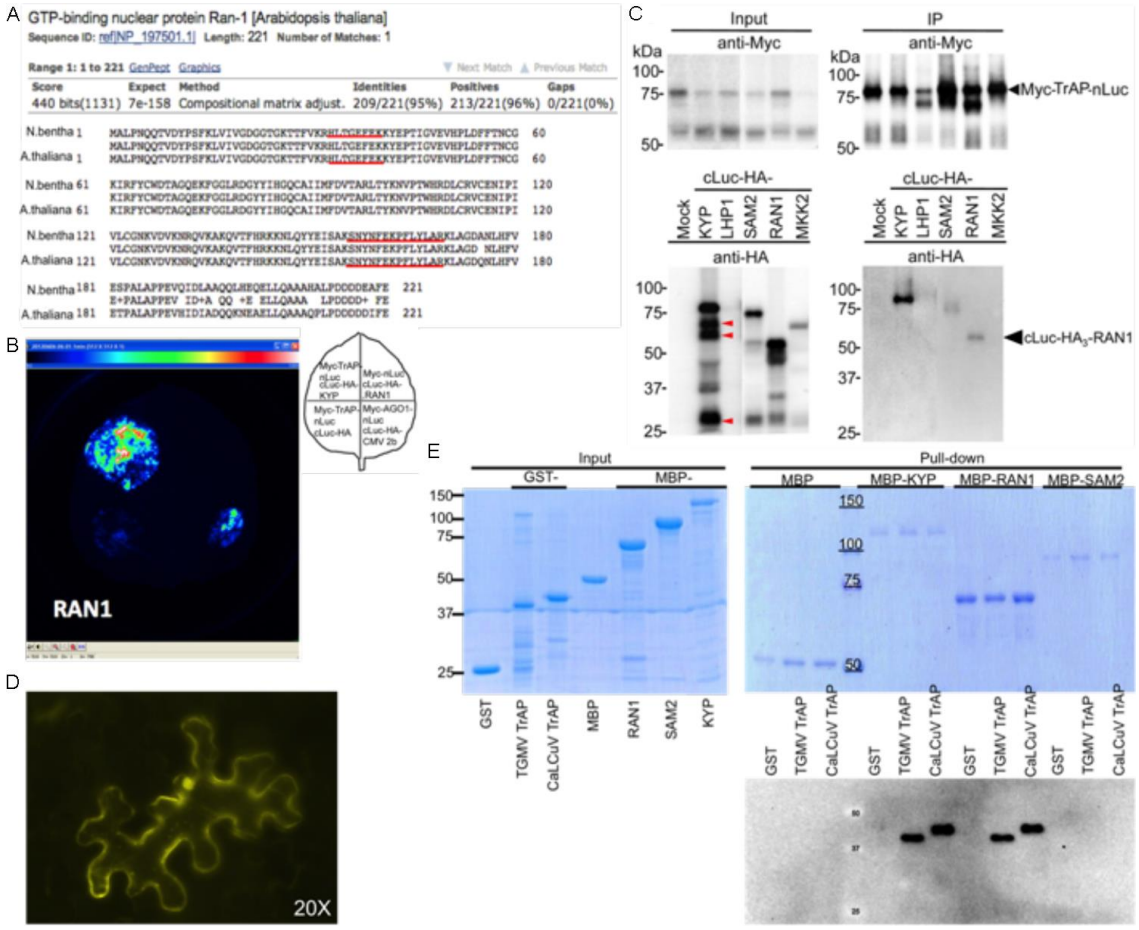


Figure 40. RAN1 and TrAP interact *in vivo* and *in vitro*.
(A) BLAST Alignment of *N. benthamiana* and *A. thaliana* RAN1 proteins, the peptides obtained by LC/MS/MS are underlined in red in both proteins. (B) Luciferase Complementation Imaging Assay. (C) Co-immunoprecipitation of RAN1 and TrAP, left panel shows input proteins while right panel shows IP. (D) Localization of YFP-RAN1 is both nuclear and cytoplasmic; note the discrete spots of YFP-RAN1 in the cytosol. (E) *In vitro* pull-down assay; left panel shows the input proteins, while right panel shows the proteins pulled down using amylose resin. MBP-RAN1 can pull down both CaLCuV and TGMV TrAP proteins.

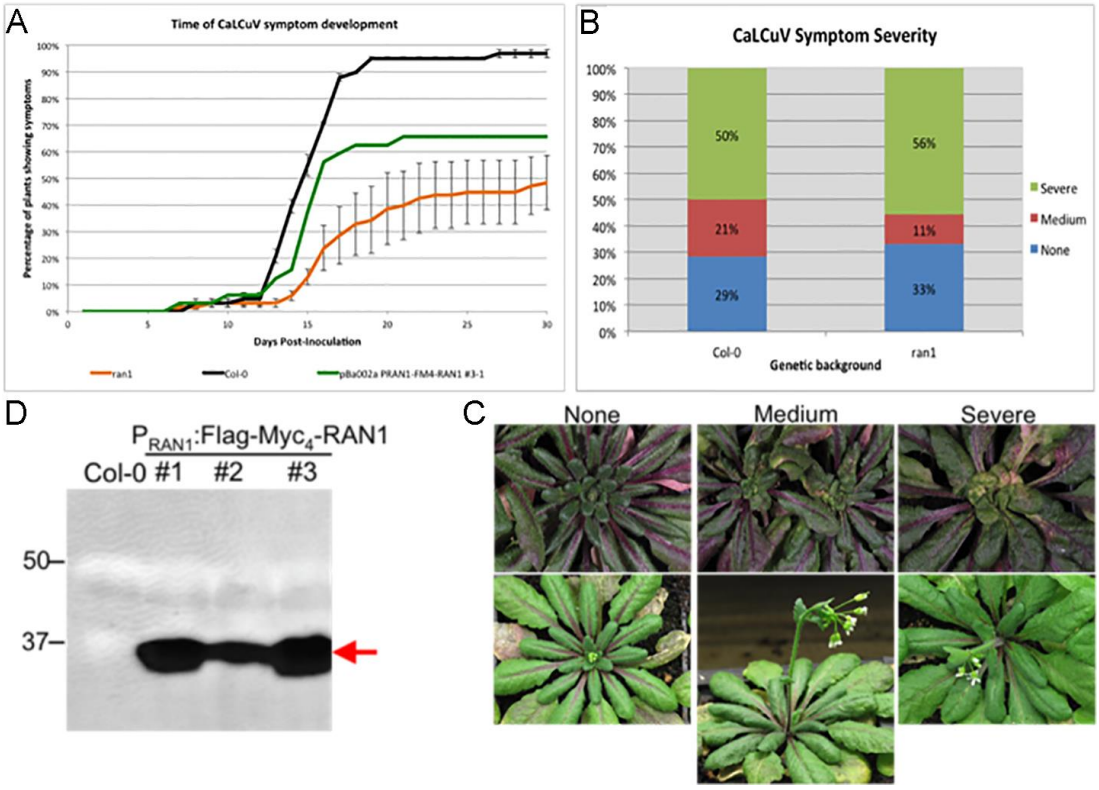


Figure 41. RAN1 is involved in the CaLCuV pathogenesis.
(A) Time of CaLCuV symptom development in different genetic backgrounds. (B) There was no obvious difference in the severity of CaLCuV infection in *ran1* mutants. (C) Scale of CaLCuV symptoms in *ran1* mutant plants. (D) Western blot using anti-Myc antibody to detect expression of FM-RAN1 in the complementation lines.

We have produced several independent RAN1 complementation lines (Figure 41D), which express the tagged FM-RAN1 gene from its native P_{RAN1} promoter. Importantly, the plants show no developmental abnormalities, but only show partial complementation of the virus infection phenotype (Figure 41A). The lack of complementation might be a consequence of the FM-tag, or alternatively to the insertion site of the construct. Furthermore, it would be really interesting to see the effect of the overexpression of RAN1 in the development of infection.

3.3.5 TrAP interacts with the RNA processing protein ESP3

Enhanced Silencing Phenotype 3 (ESP3) is a DEAH RNA helicase domain containing protein of 118.8KDa, known as the only homologue of the yeast PRP2 in Arabidopsis. There are four related DEAH RNA helicases in yeast known to be essential for splicing, PRP16, 22, 43, and 2. However, the function of ESP3 has not been characterized in Arabidopsis. ESP3 was originally identified as part of the putative minimal set of plant specific genes required for normal embryo development (Tzafrir et al., 2004). Later, ESP3 was found to be involved in RNA metabolism, through a mutant screening for an enhanced silencing phenotype. In the study, seven ESP genes were identified and all of them were related to RNA processing. Interestingly, the other six turned out to be part of transcription termination and 3'-end formation, exhibiting accumulation of read-through RNAs; *esp3* mutants, on the other hand, did not show such accumulation. As a homolog of yeast PRP2, ESP3 was proposed to prevent the splicing RNAs to enter the RNAi pathway, and therefore protect them from silencing (Herr et al., 2006). The T-DNA insertion line (a null mutant) shows early flowering, reduced stature and altered leaf morphology at the seedling stage, similar to the TrAP transgenic lines.

We started with experimental validation of the ESP3-TrAP interaction. During the preliminary screening for TrAP interacting factors, we evidenced Luciferase complementation in *N. benthamiana* (Figure 42A). However, we were unable to further confirm this interaction using CoIP possibly due to the extremely low expression of the ESP3 fusion proteins in the plants. Then we asked whether TrAP and ESP3 located to the same cellular compartments. To this end, we co-expressed TrAP-CFP and YFP-ESP3 in

N. benthamiana and evaluated the cellular localization of the fusion proteins by confocal microscopy. When expressed alone YFP-ESP3 localized to specific nuclear compartments, forming speckles that vary in number, size and distribution among nuclei (Figure 42B). On the other hand, TrAP-CFP was more uniformly distributed in the nucleus with higher concentration in the nucleolus, and the phenotype was almost invariable among nuclei and among TrAP from different viruses (i.e. TGMV or CaLCuV) (Figure 42C). When co-expressed, the accumulation of the CFP-TrAP signal to the precise location of the YFP-ESP3 was conspicuous (Figure 42D). Importantly, this was a phenotype unique to two proteins in our screening of more than forty proteins. Together, these results suggested a possible involvement of ESP3 with TrAP during the virus infection.

esp3 mutants showed enhanced silencing of Potato Virus X (PVX) derived transcripts, so we asked whether *esp3* plants would also show enhanced silencing of begomovirus derived RNAs, possibly rendering them more resistant to infection. We obtained seeds for the homozygous knockout stock SALK_021156C from the ABRC, and tested them for progression and severity of the CaLCuV infection. Interestingly, we observed hypersensitivity to the virus infection regarding both time of symptom development (Figure 43A) and severity of the disease (Figure 43B-C). Then, we made the pBA-ESP3-HA_3 and pBA002a-P_{ESP3}:FM-ESP3 constructs and used them to complement the mutant background. We were unable to detect 35S-ESP3-HA_3 expression in the transformants; however, we obtained several lines expressing FM-ESP3 from the native promoter P_{ESP3} (Figure 43D). We chose two of them to test the susceptibility to the

CaLCuV infection. We observed a delay in the symptom development when compared to the mutant background (Figure 43E).

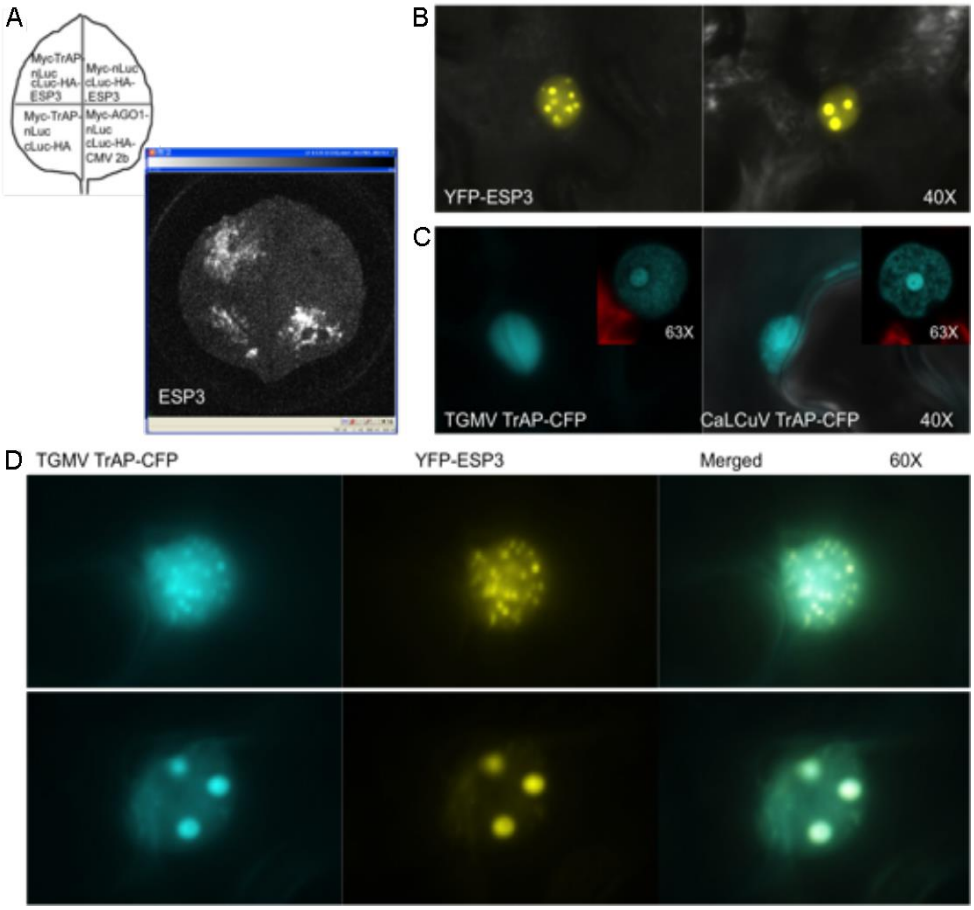


Figure 42. YFP-ESP3 and TrAP-CFP co-localize in nuclear speckles.
(A) Luciferase complementation imaging assay. (B) Nuclear localization of YFP-ESP3 in defined speckles varying in number and size. (C) TrAP-CFP localization is mostly nuclear, and uniformly distributed with higher abundance in the nucleolus. (D) Co-expression of YFP-ESP3 and TrAP-CFP changes TrAP-CFP nuclear localization pattern to accumulate in the YFP-ESP foci.

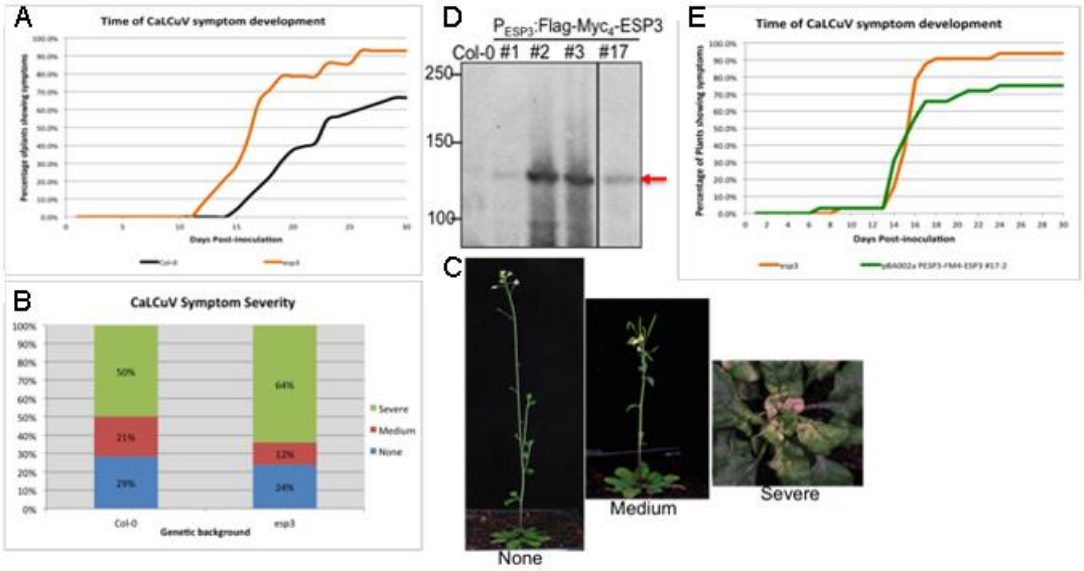


Figure 43. *esp3* mutant *Arabidopsis* plants are hypersusceptible to virus infection.
(A) CaLCuV symptoms develop faster in esp3 mutants than in WT plants. (B) CaLCuV infection is more severe in *esp3* mutants than in WT plants. (C) Scale of CaLCuV symptom severity. (D) Western Blot using anti-Myc antibody to detect the expression of FM-ESP3 protein in the complementation lines. (E) The CaLCuV infection is slower in the P_{ESP3}:FM-ESP3 complementation line than in the *esp3* mutant.

3.3.6 TrAP interacts with the histone demethylase REF6

Previous studies have reported a striking resemblance of the TrAP transgenic plants to several mutants in the TGS pathway, including but not limited to *lhp1* and *clf* mutants (Castillo-González et al., 2015). TrAP transgenic plants exhibit early flowering phenotype and upward curling of leaves, accompanied with loss of H3K27 methylation. CLF catalyzes the deposition of H3K27me3 heterochromatic mark. LHP1, also termed TERMINAL FLOWER 2 (TFL2) (Gaudin et al., 2001; Kotake et al., 2003; Nakahigashi et al., 2005; Sung et al., 2006; Zhang et al., 2007b) associates to heterochromatin and can direct the spreading of the silent status to adjacent euchromatin (Turck et al., 2007; Zhang

et al., 2007b). The coordinated activities of CLF and LHP1 result in chromatin methylation and transcriptional repression (Saleh et al., 2007).

In the same genetic pathway of *CLF* and *LHP1* genes is a gene called Relative of Early Flowering phenotype 6 (REF6), REF6 encoded a Jumonji domain involved in H3K27 demethylation. Strikingly, we observed phenotypic similarity between REF6 and TrAP overexpression plants in *A. thaliana*, suggesting that they might have genetic or even biochemical interaction. In line with this hypothesis, the two proteins co-localized in nuclear speckles (Figure 44). Further biochemical assays to validate the interaction between REF6 and TrAP are in progress.

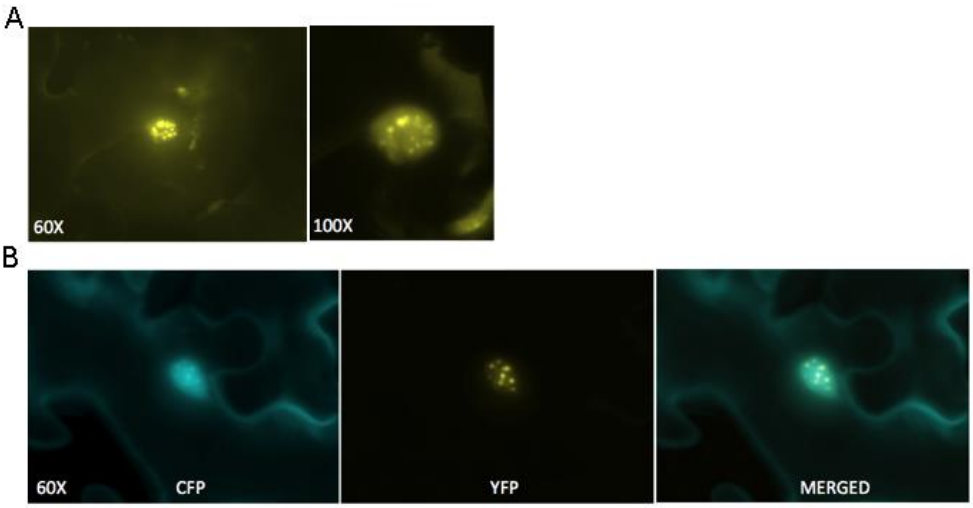


Figure 44. TrAP-CFP and YFP-REF6 co-localize in nuclear speckles.
(A) Nuclear localization of YFP-REF6 in defined speckles varying in number and size. (B) Co-expression of YFP-REF6 and TrAP-CFP changes TrAP-CFP nuclear localization pattern to accumulate in the YFP-REF6 foci.

3.4 Discussion

3.4.1 TrAP is engaged in multiple protein complexes and interferes with a plethora of cellular processes

Arguably, early expression of TrAP would result in suboptimal infection due to very low accumulation of virus templates for transcription and replication. Thus, TrAP function must be regulated. Two main ways of regulation have been proposed and studied: transcriptional regulation of *TrAP* expression, and functional regulation after translation. *TrAP* gene is expressed from the second ORF in the counter virion sense strand on the virus genome. This ORF overlaps at its 5'end with the 3'end of the Replicase gene, and at its 5'end with the 3'end of the replication enhancer. Both Rep and REn are essential for virus infection, and unlike TrAP, they are early-expressed genes. The Rep protein inhibits its own expression by binding to its own promoter and shifting to the expression of the *TrAP* ORF once a threshold amount of Rep, and therefore dsDNA template has been accumulated. Hence the timing of TrAP expression is regulated by the expression of the viral proteins.

3.4.2 RAN1 modulates geminivirus pathogenesis

The RAN protein is a highly conserved eukaryotic GTPase factor of the Ras-family of proteins, which is essential for the transport of proteins and RNA through the nuclear pores. The mammalian and yeast homologues of RAN have been involved with the cell cycle regulation (Bischoff and Ponstingl, 1991; Coutavas et al., 1993). The Arabidopsis genome codes for four RAN genes, numbered 1-4, which are located on the chromosome 5, and produce four RAN proteins with 70-98% identities. Specifically, RAN1, 2, and 3

share 93-98% sequence identities, and RAN4 is the most divergent, sharing only 70% of its protein sequence.

The canonical mechanism of nucleo-cytoplasmic movement of cargos is RAN dependent and is regulated by two factors: the gradient of RAN-GTP/RAN-GDP around the nuclear pore, and the specific binding of the cargo to the receptor (importin/exportin) through the signal peptide (NLS/NES). The cargo does not typically bind to RAN, but to the importin/exportin protein, while RAN serves as the energy provider for the translocation. The importin/exportin binds to proteins in the nuclear pore called nucleoporins that provide a transient docking site for the moving receptor-cargo complex through the pore. The receptors then bind to the RAN protein to release the cargo at its destination. The interaction of TrAP with RAN is not really expected if TrAP is a canonical cargo.

The nuclear pore structure can be imagined as a sieve made of nucleoporins interacting among them through a Phe-Gly rich motif; the normal (closed) pore allows the diffusion of small molecules (<30KDa), but the "aperture" size can be regulated by the interaction of the Phe-Gly nucleoporins with the receptors in the receptor-cargo complex, allowing the translocation of large cargos (Doucet and Hetzer, 2019).

TrAP is a very small protein of only 15kDa. That TrAP contains a NLS is unexpected because it could diffuse freely through the pore, even as a TrAP dimer. The requirement of NLS for the nuclear localization of TrAP suggests that TrAP is by default in a protein complex larger than the TrAP dimer and that the cellular localization is regulated. This is consistent with the proposed phosphorylation-dependent localization of

TrAP. Previous studies have provided genetic evidence of the TrAP/C2 interaction with the importin Karyopherin α during infection, however that interaction seemed to interfere rather than promote virus infection in several geminivirus species (Hanley-Bowdoin et al., 2013).

The virus translocation depends on the expression of the TrAP-dependent BR1 gene, which codes for the nuclear shuttle protein (NSP). NSP interacts with a host GTPase, NIG (for NSP-interacting GTPase), which provides the energy for the NSP-virus nuclear translocation. In fact, overexpression of NIG enhances susceptibility to geminivirus infection in Arabidopsis and *N. benthamiana* systems. It has been proposed that NSP-virus nuclear translocation can function independently from RAN-GTPase (Carvalho and Lazarowitz, 2004; Carvalho et al., 2008a, 2008b).

On the other hand, the Arabidopsis genome contains four *RAN* genes that code for almost identical proteins, which are believed to be functionally redundant. In fact the *ran1* mutant has not shown any phenotype in the plant regarding development, growth or viability, suggesting that the remaining pool of RAN protein is sufficient to sustain the plant survival. However, *ran1* plants show enhanced resistance to the virus infection indicating that each of these RAN proteins might also have functional specificity. The hypersusceptibility of *ran1* to geminivirus infection raised several immediate questions to be addressed in the future: Is this phenotype related to RAN1-TrAP interaction? Can TrAP act as a nuclear importin/exportin? Is TrAP localization regulated by RAN1?

Interestingly, the expression of plant RAN proteins (i.e. tomato, wheat and rice) is able to complement the cell cycle phenotype of the yeast RAN homologue mutant, *pim46*,

further implicating the plant RAN proteins in cell cycle regulation (Ach and Gruissem, 1994). Previous reports in *A. thaliana* have shown that the expression of RAN proteins is highest in the meristematic tissue, and it is coordinated with the expression of Ran Binding Proteins (RanBP), that are also conserved throughout eukaryotes (Haizel et al., 1997). Indeed, other studies overexpressing the heterologous wheat and rice RAN1 in *A. thaliana* have evidenced a prolonged life cycle in the transgenic lines, concomitant with increased primordial tissue, fewer lateral roots, and hypersensitivity to auxin (Liu et al., 2014b; Wang et al., 2006) Taken together, it seems RAN1 protein is a nuclear transport protein involved in the regulation of cell cycle progression in meristematic tissues, where auxin signaling is critical.

Considering the total dependence of geminivirus on the plant machinery for replication, it is not surprising that the infection process is more efficient in replicating cells. Noteworthy, the most severe effects of the disease are seen in the meristems (Carvalho et al., 2008a, 2008b, Gutierrez, 2000a, 2000b). To achieve infection, the geminivirus genome must be transported to the nucleus, just as the virus encoded proteins AC1, AC2, AC3 and BV1, presumably requiring the interaction with a host nuclear transporter to cross the nuclear pore (Jeffrey et al., 1996). Our preliminary data is consistent with a deficient RAN activity in the cell; we envision a situation in which TrAP-RAN interaction is enough to affect auxin response and cell cycle progression in the infected cells. We hypothesize that RAN1 might regulate TrAP shuttling in and out of the nucleus, and therefore is a likely regulator of the TrAP activity during infection.

Alternatively, RAN1 can serve as an effector of TrAP to regulate cell cycle progression and to promote a cellular environment permissive to the virus replication and assembly.

3.4.3 ESP3 modulates virus infection

Transcription from the complementary sense strand of begomovirus genomes is very complex; it consists of multiple overlapping RNAs with different 5'-ends and a co-terminal 3'-end (Bisaro et al., 1990b; Hanley-Bowdoin et al., 1988; Revington et al., 1989; Sunter and Bisaro, 1989; Sunter et al., 1989). All begomoviruses produce polycistronic transcripts, but to date no splicing or processing site in the transcripts has been reported, and the precise regulation of virus gene expression is not yet understood. Studies in TGMV have revealed three transcription start sites on the complementary strand: AC2515, AC1935, and AC1629. AC2515 is a strong promoter with a canonical TATA box that drives the expression of AC1, AC3 and AC4, although the regulation of AC4 is not known. AC1935 does not have a canonical TATA box and only allows the expression of AC3. Finally, AC1629 exhibits a strong promoter activity, it contains a canonical TATA box, and leads the differential expression of both AC2 and AC3 genes, but it produces AC3 up to four-fold more than it does AC2 (Shung et al., 2006). Although the prevailing dogma is that viral transcripts do not undergo splicing processing, the AC4 gene is entirely contained in the AC1 transcript, and AC2/AC3 are expressed alternatively from a single transcript, which suggests the possible requirement of RNA processing.

In yeasts, PRP2 is required for the destabilization of the spliceosome-RNA complex, allowing it to fulfill its catalytic activity (Wlodaver and Staley, 2014). If the splicing activity of ESP3 were required for the proper expression of viral genes, we would

have seen a more resistant phenotype in the *esp3* mutants regarding the virus infection, contrary to the hypersensitive phenotype observed. However, many pleiotropic effects of the *esp3* knockout can still play a role in the pathogenesis. Therefore, the possibility of ESP3 involvement in the processing of CaLCuV transcripts cannot be discarded; specially taking into account that TrAP is a virus transcription factor essential for the expression of several viral genes. Thus, we propose that TrAP might modulate ESP3-mediated processing of viral transcripts. Moreover, TrAP-ESP3 co-localization together with the hyper-susceptibility of the *esp3* mutant to the virus infection is that ESP3 can be involved in the defense mechanism of the plant against the viral threats. Previous reports in plants and animals have found splicing factors to be involved in the RNA-directed DNA Methylation pathway (RdDM) of TGS. Although the precise function of the splicing factors in the TGS is not yet known, it has been proposed that defective RNA processing provides substrates for RNA-dependent RNA Polymerases (RdRP), which synthesize dsRNA that can be a source of siRNAs to direct gene silencing. Also, stalled spliceosomes have been related to enhanced silencing phenotypes. Therefore, we propose a potential role of ESP3 in the TGS pathway and as a target or TrAP to prevent TGS of the viral genome. The proposed ESP3 function could be general and function in the maintenance of genomic integrity, or it could be specific and function as a defense mechanism against geminivirus. In order to discern between these possibilities, further studies on the transcription of otherwise silent transposons should be performed, as well as to determine whether or not ESP3 associates with the viral minichromosome. Importantly, several geminivirus specific siRNAs have been detected during infection, and many are

specifically are generated from the TrAP-dependent AR1 gene transcripts (coat protein) and other regions in the viral genome. Future studies should include the assessment of the production of CaLCuV-derived siRNAs in the *esp3*, Col-0, and ESP3 complemented lines, to determine its effect on the pathogenesis of the begomovirus.

3.5 Materials and Methods

3.5.1 DNA constructions

We used the Gateway system (Invitrogen) (Zhang et al., 2005) for the plant constructs. The destination vectors (containing the destination cassette–DC-) pBA-DC-CFP, pBA-Flag-4Myc-DC, pER10-YFP-DC (Zhang et al., 2006a), pCambia Myc-DC-nLUC (Zhang et al., 2011c), and pER10cLUC-3HA-DC (Castillo-González et al., 2015) were used for transient expression in *N. benthamiana* and for stable *Arabidopsis* transformation. The cDNA or DNA fragments were cloned into pENTR/D vectors, sequencing confirmed; and then transferred to the appropriate destination vectors by recombination using the LR Clonase (Invitrogen).

For expression of recombinant proteins, the cDNA or DNA fragments were cloned into pMAL-DC or pMCSG10 vectors to produce MBP- or His-GST-tagged proteins respectively. The plasmids were further confirmed by sequencing and transformed into *E. coli* BL21 Rossetta DE3 for expression.

3.5.2 Transgenic plants

A. thaliana (Col-0) mutant plants *hog1* (CS1892, and SALK_023915C), *sam2* (SALK_597697C), *ran1* (SALK_138680C), *esp3* (SALK_021856C), and *ref6* (SALK_001018C) mutants were obtained from the Arabidopsis Biological Resource Center and confirmed by genotyping. The overexpression (wild type Col-0 background and driven by a 35S promoter) and complementation (mutant background and driven by the native promoter) constructs were transformed with binary vectors by the floral-dip method (Clough and Bent, 1998; Zhang et al., 2006b). The transgenic seeds were selected on standard MS medium (Murashige and Skoog, 1962) containing the appropriate selective agents: 10 mg/l glufosinate ammonium (Sigma) or 10 mg/l kanamycin (Sigma), together with 100 mg/l carbenicillin (Sigma), for pBA and pER10 constructs, respectively.

3.5.3 Two-step immunoprecipitation for mass spectrometry analysis

FM-TrAP was transiently expressed in true leaves of 3-week-old *N. benthamiana* plants by infiltration with *A. tumefasciens* carrying the vector *XVE-FM-TrAP*. Briefly, the agrobacteria was grown in LB media supplemented with 100 μM Acetosyringone (Sigma) until it reached OD_{600} = 1.0. The bacteria were harvested by centrifugation at 4000 × rcf in 4°C, and resuspended in 10 mM $MgCl_2$ supplemented with 150 μM Acetosyringone and 25 μM ß-estradiol (Sigma). The leaves were harvested 48 hours after inoculation, ground in liquid nitrogen and stored at −80°C until use. Similarly, FM-TrAP expression was induced for 16 hours in 9-day-old wild-type control and *Arabidopsis* transgenic plants

expressing *XVE-FM-TrAP* with 25 µM ß-estradiol in liquid MS media, ground in liquid nitrogen and stored at −80°C until use.

For the assay, total proteins were extracted from 10 g of ground powder in 40 ml (4 volumes) of IP buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 4 mM $MgCl_2$, 50 µM $ZnCl_2$, 0.1% Triton X-100, 1 mM PMSF, 1% glycerol, 3× Roche Complete EDTA-free protease inhibitor, 15 µM MG132); then, the soluble proteins were cleared twice by ultracentrifugation at 20,000 × rcf for 15 min at 4°C. The protein complexes were first immunoprecipitated using 500 µl of Anti-FLAG M2 magnetic beads (Sigma, Cat# M8823) and incubated in slow rotation for 2 hr at 4°C, the nonspecific-bound proteins were removed by three consecutive washes with 15 ml of IP buffer for 10 min incubation each at 4°C. The protein complexes were then eluted by competition with 100 mg/ml FLAG peptide and subsequently immunoprecipitated with 100 µl Anti-c-Myc-agarose affinity gel (Sigma–Aldrich #A7470) at 4°C for 1.5 hr, the nonspecific-bound proteins were removed by five consecutive washes with the IP buffer with 5 min incubation each at 4°C. The protein complexes were eluted in 200 µl of elution buffer (5 mM EDTA, 200 mM NH_4OH) for 20 min. The supernatant was collected, frozen in liquid nitrogen and dried using the Savant SpeedVac concentrator; finally, the sample was solubilized in 30 µl of 2× SDS-loading buffer and run in 10% SDS-PAGE. The samples were run to one-third of the gel, stained with Coomassie blue and collected by excising the whole lane for mass spectrometry analysis in the Taplin Mass Spectrometry Facility at Harvard Medical School.

3.5.4 Expression and purification of recombinant proteins

For *in vitro* pull down assay, the prey (His-GST-CaLCuV-TrAP, His-GST-TGMV-TrAP, and His-GST) proteins were purified by IMAC using the lysis buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 2 mM PMSF), incubated with the Ni-NTA resin (at 4°C for 1 hr, eluted with 300 mM imidazole and immediately dialyzed in storage buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM 2-mercaptoethanol, 50% Glycerol) at 4°C overnight. On the other hand, the bait (MBP-Bait or MBP) proteins were purified by affinity chromatography using the lysis buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM EDTA, 2 mM PMSF, 10 mM 2-mercaptoethanol), incubated with the amylose resin (at 4°C for 2 hr, eluted with 10 mM maltose and immediately dialyzed in storage buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM 2-mercaptoethanol, 50% Glycerol) at 4°C overnight.

3.5.5 In vitro pull down and Co-IP assays

In vitro pull-down assays and *in vivo* Co-IP were done exactly as described (Castillo-González et al., 2015).

3.5.6 Luciferase complementation imaging (LCI) assay

The LCI was performed on 4-week-old *N. benthamiana* leaves infiltrated with various combinations of *A. tumefaciens* GV3101 harboring pCambia Myc-TrAP-nLUC or pCambia Myc-nLUC and *A. tumefaciens* ABI carrying pER10cLUC-3HA or pER10cLUC-3HA-candidate proteins. The agrobacteria containing the pER10 plasmids were incubated with 25µM beta-estradiol for 3 hr prior infiltration, and all the cultures were adjusted to OD_{600} = 0.8. The transfected leaves were assayed 2 days after

agroinfiltration by adding the substrate (10 mM luciferin). The sprayed leaves were incubated in total darkness for 5 min and photographed using an electron multiplying charge-coupled device (EMCCD) camera, Cascade II 512, from Photomerics (Roper Scientific). The images were processed with WinView32 Ver 2.5.19.7 (Roper Scientific).

3.5.7 Confocal microscopy

Leaves of 4-week-old *N. benthamiana* plants were agroinfiltrated with syringe without needle as previously described (Zhang et al., 2005) with *A. tumefasciens* ABI carrying pBA-TrAP-CFP and pER10-YFP-Test protein. The agrobacteria containing the pER10 plasmids were incubated with 25 µM beta-estradiol for 3 hr prior infiltration, and all the cultures were adjusted to OD_{600} = 0.8. The plants were maintained for 2 days at 24°C (16 hr light/8 hr dark). The co-localization was evaluated using a Nikon inverted microscope Eclipse T*i*-E. CFP signal was measured by excitation with Shutter 10-3 filter 3 (CFPHQ [Ex]), and emission was detected at 485 nm; YFP used Shutter 10-3 filter 4 (YFPHQ [Ex]) and emission was detected at 540 nm. The images were processed using NIS-Elements-AR 4.30.01 (Nikon) and Adobe Photoshop software.

3.5.8 CaLCuV pathogenesis assays

The virus infection assays were performed as in (Castillo-González et al., 2015). Briefly, 3-week-old plants grown in short day conditions (8 hr light/16 hr dark) were infected by agroinfiltration of the CaLCuV infective clones of DNAA and DNAB genomic particles, pNSB1090 and pNSB1091 respectively (Arguello-Astorga et al., 2007). For this, the agrobacteria strains were grown separately to O.D.600 = 1.0 in LB media supplemented with 100 µM Acetosyringone. The bacteria were harvested by

centrifugation at 4000 rcf in 4^{o}C, and resuspended in 10 mM $MgCl_2$ supplemented with 150 μM Acetosyringone. The resuspended bacteria were incubated for three hours in room temperature and, prior to inoculation, were mixed in a 3:1 (DNA A to DNA B) ratio to a final O.D.600 = 1.0. The mixture was combined with super fine silicon carbide powder (600 grit, Alfa Aesar) to a final concentration of 1% (v/v), and sprayed on the seedlings using an airbrush at an output pressure of 80-90 psi. Col-0 wild type, mutants, and *Flag-4Myc*- complementation lines were all inoculated at the same time and maintained under the same conditions. The progression of the disease was evaluated by daily observation to determine the offset of viral infection. Two indicators were taken into consideration, the time of symptom development and the severity of the observed symptoms. The assays were replicated at least 3 times on 32-36 plants of each genotype per assay. The data was analyzed using Microsoft Excel. The significance of the observed phenotypic differences among genotypes was determined using Student's T-test.

4. CONCLUSIONS AND FUTURE WORK

Geminivirus are important plant pathogens that wreak havoc in agriculture every year. Their small genomes (2-3kb) allow for very limited coding capacity, producing multifunctional proteins able to simultaneously interfere with many host cellular processes and take over the host. We studied the molecular function of the geminivirus silencing suppressor TrAP. TrAP is involved in the transcriptional activation of viral genes and it was one of the first viral proteins shown to interfere with the TGS pathway. Since TrAP does not bind to DNA in a sequence-specific manner, we hypothesized that its functions are mediated by the interaction with host proteins. In order to identify such TrAP-interacting proteins (TrIPs), we expressed TrAP in the host plants *Nicotiana benthamiana* and *Arabidopsis thaliana* to isolate TrAP-complexes and analyze their composition by mass spectrometry. Subsequently, we confirmed the interactions *in vivo* and *in vitro*. In order to assess the significance of these interactions for viral pathogenesis, we obtained *A. thaliana* TrIP-mutants and produced TrIP-complemented and TrIP-overexpression lines to be used in geminivirus pathogenesis assays. In parallel, we produced a series of *A. thaliana* TrAP-transgenic lines that enabled us to study TrAP function in isolation from other viral components. Notably, our collection of TrAP-transgenic lines comprises a range of expression profiles as a result of using different promoters, such as 35S (constitutive, strong) and XVE (inducible, tunable), which are also useful resources for the scientific community.

The TrAP-mediated TGS suppression has been attributed to pleiotropic effects of TrAP interaction with proteins in the methyl cycle, specifically ADK2 and SAMDC1

(Moffatt et al. 2002; Zhang et al. 2011). Our proteomics data expanded on this model, as we recovered several proteins involved in the methyl cycle (Figure 16). We confirmed the *in vivo* interaction of TrAP with the S-adenosyl methionine synthetase SAM2 and provide evidence of the relevance of SAM2 function in plant defense, as *sam2* null mutants are hypersensitive to virus infection (Figure 39). Moreover, we recovered the S-adenosyl homocysteine hydrolase, SAHH1, from the proteomics data of TrAP complexes obtained from two different host plants. We further confirmed the *in vivo* and *in vitro* interaction of TrAP with SAHH1/HOG1 and showed the physiological relevance of SAHH1/HOG1 function during pathogenesis (Figure 38). *sahh1/hog1* null mutants were hypersensitive to virus infection and plants overexpressing SAHH1/HOG1 were more resistant to virus infection than the wild type (Figure 38). Our findings not only provide strong evidence for the effect of TrAP in the regulation of the methyl cycle, but also serve as the first evidence of a viral suppressor directly interfering with the methyl cycle in the host plant (Figure 16). Furthermore, SAMe is also the cellular precursor of ethylene and polyamines. Both of which are essential metabolites in the plant stress response (Liu et al. 2015). As such, the interference of geminivirus with the SAMe cycle might have other implications not yet pursued. Our transcriptome data from TrAP transgenic plant (Supplementary File 1) shows only one differentially expressed gene related to the ethylene pathway, the ethylene-responsive element binding factor 15 (ERF15). Not only is this transcript downregulated in TrAP transgenic plants, but it has also been directly implicated in the regulation of the immune response against bacterial and fungal pathogens (Zhang et al. 2015). It is worth mentioning that a major disadvantage of using our transcriptome data

set as the only reference for deregulation of the ethylene pathway is that we used 7-days old seedlings, which are at a developmental stage where ethylene biosynthesis is marginal. Furthermore, our materials are constitutively expressing TrAP from the strong promoter 35S, and may not represent the natural conditions in which TrAP would be expressed during virus infection. Future work should consider the effects of TrAP interactions with SAM2 and SAHH1 on their cellular function, taking into consideration the plant organ and the developmental stage.

Notably, although the evidence for TrAP interference with the methyl cycle is substantial, none of the null mutants in the identified TrIPs show a phenotype resembling that of the TrAP transgenic plants. Moreover, TrAP cellular localization is mostly nuclear, while the methyl cycle occurs in the cytosol. These observations suggest the involvement of TrAP in other cellular processes. Indeed, our proteomics approach provided us with multiple hints for the discovery of such processes. We recovered and confirmed the *in vivo* and *in vitro* interaction of TrAP with the nuclear transport factor RAN 1 (Figure 40) and provide evidence of the requirement of RAN1 function for pathogenesis, as null mutants are more resistant to virus infection than the wild type plants (Figure 41). This finding is striking because it points to specialization within the RAN gene family and it implies that TrAP moves through the nuclear pore as a large molecular complex. An initial hypothesis for the relevance of this interaction is that TrAP localization regulated by RAN1, since active transport across the nuclear pore derives energy from the hydrolysis of GTP by RAN. However, the RAN (GTPase) interacts with a transportin bound to a cargo rather than to the cargo itself. Could TrAP act as a nuclear importin/exportin? If so, what does it

transport? An alternative hypothesis is that the requirement of RAN for pathogenesis is due to an indirect effect. In this regard, RAN has been implicated in the regulation of the cell cycle in yeast and plants (Ach and Gruissem, 1994; Liu et al., 2014b; Wang et al., 2006). RAN expression is highest at the meristems, where it regulates life cycle length and sensitivity to auxin. Geminivirus infection depends entirely on the cellular replication machinery; as such, viral infection is highest in meristematic tissues. Gene ontology analysis of the differentially expressed genes in the TrAP-transgenic plants showed significant deregulation of auxin-responsive genes (Supplementary file 1, Figure 19D), suggesting a possible insensitivity to the plant hormone. This observation is consistent with loss of RAN function, but not necessarily with the resistant phenotype of the *ran1* mutant plants when challenged with the geminivirus pathogen. Whether RAN1 modulates TrAP localization and activity during infection, or TrAP promotes a permissive cellular environment through its interaction with RAN1 are critical questions that remain open for future research. To our knowledge, no nuclear transporter has been directly related to the TGS regulation. Since TrAP is a TGS suppressor (Figure 19), this is a possibility that should be explored. Moreover, can TrAP directly inhibit TGS in the nucleus?

The highest peptide recovery for a potential TrIP in our proteomics analysis was from the histone methyltransferase KYP (Figure 21, Table 3). KYP is an essential TGS effector that catalyzes the deposition of the repressive H3K9me2 mark. H3K9me2 marks are recognized by the DNA methyltransferases CMT2 and CMT3 to direct DNA methylation in non-CG sequence contexts. KYP further reinforces gene silencing by recognizing non-CG DNA methylation and catalyzing the H3K9 dimethylation on the

associated nucleosomes. We confirmed the direct interaction between TrAP and KYP *in vivo* and *in vitro*, and we mapped the interacting region to the catalytic SET domain in KYP (Figure 20, Figure 23). Moreover, we demonstrated that TrAP inhibits the histone methyltransferase activity of KYP *in vitro* (Figure 25), while TrAP expression is consistent with the reduction of H3K9me2 marks (Figure 28) and transcriptional activation of KYP-repressed loci (Figure 19) *in vivo*. Since KYP activity promotes DNA methylation in non-CG contexts, we studied the genome methylation status of the TrAP-transgenic plants by whole genome bisulfite sequencing, and we found that TrAP expression is consistent with a reduction in CHH methylation at KYP-regulated gene-rich regions (Figure 29). We further studied the role of KYP during geminivirus pathogenesis and observed that *kyp* null mutants are hypersusceptible to viral infection, whereas overexpression of KYP rendered the plants more resistant than the wild type to the viral threat (Figure 31). Moreover, the geminivirus minichromosome is directly bound and methylated by KYP (Figure 31). We explored the possibility of KYP being the ultimate target of TrAP in the host cell, so we produced a mutant geminivirus lacking a functional TrAP protein and challenged wild type and *kyp* null mutant plants with it, We found that systemic infection was only permissible in *kyp* mutants, although virus accumulation was very low (Figure 33). Our findings demonstrate the role of KYP in viral defense as an effector of epigenetic silencing on the viral chromatin. This implies that TrAP inhibition of KYP is a new viral counter-defense mechanism to enable infection. Thus, direct inhibition of KYP represents a novel counter-defense mechanism for virus survival in the

hosts (Figure 35). To the extent of our knowledge, this is the first account for direct suppression of a histone methyltransferase in the epigenetic pathway by a viral protein.

Direct interference with the TGS pathway by TrAP in order to allow viral gene expression could explain the transcriptional activity of TrAP and its essential role in the expression of viral genes (i.e. coat and nuclear shuttle proteins (Hanley-Bowdoin et al., 2013)). In other words, could TrAP inhibition of KYP be the mechanism of TrAP-dependent transcriptional activation of viral genes? Previous studies have mapped the transcriptional activation domain of TrAP to the acidic region at its C-terminus; preliminary data have shown that neither the C-terminus nor the transactivation domains of TrAP are involved in the interaction with KYP (data not shown), instead the basic region at the N-terminus seems to be essential for interaction. This suggests a possible biochemical separation of the two functions in TrAP. However, detailed assessment of viral gene expression in the *trap-kyp* pathosystem is necessary.

We showed that TrAP directly interact with the SET domain of KYP; however, this domain is highly conserved throughout eukaryotes. In fact, Arabidopsis codes for more than 31 SET proteins (Pontivianne et al. 2010) with histone methyltransferase activity. Hence, it is relevant to address the potential promiscuity of TrAP as an inhibitor of multiple SET domain HMTases. Indeed, we have shown direct *in vitro* interaction between TrAP and three other SET proteins (Figure 24), but no enzymatic assays have been conducted, making it very difficult to determine whether KYP is the main or only SET target of TrAP in the cells (Castillo-González et al. 2015).

On the other hand, since our findings can be easily translated to other organisms, it is of great interest to understand the inhibitory mechanism of TrAP. This can provide the scientific community with useful tools for regulating gene expression, for example by using TrAP as an inhibitor of H3K9 methylation, or targeting it to specific heterochromatic regions. One particular use could be targeting of TrAP to surreptitious pathogens, such as latent HIV, to induce their expression and enable treatment. To this end, we must determine the TrAP structural requirements for interaction with KYP. Since TrAP is largely unstructured, we hypothesize that it can serve as a scaffold for the formation of protein complexes in diverse cellular environments, with a highly dynamic tertiary structure.

Notably, the phenotype of TrAP transgenic plants does not resemble that of *kyp* mutant or *kyp suvh5 suvh6* triple mutant plants, suggesting that KYP inhibition is not the entire function of TrAP. Particularly challenging is the explanation of the loss of H3K27 tri-methylation and early flowering in the TrAP transgenic lines, both of which have been related to functions of complexes other than those formed by KYP and its paralogs in Arabidopsis (i.e. PRC2 and LHP1 complexes). These phenotypic differences hint to a broader function of TrAP, which might include interference with multiple TGS pathways; this potential promiscuity awaits further investigation. In this regard, we noticed that TrAP-overexpression plants are morphologically similar to *lhp1* null mutants, which is a protein characteristically associated with the H3K27me3 pathway. In fact, our transcriptome data evidenced the highly significant overlap between differentially expressed genes in TrAP-transgenic plants and *lhp1* null mutants (Figure 19). This finding

genetically involved TrAP in the LHP1 pathway; alas, the mechanism is still elusive. We did not recover LHP1 from the proteomics data, but other experimental approaches (i.e. coimmunoprecipitation, luciferase complementation imaging, and co-localization by confocal microscopy) provided some evidence for TrAP-LHP1 interaction *in vivo* (Figure 20). However, we failed to demonstrate TrAP-LHP1 interaction *in vitro* (Figure 22), suggesting that the two proteins may not directly interact *in vivo*. We did not investigate a direct function of TrAP as a regulator of H3K27me3 HMTases, but it would not be surprising that TrAP inhibits both pathways. Notably, when we performed chromatin immunoprecipitation (CHIP) assays on the TrAP-transgenic lines, besides the loss of H3K9me2 marks, we observed a severe decrease in H3K27me3 similarly to the molecular phenotype of the *lhp1* mutant (Figure 28). While KYP and H3K9me2 marks are mainly localized in transcriptionally silent heterochromatic regions, LHP1 and H3K27me3 marks are mostly found in euchromatic regions (Black et al., 2012; Du et al., 2014; Liu et al., 2010, 2014d; Turck et al., 2007; Zhang et al., 2007a, 2007b). Interestingly, our CHIP data did not support any effect of TrAP expression on marks associated to transcriptionally active euchromatin (i.e. H3K4me3 and H3K36me3, (Figure 23), implying a specific mechanism for TrAP-mediated HMTase inhibition.

In the same genetic pathway of *CLF* and *LHP1* genes is a gene called Relative of Early Flowering phenotype 6 (REF6), REF6 encoded a Jumonji domain involved in H3K27 demethylation. Strikingly, we observed phenotypic similarity between REF6 and TrAP overexpression plants in *A. thaliana*, suggesting that they might have genetic or even biochemical interaction. In line with this hypothesis, the two proteins co-localized in

nuclear speckles (Figure 44). Further biochemical assays to validate the interaction between REF6 and TrAP are in progress.

TrAP has been shown to shuttle in and out of the nucleus, to interfere directly and indirectly with the TGS pathway and to regulate gene expression and defense response in the host, a critical question to address is what is the size of TrAP pool in the cells during infection? How much TrAP is required to really deplete H3K9me2 from the cells, or to inhibit ADK2, or to activate transcription of the viral genes AR1 and BR1? How much TrAP is necessary to shift the balance towards infection? The use of constitutive promoters and truncated TrAP proteins introduces inherent caveats in the experimental design. To better assess the physiological relevance of our findings, we developed a model of transgenic Arabidopsis plants that can express FM-TrAP from a gradual inducible promoter, which can serve in the determination of how much TrAP is necessary to attain the different phenotypes reported; equally important is the evaluation of TrAP stability through the measurement of protein decay after translation inhibition with cyclohexamide. Other related questions pertain local concentration of TrAP, and how is the expression of viral TrAP regulated?

In conclusion, our research provides ample evidence for the interference with the TGS pathway by the geminiviral TrAP protein. On one side, TrAP interfering with the methyl cycle through its interactions with SAHH1 and SAM2, indirectly preventing gene silencing. We then demonstrated that TrAP interacts with the nuclear transporter RAN1 and that its function positively regulates virus infection, suggesting that TrAP cellular localization is essential for its function. And finally, we identified the nuclear TGS effector

KYP as a direct target of TrAP in the nucleus, by which TrAP directly abrogates epigenetic silencing. As such, we demonstrate that KYP plays a critical role in the immune response to invading nucleic acids in plants. Interestingly, the human homolog of KYP, SUVH39H1 has been reported to be required for the maintenance of latency in HIV and Epsteinn-Barr virus infections (du Chéné et al., 2007; Imai et al., 2014). This suggests that SUVH39H1 fulfills a similar function in human pathogenesis as the one we identified for KYP in plants. In sum, KYP inhibition is an effective tactic to abrogate TGS. We predict that TGS suppression is a widespread counter-defense strategy, and that inhibition of KYP is only the tip of the iceberg. Furthemore, TrAP inhibition of histone lysine methyltransferases and its direct effect on DNA methylation gene expression may serve as a stepping stone for the development of epigenetic therapeutics addressing both human health and agricultural productivity.

REFERENCES

Ach, R.A., and Gruissem, W. (1994). A small nuclear GTP-binding protein from tomato suppresses a Schizosaccharomyces pombe cell-cycle mutant (R/spil/TC4/plant cell cyde/Lycopersicon escukntum). Proc. Natl. Acad. Sci. USA *91*, 5863–5867.

Ahlquist, P. (2002). RNA-Dependent RNA Polymerases, Viruses, and RNA Silencing. Science (80-.). *296*.

Aida, M., Ishida, T., and Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. Development *126*, 1563–1570.

Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell *121*, 207–221.

Alvarez, M.E., Nota, F., and Cambiagno, D.A. (2010). Epigenetic control of plant immunity. Mol. Plant Pathol. *11*, 563–576.

Aregger, M., Borah, B.K., Seguin, J., Rajeswaran, R., Gubaeva, E.G., Zvereva, A.S., Windels, D., Vazquez, F., Blevins, T., Farinelli, L., et al. (2012). Primary and Secondary siRNAs in Geminivirus-induced Gene Silencing. PLoS Pathog. *8*.

Arguello-Astorga, G., Ascencio-Ibáñez, J.T., Dallas, M.B., Orozco, B.M., and Hanley-Bowdoin, L. (2007). High-frequency reversion of geminivirus replication protein mutants during infection. J. Virol. *81*, 11005–11015.

Argüello-Astorga, G.R., Guevara-González, R.G., Herrera-Estrella, L.R., and Rivera-Bustamante, R.F. (1994). Geminivirus Replication Origins Have a Group-Specific Organization of Iterative Elements: A Model for Replication. Virology *203*, 90–100.

Von Arnim, A., and Stanley, J. (1992). Determinants of tomato golden mosaic virus symptom development located on DNA B. Virology *186*, 286–293.

Aukerman, M.J., and Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2 -Like Target Genes. Plant Cell *15*, 2730–2741.

Axtell, M.J. (2013). Classification and Comparison of Small RNAs from Plants. Annu. Rev. Plant Biol. *64*, 137–159.

Azevedo, J., Garcia, D., Pontier, D., Ohnesorge, S., Yu, A., Garcia, S., Braun, L., Bergdoll,

M., Hakimi, M.A., Lagrange, T., et al. (2010). Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev. *24*, 904–915.

Azizi, P., Rafii, M.Y., Maziah, M., Abdullah, S.N.A., Hanafi, M.M., Latif, M.A., Rashid, A.A., and Sahebi, M. (2015). Understanding the shoot apical meristem regulation: A study of the phytohormones, auxin and cytokinin, in rice. Mech. Dev. *135*, 1–15.

Baliji, S., Sunter, J., and Sunter, G. (2007). Transcriptional analysis of complementary sense genes in Spinach curly top virus and functional role of C2 in pathogenesis. Mol. Plant. Microbe. Interact. *20*, 194–206.

Baliji, S., Lacatus, G., and Sunter, G. (2010). The Interaction between geminivirus pathogenicity proteins and adenosine kinase leads to increased expression of primary cytokinin-responsive genes. Virology *402*, 238–247.

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

Baumberger, N., and Baulcombe, D.C. (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc. Natl. Acad. Sci. U. S. A. *102*, 11928–11933.

Baumberger, N., Tsai, C.H., Lie, M., Havecker, E., and Baulcombe, D.C. (2007). The Polerovirus Silencing Suppressor P0 Targets ARGONAUTE Proteins for Degradation. Curr. Biol. *17*, 1609–1614.

Bazin, J., Khan, G.A., Combier, J., Bustos-Sanmamed, P., Debernardi, J.M., Palatnik, J., Hartmann, C., Crespi, M., Rodriguez, R., Diderot, P., et al. (2013). miR396 affects mycorrhization and root meristem activity in the legume Medicago truncatula. Plant J. *74*, 920–934.

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

Berger, M.R., and Sunter, G. (2013). Identification of sequences required for AL2-mediated activation of the tomato golden mosaic virus-yellow vein BR1 promoter. J. Gen. Virol. *94*, 1398–1406.

Binda, O. (2013). On your histone mark, SET, methylate! Epigenetics *8*, 457–463.

Bisaro, D.M. (2006). Silencing suppression by geminivirus proteins. Virology *344*, 158–168.

Bisaro, D.M., Sunter, G., Revington, G.N., Brough, C.L., Hormuzdi, S.G., and Hartitz, M.

(1990a). Molecular Genetics of Tomato Golden Mosaic Virus Replication: Progress Toward Defining Gene Functions, Transcription Units and the Origin of DNA Replication. In Viral Genes and Plant Pathogenesis, (New York, NY: Springer New York), pp. 89–105.

Bisaro, D.M., Sunter, G., Revington, G.N., Brough, C.L., Hormuzdi, S.G., and Hartitz, M. (1990b). Molecular Genetics of Tomato Golden Mosaic Virus Replication: Progress Toward Defining Gene Functions, Transcription Units and the Origin of DNA Replication. In Viral Genes and Plant Pathogenesis, (New York, NY: Springer New York), pp. 89–105.

Bischoff, F.R., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature *354*, 80–82.

Black, J.C., and Whetstine, J.R. (2011). Chromatin landscape: methylation beyond transcription. Epigenetics *6*, 9–15.

Black, J.C., Van Rechem, C., and Whetstine, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol. Cell *48*, 491–507.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Banning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. EMBO J. *17*, 170–180.

Bolduc, N., Yilmaz, A., Mejia-Guerra, M.K., Morohashi, K., O'Connor, D., Grotewold, E., and Hake, S. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. Genes Dev. *26*, 1685–1690.

Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu. Rev. Plant Biol. *65*, 473–503.

Borges, F., and Martienssen, R.A. (2015). The expanding world of small RNAs in plants. Nat. Publ. Gr. *16*, 1–15.

Borges, F., and Martienssen, R. a (2013). Establishing epigenetic variation during genome reprogramming. RNA Biol. *10*, 490–494.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.-K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell *123*, 1279–1291.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R. (2000). Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by

CLV3 activity. Science *289*, 617–619.

Briddon, R.W., and Markham, P.G. (2001). Complementation of bipartite begomovirus movement functions by topocuviruses and curtoviruses. Arch. Virol. *146*, 1811–1819.

Briddon, R.W., Bedford, I.D., Tsai, J.H., and Markham, P.G. (1996). Analysis of the Nucleotide Sequence of the Treehopper-Transmitted Geminivirus, Tomato Pseudo-Curly Top Virus, Suggests a Recombinant Origin. Virology *219*, 387–394.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. Science (80-.). *320*, 1185–1190.

Brough, C.L., Gardiner, W.E., Inamdar, N.M., Zhang, X.-Y., Ehrlich, M., and Bisaro, D.M. (1992). DNA methylation inhibits propagation of tomato golden mosaic virus DNA in transfected protoplasts. Plant Mol. Biol. *18*, 703–712.

Buchmann, R.C., Asad, S., Wolf, J.N., Mohannath, G., and Bisaro, D.M. (2009). Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. J. Virol. *83*, 5005–5013.

Bustamante, M., Matus, J.T., and Riechmann, J.L. (2016). Genome-wide analyses for dissecting gene regulatory networks in the shoot apical meristem. J. Exp. Bot. *67*, 1639–1648.

Carrillo-Tripp, J., Shimada-Beltr??n, H., and Rivera-Bustamante, R. (2006). Use of geminiviral vectors for functional genomics. Curr. Opin. Plant Biol. *9*, 209–215.

Cartolano, M., Castillo, R., Efremova, N., Kuckenberg, M., Zethof, J., Gerats, T., Schwarz-Sommer, Z., and Vandenbussche, M. (2007). A conserved microRNA module exerts homeotic control over Petunia hybrida and Antirrhinum majus floral organ identity. Nat. Genet. *39*, 901–905.

Carvalho, M.F., and Lazarowitz, S.G. (2004). Interaction of the Movement Protein NSP and the Arabidopsis acetyltransferase AtNSI is necessary for Cabbage Leaf Curl geminivirus infection and pathogenicity. J. Virol. *78*, 11161–11171.

Carvalho, C.M., Santos, A.A., Pires, S.R., Rocha, C.S., Saraiva, D.I., Machado, J.P.B., Mattos, E.C., Fietto, L.G., and Fontes, E.P.B. (2008a). Regulated nuclear trafficking of rpL10A mediated by NIK1 represents a defense strategy of plant cells against virus. PLoS Pathog. *4*.

Carvalho, C.M., Machado, J.P.B., Zerbini, F.M., and Fontes, E.P.B. (2008b). NSP-interacting GTPase. A cytosolic protein as cofactor for nuclear shuttle proteins.

Plant Signal. Behav. *3*, 752–754.

Castillo-González, C., Liu, X., Huang, C., Zhao, C., Ma, Z., Hu, T., Sun, F., Zhou, Y., Zhou, X., Wang, X.-J.J., et al. (2015). Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. Elife *4*, 1–30.

Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E. (2004). RNA Silencing Genes Control de novo DNA Methylation. Science. *303*, 1336.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.-H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development *131*, 5263–5276.

Chaturvedi, S., Kalantidis, K., and Rao, A.L.N. (2014). A Bromodomain-Containing Host Protein Mediates the Nuclear Importation of a Satellite RNA of Cucumber Mosaic Virus. Jounal Virol. *88*, 1890–1896.

Chen, X. (2004). A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development. Science (80-.). *303*, 2022–2025.

Chen, X. (2005). MicroRNA biogenesis and function in plants. FEBS Lett. *579*, 5923–5931.

Chen, H.-M., Chen, L.-T., Patel, K., Li, Y.-H., Baulcombe, D.C., and Wu, S.-H. (2010). 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc. Natl. Acad. Sci. *107*, 15269–15274.

du Chéné, I., Basyuk, E., Lin, Y.-L., Triboulet, R., Knezevich, A., Chable-Bessia, C., Mettling, C., Baillat, V., Reynes, J., Corbeau, P., et al. (2007). Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. EMBO J. *26*, 424–435.

Chinnusamy, V., and Zhu, J.K. (2009). Epigenetic regulation of stress responses in plants. Curr. Opin. Plant Biol. *12*, 133–139.

Chiu, M.H., Chen, I.H., Baulcombe, D.C., and Tsai, C.H. (2010). The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway. Mol. Plant Pathol. *11*, 641–649.

Chuck, G., Cigan, a M., Saeteurn, K., and Hake, S. (2007). The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. Nat. Genet. *39*, 544–549.

Chung, H.Y., and Sunter, G. (2014). Interaction between the transcription factor AtTIFY4B and begomovirus AL2 protein impacts pathogenicity. Plant Mol. Biol. *86*, 185–200.

Chung, H.Y., Lacatus, G., and Sunter, G. (2014). Geminivirus AL2 protein induces expression of, and interacts with, a calmodulin-like gene, an endogenous regulator of gene silencing. Virology *460–461*, 108–118.

Clavel, M., Pélissier, T., Montavon, T., Tschopp, M.A., Pouch-Pélissier, M.N., Descombin, J., Jean, V., Dunoyer, P., Bousquet-Antonelli, C., and Deragon, J.M. (2016). Evolutionary history of double-stranded RNA binding proteins in plants: identification of new cofactors involved in easiRNA biogenesis. Plant Mol. Biol. *91*, 131–147.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. *16*, 735–743.

Combier, J.-P., Frugier, F., de Billy, F., Boualem, A., El-yahyaoui, F., Moreau, S., Vernié, T., Ott, T., Gamas, P., Crespi, M., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. Genes Dev. *20*, 3084–3088.

Conway, L.J., and Poethig, R.S. (1997). Mutations of Arabidopsis thaliana that transform leaves into cotyledons. Proc. Natl. Acad. Sci. U. S. A. *94*, 10209–10214.

Coutavas, E., Ren, M., Oppenheim, J.D., D'Eustachio, P., and Rush, M.G. (1993). Characterization of proteins that interact with the cell-cycle regulatory protein Ran/TC4. Nature *366*, 585–587.

Creasey, K.M., Zhai, J., Borges, F., Van Ex, F., Regulski, M., Meyers, B.C., and Martienssen, R. a (2014). miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. Nature *508*, 411–415.

Csorba, T., Lózsa, R., Hutvágner, G., and Burgyán, J. (2010). Polerovirus protein P0 prevents the assembly of small RNA-containing RISC complexes and leads to degradation of ARGONAUTE1. Plant J. *62*, 463–472.

Csorba, T., Kontra, L., and Burgyán, J. (2015). Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. Virology *479–480*, 85–103.

Cui, X., and Cao, X. (2014). Epigenetic regulation and functional exaptation of transposable elements in higher plants. Curr. Opin. Plant Biol. *21*, 83–88.

Daelemans, D., Costes, S. V, Cho, E.H., Erwin-Cohen, R.A., Lockett, S., and Pavlakis, G.N. (2004). *In vivo* HIV-1 Rev multimerization in the nucleolus and cytoplasm

identified by fluorescence resonance energy transfer. J. Biol. Chem. *279*, 50167–50175.

Dalmay, T., Horse®eld, R., Hartig Braunstein, T., and Baulcombe, D.C. (2001). SDE3 encodes an RNA helicase required for posttranscriptional gene silencing in ArabidopsiS. EMBO J. *20*, 2069–2077.

Deal, R.B., and Henikoff, S. (2011). Histone variants and modifications in plant gene regulation. Curr. Opin. Plant Biol. *14*, 116–122.

Delaux, P.-M., Guillaume, B., and Combier, J.-P. (2013). Rapid report NSP1 is a component of the Myc signaling pathway. New Phytol. *199*, 59–65.

Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K.D., Carrington, J.C., and Voinnet, O. (2006). Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science (80-.). *313*, 68–71.

Ding, S.-W. (2010). RNA-based antiviral immunity. Nat. Rev. Immunol. *10*, 632–644.

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

Doucet, C.M., and Hetzer, M.W. (2019). Nuclear pore biogenesis into an intact nuclear envelope. Chromosoma *119*, 469–477.

Douglas, R.N., Wiley, D., Sarkar, A., Springer, N., Timmermans, M.C.P., and Scanlon, M.J. (2010). ragged seedling2 Encodes an ARGONAUTE7-like protein required for mediolateral expansion, but not dorsiventrality, of maize leaves. Plant Cell *22*, 1441–1451.

Du, J., Zhong, X., Bernatavichute, Y. V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., et al. (2012). Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell *151*, 167–180.

Du, J., Johnson, L.M., Groth, M., Feng, S., Hale, C.J., Li, S., Vashisht, A.A., Gallego-Bartolome, J., Wohlschlegel, J.A., Patel, D.J., et al. (2014). Mechanism of DNA methylation-directed histone methylation by KRYPTONITE. Mol. Cell *55*, 495–504.

Du, J., Johnson, L.M., Jacobsen, S.E., and Patel, D.J. (2015). DNA methylation pathways and their crosstalk with histone methylation. Nat. Rev. Mol. Cell Biol. *16*, 519–532.

Dubin, M.J., Zhang, P., Meng, D., Remigereau, M.-S., Osborne, E.J., Paolo Casale, F.,

Drewe, P., Kahles, A., Jean, G., Vilhjálmsson, B., et al. (2015). DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. Elife *4*, e05255.

Efroni, I., Blum, E., Goldshmidt, A., and Eshed, Y. (2008). A Protracted and Dynamic Maturation Schedule Underlies Arabidopsis Leaf Development. Plant Cell *20*, 2293–2306.

Efroni, I., Han, S.K., Kim, H.J., Wu, M.F., Steiner, E., Birnbaum, K.D., Hong, J.C., Eshed, Y., and Wagner, D. (2013). Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. Dev. Cell *24*, 438–445.

Ellendorff, U., Fradin, E.F., De Jonge, R., and Thomma, B.P.H.J. (2009). RNA silencing is required for Arabidopsis defence against Verticillium wilt disease. J. Exp. Bot. *60*, 591–602.

Elmer, J.S., Brand, L., Sunter, G., Gardiner, W.E., Bisaro, D.M., and Rogers, S.G. (1988). Genetic Analysis of the tomato golden mosaic virus II. The product of the AL1 coding sequence is required for replication. Nucleic Acids Res. *16*, 4465–4482.

Endo, Y., Iwakawa, H., and Tomari, Y. (2013). Arabidopsis ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. EMBO Rep. *14*, 652–658.

Eschenfeldt, W.H., Lucy, S., Millard, C.S., Joachimiak, A., and Mark, I.D. (2009). A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. pp. 105–115.

Etemadi, M., Gutjahr, C., Couzigou, J., Zouine, M., Lauressergues, D., Timmers, A., Audran, C., Bouzayen, M., Bécard, G., and Combier, J. (2014). Auxin Perception Is Required for Arbuscule Development in Arbuscular Mycorrhizal Symbiosis 1 [W]. Plant Physiol. *166*, 281–292.

Fahlgren, N., Montgomery, T.A., Howell, M.D., Allen, E., Dvorak, S.K., Alexander, A.L., and Carrington, J.C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. Curr. Biol. *16*, 939–944.

Fang, X., and Qi, Y. (2016). RNAi in Plants: An Argonaute-Centered View. Plant Cell *28*, 272–285.

Farrona, S., Coupland, G., and Turck, F. (2008). The impact of chromatin regulation on the floral transition. Semin. Cell Dev. Biol. *19*, 560–573.

Fátyol, K., Ludman, M., and Burgyán, J. (2016). Functional dissection of a plant

Argonaute. Nucleic Acids Res. *44*, 1384–1397.

Fei, Q., Xia, R., and Meyers, B.C. (2013). Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. Plant Cell *25*, 2400–2415.

Feng, S., and Jacobsen, S.E. (2011). Epigenetic modifications in plants: an evolutionary perspective. Curr. Opin. Plant Biol. *14*, 179–186.

Feng, S., Rubbi, L., Jacobsen, S.E., and Pellegrini, M. (2011). Determining DNA Methylation Profiles Using Sequencing. pp. 223–238.

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

Fischer, A., Hofmann, I., Naumann, K., and Reuter, G. (2006). Heterochromatin proteins and the control of heterochromatic gene silencing in Arabidopsis. J. Plant Physiol. *163*, 358–368.

Fondong, V.N. (2013). Geminivirus protein structure and function. Mol. Plant Pathol. *14*, 635–649.

Fouracre, J.P., and Poethig, R.S. (2016). The role of small RNAs in vegetative shoot development. Curr. Opin. Plant Biol. *29*, 64–72.

Franssen, H.J., Xiao, T.T., Kulikova, O., Wan, X., Bisseling, T., and Scheres, B. (2015). Root developmental programs shape the Medicago truncatula nodule meristem. Development *142*, 2941–2950.

Furner, I.J., Sheikh, M.A., and Collett, C.E. (1998). Gene Silencing and Homology-Dependent Gene Silencing in Arabidopsis: Genetic Modifiers and DNA Methylation. Genetics *149*, 651–662.

Galli, M., and Gallavotti, A. (2016). Expanding the Regulatory Network for Meristem Size in Plants. Trends Genet. *32*, 372–383.

Gan, E.-S., Huang, J., and Ito, T. (2013). Functional Roles of Histone Modification, Chromatin Remodeling and MicroRNAs in Arabidopsis Flower Development. In International Review of Cell and Molecular Biology, pp. 115–161.

Gardiner, W.E., Sunter, G., Brand, L., Elmer, J.S., Rogers, S.G., and Bisaro, D.M. (1988). Coat Protein Is Not Required for Systemic Spread or Symptom Development. EMBO J. *7*, 899–904.

Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G., Lefebvre, D., and Grandjean, O.

(2001). Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Development *128*, 4847–4858.

Gobbato, E. (2015). Recent developments in arbuscular mycorrhizal signaling. Curr. Opin. Plant Biol. *26*, 1–7.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat. Rev. Genet. *13*, 343–357.

Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M.A., Tonti-Filippini, J., Chen, H., Millar, A.H., and Ecker, J.R. (2008). A link between RNA metabolism and silencing affecting Arabidopsis development. Dev. Cell *14*, 854–866.

Guo, H., Song, X., Xie, C., Huo, Y., Zhang, F., Chen, X., Geng, Y., and Fang, R. (2013). Rice yellow stunt rhabdovirus protein 6 suppresses systemic RNA silencing by blocking RDR6-mediated secondary siRNA synthesis. Mol. Plant. Microbe. Interact. *26*, 927–936.

Guo, W., Liew, J.Y., and Yuan, Y.A. (2014). Structural insights into the arms race between host and virus along RNA silencing pathways in Arabidopsis thaliana. Biol. Rev. *89*, 337–355.

Gutierrez, C. (2000a). Geminiviruses and the plant cell cycle. Plant Mol. Biol. *43*, 763–772.

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

Haas, G., Azevedo, J., Moissiard, G., Geldreich, A., Himber, C., Bureau, M., Fukuhara, T., Keller, M., and Voinnet, O. (2008). Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. EMBO J. *27*, 2102–2112.

Haizel, T., Merkle, T., Pay, A., Fejes, E., and Nagy, F. (1997). Characterization of proteins that interact with the GTP-bound form of the regulatory GTPase Ran in Arabidopsis. Plant J. *11*, 93–103.

Hajimorad, M.R., Ghabrial, S.A., and Roossinck, M.J. (2009). De novo emergence of a novel satellite RNA of cucumber mosaic virus following serial passages of the virus derived from RNA transcripts. Arch. Virol. *154*, 137–140.

Hanley-Bowdoin, L., Elmer, J.S., and Rogers, S.G. (1988). Nucleic Acids Research Transient expression of heterologous RNAs using tomato golden mosaic virus. Nucleic Acids Res. *16*, 10511–10528.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D., and Mansoor, S. (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. Nat. Rev. Microbiol. *11*, 777–788.

Hao, L., Wang, H., Sunter, G., and Bisaro, D.M. (2003). Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. Plant Cell *15*, 1034–1048.

Hartitz, M.D., Sunter, G., and Bisaro, D.M. (1999). The tomato golden mosaic virus transactivator (TrAP) is a single-stranded DNA and zinc-binding phosphoprotein with an acidic activation domain. Virology *263*, 1–14.

Havecker, E.R., Wallbridge, L.M., Hardcastle, T.J., Bush, M.S., Kelly, K.A., Dunn, R.M., Schwach, F., Doonan, J.H., and Baulcombe, D.C. (2010). The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell *22*, 321–334.

Hayes, R.J., and Buck, K.W. (1989). Replication of tomato golden mosaic virus DNA B in transgenic plants expressing open reading frames (ORFs) of DNA A: requirement of ORF AL2 for production of single-stranded DNA. Nucleic Acids Res. *17*, 10213–10222.

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

Herr, A.J., Molnar, A., Jones, A., and Baulcombe, D.C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proc Natl Acad Sci U S A *103*, 14994–15001.

Højfeldt, J.W., Agger, K., and Helin, K. (2013). Histone lysine demethylases as targets for anticancer therapy. Nat. Rev. Drug Discov. *12*, 917–930.

Hormuzdi, S.G., and Bisaro, D.M. (1993). Genetic analysis of beet curly top virus: evidence for three virion sense genes involved in movement and regulation of single- and double-stranded DNA levels. Virology *193*, 900–909.

Hormuzdi, S.G., and Bisaro, D.M. (1995). BCTV examination of the roles of L2 and L3 genes in viral pathogenesis. Virology 1044–1054.

Huang, W., Pitorre, D., Poretska, O., Marizzi, C., Winter, N., Poppenberger, B., and Sieberer, T. (2015). ALTERED MERISTEM PROGRAM1 suppresses ectopic stem cell niche formation in the shoot apical meristem in a largely cytokinin-independent manner. Plant Physiol. *167*, 1471–1486.

Imai, K., Kamio, N., Cueno, M.E., Saito, Y., Inoue, H., Saito, I., and Ochiai, K. (2014). Role of the histone H3 lysine 9 methyltransferase Suv39 h1 in maintaining

Epsteinn-Barr virus latency in B95-8 cells. FEBS J. *281*, 2148–2158.

Itoh, J.-I., Hibara, K.-I., Sato, Y., and Nagato, Y. (2008). Developmental role and auxin responsiveness of Class III homeodomain leucine zipper gene family members in rice. Plant Physiol. *147*, 1960–1975.

Iwakawa, H. oki, and Tomari, Y. (2013). Molecular insights into microRNA-mediated translational repression in plants. Mol. Cell *52*, 591–601.

Iwata, Y., Takahashi, M., Fedoroff, N. V., and Hamdan, S.M. (2013). Dissecting the interactions of SERRATE with RNA and DICER-LIKE 1 in Arabidopsis microRNA precursor processing. Nucleic Acids Res. *41*, 9129–9140.

Jackel, J.N., Buchmann, R.C., Singhal, U., and Bisaro, D.M. (2014). Analysis of Geminivirus AL2 and L2 Proteins Reveals a Novel AL2 Silencing Suppressor Activity. J. Virol. *89*, 3176–3187.

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

Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P., and Tsiantis, M. (2005). KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. Curr. Biol. *15*, 1560–1565.

Jeffrey, J.L., Pooma, W., and Petty, I.T.D. (1996). Genetic Requirements for Local and Systemic Movement of Tomato Golden Mosaic Virus in Infected Plants. Virology *223*, 208–218.

Ji, L., Yu, Y., Zhai, J., Luscher, E., Gao, L., Liu, C., Le, B., Cao, X., Mo, B., Meyers, B.C., et al. ARGONAUTE10 promotes the degradation of miR165 / 6 through the SDN1 and SDN2 exonucleases in Arabidopsis. Draft Manuscr.

Jones, A.L., and Sung, S. (2014). Mechanisms underlying epigenetic regulation in arabidopsis thaliana. Integr. Comp. Biol. *54*, 61–67.

Juarez, M.T., Kui, J.S., Thomas, J., Heller, B.A., and Timmermans, M.C.P. (2004). microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. Nature *428*, 84–88.

Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J., et al. (2005). Nodulation Signaling in Legumes Requires NSP2, a Member of the GRAS Family of Transcriptional Regulators. Science (80-.). *308*, 1786–1789.

Kammann, M., Schalk, H.J., Matzeit, V., Schaefer, S., Schell, J., and Gronenborn, B. (1991). DNA replication of wheat dwarf virus, a geminivirus, requires two cis-acting signals. Virology *184*, 786–790.

Kenworthy, A.K. (2001). Imaging Protein-Protein Interactions Using Fluorescence Resonance Energy Transfer Microscopy. Methods *24*, 289–296.

Knauer, S., Holt, A.L., Rubio-Somoza, I., Tucker, E.J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M.C., Tucker, M.R., and Laux, T. (2013). A Protodermal miR394 Signal Defines a Region of Stem Cell Competence in the Arabidopsis Shoot Meristem. Dev. Cell *24*, 125–132.

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K. (2003). Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. Plant Cell Physiol. *44*, 555–564.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Koyama, T., Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2010). TCP transcription factors regulate the activities of ASYMMETRIC LEAVES1 and miR164, as well as the auxin response, during differentiation of leaves in Arabidopsis. Plant Cell *22*, 3574–3588.

Kreuze, J.F., Savenkov, E.I., Cuellar, W., Li, X., and Valkonen, J.P.T. (2005). Viral Class 1 RNase III Involved in Suppression of RNA Silencing Viral Class 1 RNase III Involved in Suppression of RNA Silencing. J. Virol. *79*, 7227–7238.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

Kumar, V., Mishra, S.K., Rahman, J., Taneja, J., Sundaresan, G., Mishra, N.S., and Mukherjee, S.K. (2015). Mungbean yellow mosaic Indian virus encoded AC2 protein suppresses RNA silencing by inhibiting Arabidopsis RDR6 and AGO1 activities. Virology *486*, 158–172.

Lacatus, G., and Sunter, G. (2009). The Arabidopsis PEAPOD2 transcription factor interacts with geminivirus AL2 protein and the coat protein promoter. Virology *392*, 196–202.

Laliberté, J.-F., and Sanfaçon, H. (2010). Cellular remodeling during plant virus infection. Annu. Rev. Phytopathol. *48*, 69–91.

Larue, C.T., Wen, J., and Walker, J.C. (2009). A microRNA-transcription factor module regulates lateral organ size and patterning in Arabidopsis. Plant J. *58*, 450–463.

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rätsch, G., and Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. *105*, 8795–8800.

Laufs, P., Peaucelle, A., Morin, H., and Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. Development *131*, 4311–4322.

Lauressergues, D., Delaux, P.-M., Formey, D., Lelandais-Bière, C., Fort, S., Cottaz, S., Bécard, G., Niebel, A., Roux, C., and Combier, J. (2012). The microRNA miR171h modulates arbuscular mycorrhizal colonization of Medicago truncatula by targeting NSP2. Plant J. *72*, 512–522.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.

Leibfried, A., To, J.P.C., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U. (2005). WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. Nature *438*, 1172–1175.

Li, H., Deng, Y., Wu, T., Subramanian, S., and Yu, O. (2010). Misexpression of miR482, miR1512, and miR1515 Increases Soybean Nodulation. Plant Physiol. *153*, 1759–1770.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. Curr. Biol. *15*, 1501–1507.

Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., Ji, L., Pan, Z., Cao, X., Mo, B., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in arabidopsis. Cell *153*, 562–574.

Li, S., Castillo-González, C., Yu, B., and Zhang, X. (2017). The functions of plant small RNAs in development and in stress responses. Plant J. *90*, 654–670.

Li, Z., Li, B., Shen, W.H., Huang, H., and Dong, A. (2012). TCP transcription factors interact with AS2 in the repression of class-I KNOX genes in Arabidopsis thaliana. Plant J. *71*, 99–107.

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. Annu. Rev. Plant Biol. *61*, 395–420.

Liu, J., Rice, J.H., Chen, N., Baum, T.J., and Hewezi, T. (2014a). Synchronization of developmental processes and defense signaling by growth regulating transcription

factors. PLoS One *9*, 1–14.

Liu, J-H., Wang, W., Wu, H., Gong, X., and Moriguchi, T. (2015). Polyamines function in stress tolerance: from synthesis to regulation. Front. Plant Sci. *6*, 8-27.

Liu, P., Qi, M., Wang, Y., Chang, M., Liu, C., Sun, M., Yang, W., and Ren, H. (2014b). Arabidopsis RAN1 mediates seed development through its parental ratio by affecting the onset of endosperm cellularization. Mol. Plant *7*, 1316–1328.

Liu, Q., Yao, X., Pi, L., Wang, H., Cui, X., and Huang, H. (2009). The ARGONAUTE10 gene modulates shoot apical meristem maintenance and establishment of leaf polarity by repressing miR165/166 in Arabidopsis. Plant J. *58*, 27–40.

Liu, W., Yu, W., Hou, L., Wang, X., Zheng, F., Wang, W., Liang, D., Yang, H., Jin, Y., and Xie, X. (2014c). Analysis of miRNAs and their targets during adventitious shoot organogenesis of Acacia crassicarpa. PLoS One *9*.

Liu, Z.W., Shao, C.R., Zhang, C.J., Zhou, J.X., Zhang, S.W., Li, L., Chen, S., Huang, H.W., Cai, T., and He, X.J. (2014d). The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci. PLoS Genet. *10*.

Loach, R., Moffatt, B., and Wagner, C. (2005). The Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1 Gene Codes for an S -Adenosyl- L -Homocysteine Hydrolase Required for DNA Methylation-Dependent Gene Silencing. Tha Plant Cell *17*, 404–417.

Lopez, J.A., Sun, Y., Blair, P.B., and Mukhtar, M.S. (2015). TCP three-way handshake: Linking developmental processes with plant immunity. Trends Plant Sci. *20*, 238–245.

Lozano-Duran, R., Rosas-Diaz, T., Gusmaroli, G., Luna, A.P., Taconnat, L., Deng, X.W., and Bejarano, E.R. (2011). Geminiviruses Subvert Ubiquitination by Altering CSN-Mediated Derubylation of SCF E3 Ligase Complexes and Inhibit Jasmonate Signaling in Arabidopsis thaliana. Plant Cell *23*, 1014–1032.

Lu, S., Sun, Y., Amerson, H., and Chiang, V.L. (2007). MicroRNAs in loblolly pine (Pinus taeda L .) and their association with fusiform rust gall development. Plant J. *51*, 1077–1098.

Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., and Barton, M.K. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. Development *126*, 469–481.

Machida, S., Chen, H.-Y., and Adam Yuan, Y. (2011). Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. Nucleic Acids Res. *39*, 7828–7836.

Mallory, A.C., Dugas, D. V., Bartel, D.P., and Bartel, B. (2004). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. Curr. Biol. *14*, 1035–1046.

Manavella, P.A., Hagmann, J., Ott, F., Laubinger, S., Franz, M., MacEk, B., and Weigel, D. (2012). Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell *151*, 859–870.

Mansoor, S., Briddon, R.W., Zafar, Y., and Stanley, J. (2003). Geminivirus disease complexes: An emerging threat. Trends Plant Sci. *8*, 128–134.

Marín-González, E., and Suárez-López, P. (2012). "And yet it moves": Cell-to-cell and long-distance signaling by plant microRNAs. Plant Sci. *196*, 18–30.

Martínez de Alba, A.E., Elvira-Matelot, E., and Vaucheret, H. (2013). Gene silencing in plants: A diversity of pathways. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 1300–1308.

Mateos, J.L., Bologna, N.G., and Palatnik, J.F. (2011). 16. Biogenesis of Plant MicroRNAs. In Non Coding RNAs in Plants, V.A. Erdmann, and J. Barciszewski, eds. pp. 251–268.

Matzeit, V., Schaefer, S., Kammann, M., Schalk, H.J., Schell, J., and Gronenborn, B. (1991). Wheat Dwarf Virus Vectors Replicate and Express Foreign Genes in Cells of Monocotyledonous Plants. PLANT CELL ONLINE *3*, 247–258.

Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat. Rev. Genet. *15*, 394–408.

Mayer, K.F.X., Schoof, H., Haecker, A., Lenhard, M., J??rgens, G., and Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell *95*, 805–815.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. *14*, 447–459.

Mirouze, M., Reinders, J., Bucher, E., Nishimura, T., Schneeberger, K., Ossowski, S., Cao, J., Weigel, D., Paszkowski, J., and Mathieu, O. (2009). Selective epigenetic control of retrotransposition in Arabidopsis. Nature *461*, 1–5.

Miyashima, S., Honda, M., Hashimoto, K., Tatematsu, K., Hashimoto, T., Sato-Nara, K.,

Okada, K., and Nakajima, K. (2013). A comprehensive expression analysis of the arabidopsis MICRORNA165/6 gene family during embryogenesis reveals a conserved role in meristem specification and a non-cell-autonomous function. Plant Cell Physiol. *54*, 375–384.

Mlotshwa, S., Pruss, G.J., Peragine, A., Endres, M.W., Li, J., Chen, X., Poethig, R.S., Bowman, L.H., and Vance, V. (2008). Dicer-like2 plays a primary role in transitive silencing of transgenes in Arabidopsis. PLoS One *3*.

Mlotshwa, S., Pruss, G.J., Gao, Z., Mgutshini, N.L., Li, J., Chen, X., Bowman, L.H., and Vance, V. (2010). Transcriptional silencing induced by Arabidopsis T-DNA mutants is associated with 35S promoter siRNAs and requires genes involved in siRNA-mediated chromatin silencing. Plant J. *64*, 699–704.

Moffatt, B.A., Stevens, Y.Y., Allen, M.S., Snider, J.D., Pereira, L.A., Todorova, M.I., Summers, P.S., Weretilnyk, E.A., Martin-mccaffrey, L., and Wagner, C. (2002). Adenosine Kinase Deficiency Is Associated with Developmental Abnormalities and Reduced. Plant Physiol. *128*, 812–821.

Mohannath, G., Jackel, J.N., Lee, Y.H., Buchmann, R.C., Wang, H., Patil, V., Adams, A.K., and Bisaro, D.M. (2014). A complex containing SNF1-related kinase (SnRK1) and adenosine kinase in arabidopsis. PLoS One *9*.

Montgomery, T.A., Howell, M.D., Cuperus, J.T., Li, D., Hansen, J.E., Alexander, A.L., Chapman, E.J., Fahlgren, N., Allen, E., and Carrington, J.C. (2008). Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. Cell *133*, 128–141.

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile Hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. Society *14*, 629–639.

Moussian, B., Schoof, H., Haecker, A., Jürgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. EMBO J. *17*, 1799–1809.

Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol. Plant. *15*, 473–497.

Nag, A., and Jack, T. (2010). Sculpting the flower; the role of microRNAs in flower development. Curr. Top. Dev. Biol. *91*, 349–378.

Nagasaki, H., Itoh, J., Hayashi, K., Hibara, K., Satoh-Nagasawa, N., Nosaka, M., Mukouhata, M., Ashikari, M., Kitano, H., Matsuoka, M., et al. (2007). The small

interfering RNA production pathway is required for shoot meristem initiation in rice. Proc. Natl. Acad. Sci. U. S. A. *104*, 14867–14871.

Nakahigashi, K., Jasencakova, Z., Schubert, I., and Goto, K. (2005). The Arabidopsis HETEROCHROMATIN PROTEIN1 Homolog (TERMINAL FLOWER2) Silences Genes Within the Euchromatic Region but not Genes Positioned in Heterochromatin. Plant Cell Physiol. *46*, 1747–1756.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell *2*, 279–289.

Narasipura, S.D., Kim, S., and Al-Harthi, L. (2014). Epigenetic regulation of HIV-1 latency in astrocytes. J. Virol. *88*, 3031–3038.

Nath, U., Nath, U., Crawford, B.C.W., Carpenter, R., and Coen, E. (2003). Genetic Control of Surface Curvature. Science (80-.). *299*, 1404–1407.

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O., and Jones, J.D.G. (2006). Plants and animals activate defenses after perceiving pathogen-associated molecular patterns (PAMPs) such as bacterial flagellin. In Arabidopsis, perception of flagellin increases resistance to the bacterium Pseudomonas syringae, although the molecular me. Science (80-.). *312*, 436–439.

Navarro, L., Jay, F., Nomura, K., He, S.Y., and Voinnet, O. (2008). Suppression of the MicroRNA Pathway by Bacterial Effector Proteins. Science (80-.).

Nikovics, K., Blein, T., Peaucelle, A., Ishida, T., Morin, H., Aida, M., and Laufs, P. (2006). The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. Plant Cell *18*, 2929–2945.

Nishimura, A., Ito, M., Kamiya, N., Sato, Y., and Matsuoka, M. (2002). OsPNH1 regulates leaf development and maintenance of the shoot apical meristem in rice. Plant J. *30*, 189–201.

Nodine, M.D., and Bartel, D.P. (2010). MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. Genes Dev. *24*, 2678–2692.

Nogueira, F.T.S., and Timmermans, M.C.P. (2007). An Interplay Between Small Regulatory RNAs Patterns Leaves. Plant Signal. Behav. *2*, 519–521.

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

Nogueira, F.T.S., Chitwood, D.H., Madi, S., Ohtsu, K., Schnable, P.S., Scanlon, M.J., and Timmermans, M.C.P. (2009). Regulation of small RNA accumulation in the maize shoot apex. PLoS Genet. *5*.

Nova-franco, B., Íñiguez, L.P., Valdés-lópez, O., Alvarado-affantranger, X., Leija, A., Fuentes, S.I., Ramírez, M., Paul, S., Reyes, J.L., Girard, L., et al. (2015). The Micro-RNA172c-APETALA2-1 Node as a Key Regulator of the Common Bean-Rhizobium etli Nitrogen Fixation Symbiosis. Plant Physiol. *168*, 273–291.

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., Kumlehn, J., and Schweizer, P. (2010). HIGS : Host-Induced Gene Silencing in the Obligate Biotrophic Fungal Pathogen Blumeria graminis. Plant Cell *22*, 3130–3141.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature *425*, 257–263.

Paprotka, T., Deuschle, K., Metzler, V., and Jeske, H. (2011). Conformation-Selective Methylation of Geminivirus DNA. J. Virol. *85*, 12001–12012.

Paprotka, T., Deuschle, K., Pilartz, M., and Jeske, H. (2015). Form follows function in geminiviral minichromosome architecture. Virus Res. *196*, 44–55.

Park, J., Lim, C.H., Ham, S., Kim, S.S., Choi, B.-S., and Roh, T.-Y. (2014). Genome-wide analysis of histone modifications in latently HIV-1 infected T cells. AIDS *28*, 1719–1728.

Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. (2005). Nuclear processing and export of microRNAs in Arabidopsis. Proc. Natl. Acad. Sci. USA *102*, 3691–3696.

Pascal, E., Sanderfoot, A.A., Ward, B.M., Medville, R., Turgeon, R., and Lazarowitz, S.G. (1994). The Geminivirus BR1 Movement Protein Binds Single-Stranded DNA and Localizes to the Cell Nucleus. Plant Cell *6*, 995–1006.

Pattanayak, D., Solanke, A.U., and Kumar, P.A. (2013). Plant RNA Interference Pathways: Diversity in Function, Similarity in Action. Plant Mol. Biol. Report. *31*, 493–506.

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes Dev. *18*, 2368–2379.

Petsch, K., Manzotti, P.S., Tam, O.H., Meeley, R., Hammell, M., Consonni, G., and Timmermans, M.C.P. (2015). Novel DICER-LIKE1 siRNAs Bypass the Requirement for DICER-LIKE4 in Maize Development. Plant Cell *27*,

tpc.15.00194-.

Pikaard, C.S., and Scheid, O.M. (2014). Epigenetic Regulation in Plants. Cold Spring Harb. Perspect. Biol. *4*, 1–31.

Pilartz, M., and Jeske, H. (1992). Abutilon mosaic geminivirus double-stranded DNA is packed into minichromosomes. Virology *189*, 800–802.

Pontivianne, F., Blevins, T., and Pikaard, C.D. (2010). Arabidopsis histone lysine methyltransferases. Adv. Bot. Res. *53*, 1–22.

Pumplin, N., and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat. Rev. Microbiol. *11*, 745–760.

Qiao, Y., Liu, L., Xiong, Q., Flores, C., Wong, J., Shi, J., Wang, X., Liu, X., Xiang, Q., Jiang, S., et al. (2013). Oomycete pathogens encode RNA silencing suppressors. Nat. Genet. *45*, 330–333.

Qiao, Y., Shi, J., Zhai, Y., Hou, Y., and Ma, W. (2015). Phytophthora effector targets a novel component of small RNA pathway in plants to promote infection. PNAS.

Qu, F., Ye, X., and Morris, T.J. (2008). Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. PNAS *105*, 14732–14737.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Raja, P., Sanville, B.C., Buchmann, R.C., and Bisaro, D.M. (2008). Viral genome methylation as an epigenetic defense against geminiviruses. J. Virol. *82*, 8997–9007.

Rajeswaran, R., and Pooggin, M.M. (2012). RDR6-mediated synthesis of complementary RNA is terminated by miRNA stably bound to template RNA. Nucleic Acids Res. *40*, 594–599.

Raman, S., Greb, T., Peaucelle, A., Blein, T., Laufs, P., and Theres, K. (2008). Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in Arabidopsis thaliana. Plant J. *55*, 65–76.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

Revington, G.N., Sunter, G., and Bisaro2, D.M. (1989). DNA Sequences Essential for Replication of the B Genome Component of Tomato Golden Mosaic Virus. Plant Cell *1*, 985–992.

Rodríguez-Negrete, E., Lozano-Durán, R., Piedra-Aguilera, A., Cruzado, L., Bejarano, E.R., and Castillo, A.G. (2013). Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytol. *199*, 464–475.

Rojas, M.R., Hagen, C., Lucas, W.J., and Gilbertson, R.L. (2005). Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu. Rev. Phytopathol. *43*, 361–394.

Roodbarkelari, F., and Groot, E.P. (2016). Regulatory function of homeodomain-leucine zipper (HD-ZIP) family proteins during embryogenesis. New Phytol.

Rubio-Somoza, I., Zhou, C.M., Confraria, A., Martinho, C., Von Born, P., Baena-Gonzalez, E., Wang, J.W., and Weigel, D. (2014). Temporal control of leaf complexity by miRNA-regulated licensing of protein complexes. Curr. Biol. *24*, 2714–2719.

Sakaguchi, J., and Watanabe, Y. (2012). miR165/166 and the development of land plants. Dev. Growth Differ. *54*, 93–99.

Saleh, A., Al-Abdallat, A., Ndamukong, I., Alvarez-Venegas, R., and Avramova, Z. (2007). The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish "bivalent chromatin marks" at the silent AGAMOUS locus. Nucleic Acids Res. *35*, 6290–6296.

Saleh, A., Alvarez-Venegas, R., and Avramova, Z. (2008). An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nat. Protoc. *3*, 1018–1025.

Sanei, M., and Chen, X. (2015). Mechanisms of microRNA turnover. Curr. Opin. Plant Biol. *27*, 199–206.

Scholthof, K.-B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., et al. (2011). Top 10 plant viruses in molecular plant pathology. Mol. Plant Pathol. *12*, 938–954.

Schommer, C., Debernardi, J.M., Bresso, E.G., Rodriguez, R.E., and Palatnik, J.F. (2014). Repression of cell proliferation by miR319-regulated TCP4. Mol. Plant *7*, 1533–1544.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Ju, G., Laux, T., Pflanzen, M. Der,

Morgenstelle, A. Der, Jürgens, G., and Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell *100*, 635–644.

Seefried, W.F., Willmann, M.R.R., Clausen, R.L., and Jenik, P.D. (2014). Global regulation of embryonic patterning in Arabidopsis by microRNAs. Plant Physiol. *165*, 670–687.

Shen, W., and Hanley-Bowdoin, L. (2006). Geminivirus infection up-regulates the expression of two Arabidopsis protein kinases related to yeast SNF1- and mammalian AMPK-activating kinases. Plant Physiol. *142*, 1642–1655.

Shen, W., Reyes, M.I., and Hanley-Bowdoin, L. (2009). Arabidopsis Protein Kinases GRIK1 and GRIK2 Specifically Activate SnRK1 by Phosphorylating Its Activation Loop. PLANT Physiol. *150*, 996–1005.

Shen, W., Dallas, M.B., Goshe, M.B., and Hanley-Bowdoin, L. (2014a). SnRK1 phosphorylation of AL2 delays Cabbage leaf curl virus infection in Arabidopsis. J. Virol. *88*, 10598–10612.

Shen, X., De Jonge, J., Forsberg, S.K.G., Pettersson, M.E., Sheng, Z., Hennig, L., and Carlborg, ??rjan (2014b). Natural CMT2 Variation Is Associated With Genome-Wide Methylation Changes and Temperature Seasonality. PLoS Genet. *10*.

Shimura, H., and Masuta, C. (2016). Plant subviral RNAs as a long noncoding RNA (lncRNA): Analogy with animal lncRNAs in host-virus interactions. Virus Res. *212*, 25–29.

Shimura, H., Pantaleo, V., Ishihara, T., Myojo, N., Inaba, J., Sueda, K., Burgyán, J., and Masuta, C. (2011). A Viral Satellite RNA Induces Yellow Symptoms on Tobacco by Targeting a Gene Involved in Chlorophyll Biosynthesis using the RNA Silencing Machinery. PLoS Pathog. *7*, 1–12.

Shung, C.Y., and Sunter, G. (2009). Regulation of Tomato golden mosaic virus AL2 and AL3 gene expression by a conserved upstream open reading frame. Virology *383*, 310–318.

Shung, C.Y., Sunter, J., Sirasanagandla, S.S., and Sunter, G. (2006). Distinct viral sequence elements are necessary for expression of Tomato golden mosaic virus complementary sense transcripts that direct AL2 and AL3 gene expression. Mol. Plant-Microbe Interact. *19*, 1394–1405.

Skopelitis, D.S., Husbands, A.Y., and Timmermans, M.C.P. (2012). Plant small RNAs as morphogens. Curr. Opin. Cell Biol. *24*, 217–224.

Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Biseeling, T., and Geurts, R. (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor–induced transcription. Science *308*, 1789–1791.

Smith, N.A., Eamens, A.L., and Wang, M. (2011). Viral Small Interfering RNAs Target Host Genes to Mediate Disease Symptoms in Plants. PLoS Pathog. *7*, 1–9.

Soyars, C.L., James, S.R., and Nimchuk, Z.L. (2016). Ready, aim, shoot: Stem cell regulation of the shoot apical meristem. Curr. Opin. Plant Biol. *29*, 163–168.

Sparks, E., Wachsman, G., and Benfey, P.N. (2013). Spatiotemporal signalling in plant development. Nat. Rev. Genet. *14*, 631–644.

Stenger, D.C., Revington, G.N., Stevenson, M.C., and Bisaro, D.M. (1991). Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. Proc. Natl. Acad. Sci. *88*, 8029–8033.

Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y. V., and Jacobsen, S.E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell *152*, 352–364.

Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D.J., and Jacobsen, S.E. (2014). Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat. Struct. Mol. Biol. *21*, 64–72.

Su, Y.H., Liu, Y.B., Zhou, C., Li, X.M., and Zhang, X.S. (2016). The microRNA167 controls somatic embryogenesis in Arabidopsis through regulating its target genes ARF6 and ARF8. Plant Cell. Tissue Organ Cult. *124*, 405–417.

Sun, Y.-W., Tee, C.-S., Ma, Y.-H., Wang, G., Yao, X.-M., and Ye, J. (2015). Attenuation of Histone Methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. Sci. Rep. *5*, 16476.

Sung, S., He, Y., Eshoo, T.W., Tamada, Y., Johnson, L., Nakahigashi, K., Goto, K., Jacobsen, S.E., and Amasino, R.M. (2006). Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nat. Genet. *38*, 706–710.

Sunter, G., and Bisaro, D.M. (1989). Transcription map of the B genome component of tomato golden mosaic virus and comparison with A component transcripts. Virology *173*, 647–655.

Sunter, G., and Bisaro, D.M. (1991). Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression. Virology *180*, 416–419.

Sunter, G., and Bisaro, D.M. (1992). Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. Plant Cell *4*, 1321–1331.

Sunter, G., and Bisaro, D.M. (1997). Regulation of a Geminivirus Coat Protein Promoter by AL2 Protein (TrAP): Evidence for Activation and Derepression Mechanisms. Virology *232*, 269–280.

Sunter, G., and Bisaro, D.M. (2003). Identification of a minimal sequence required for activation of the tomato golden mosaic virus coat protein promoter in protoplasts. Virology *305*, 452–462.

Sunter, G., Gardiner, W.E., and Bisaro, D.M. (1989). Identification of tomato golden mosaic virus-specific RNAs in infected plants. Virology *170*, 243–250.

Sunter, G., Hartitz, M.D., Hormuzdi, S.G., Brough, C.L., and Bisaro, D.M. (1990). Genetic-Analysis Of Tomato Golden Mosaic-Virus - Orf-Al2 Is Required For Coat Protein Accumulation While Orf-Al3 Is Necessary For Efficient Dna-Replication. Virology *179*, 69–77.

Sunter, G., Sunter, J.L., and Bisaro, D.M. (2001). Plants expressing tomato golden mosaic virus AL2 or beet curly top virus L2 transgenes show enhanced susceptibility to infection by DNA and RNA viruses. Virology *285*, 59–70.

Szarzynska, B., Sobkowiak, L., Pant, B.D., Balazadeh, S., Scheible, W.R., Mueller-Roeber, B., Jarmolowski, A., and Szweykowska-Kulinska, Z. (2009). Gene structures and processing of Arabidopsis thaliana HYL1-dependent pri-miRNAs. Nucleic Acids Res. *37*, 3083–3093.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001). SET Domain-containing Protein, G9a, is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3. J. Biol. Chem. *276*, 25309–25317.

Takada, S., Hibara, K., Ishida, T., and Tasaka, M. (2001). The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. Development *128*, 1127–1135.

Tanaka, W., Ohmori, Y., Ushijima, T., Matsusaka, H., Matsushita, T., Kumamaru, T., Kawano, S., and Hirano, H.-Y. (2015). Axillary Meristem Formation in Rice Requires the WUSCHEL Ortholog TILLERS ABSENT1. Plant Cell *27*, 1173–1184.

Tang, G., Reinhart, B.J., Bartel, D.P., and Zamore, P.D. (2003). A biochemical framework for RNA silencing in plants. Genes Dev. *17*, 49–63.

Tang, X., Bian, S., Tang, M., Lu, Q., Li, S., Liu, X., Tian, G., Nguyen, V., Tsang, E.W.T., Wang, A., et al. (2012). MicroRNA-Mediated Repression of the Seed Maturation Program during Vegetative Development in Arabidopsis. PLoS Genet. *8*, 20–22.

Tatematsu, K., Toyokura, K., Miyashima, S., Nakajima, K., and Okada, K. (2015). A molecular mechanism that confines the activity pattern of miR165 in Arabidopsis leaf primordia. Plant J. *82*, 596–608.

Teotia, S., and Tang, G. (2015). To bloom or not to bloom: Role of microRNAs in plant flowering. Mol. Plant *8*, 359–377.

Thompson, B.E., Basham, C., Hammond, R., Ding, Q., Kakrana, A., Lee, T.-F., Simon, S.A., Meeley, R., Meyers, B.C., and Hake, S. (2014). The dicer-like1 Homolog fuzzy tassel Is Required for the Regulation of Meristem Determinacy in the Inflorescence and Vegetative Growth in Maize. Plant Cell *26*, tpc.114.132670.

Tian, D., Traw, M.B., Chen, J.Q., Kreitman, M., and Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana. Nature *423*, 74–77.

Tian, L., Li, X., Ha, M., Zhang, C., and Chen, Z.J. (2014). Genetic and epigenetic changes in a genomic region containing MIR172 in Arabidopsis allopolyploids and their progenitors. Heredity (Edinb). *112*, 207–214.

Tretter, E.M., Alvarez, J.P., Eshed, Y., and Bowman, J.L. (2008). Activity Range of Arabidopsis Small RNAs Derived from Different Biogenesis Pathways. PLANT Physiol. *147*, 58–62.

Trinks, D., Rajeswaran, R., Shivaprasad, P. V, Oakeley, E.J., Veluthambi, K., Pooggin, M.M., Akbergenov, R., and Hohn, T. (2005). Suppression of RNA Silencing by a Geminivirus Nuclear Protein , AC2 , Correlates with Transactivation of Host Genes Suppression of RNA Silencing by a Geminivirus Nuclear Protein , AC2 , Correlates with Transactivation of Host Genes †. J. Virol. *79*, 2517–2527.

Tucker, M.R., Hinze, A., Tucker, E.J., Takada, S., Jürgens, G., and Laux, T. (2008). Vascular signalling mediated by ZWILLE potentiates WUSCHEL function during shoot meristem stem cell development in the Arabidopsis embryo. Development *135*, 2839–2843.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet. *3*, 0855–0866.

Turner, M., Nizampatnam, N.R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., Arunachalam, S.P., Yu, O., Subramanian, S., Science, P., et al. (2013). Ectopic

Expression of miR160 Results in Auxin Hypersensitivity, Cytokinin Hyposensitivity, and Inhibition of Symbiotic Nodule Development in Soybean. Plant Physiol. *162*, 2042–2055.

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T.C., Mcelver, J., Aux, G., Patton, D., et al. (2004). Identification of Genes Required for Embryo Development in Arabidopsis. Plant Physiol. *135*, 1206–1220.

Vanstraelen, M., and Benková, E. (2012). Hormonal Interactions in the Regulation of Plant Development. Annu. Rev. Cell Dev. Biol *28*, 463–487.

Várallyay, E., Válóczi, A., Agyi, A., Burgyán, J., and Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. EMBO J. *29*, 3507–3519.

Vargason, J.M., Szittya, G., Burgyán, J., and Tanaka Hall, T.M. (2003). Size Selective Recognition of siRNA by an RNA Silencing Suppressor. Cell *115*, 799–811.

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

Vogler, H., Akbergenov, R., Shivaprasad, P. V, Dang, V., Fasler, M., Kwon, M.-O., Zhanybekova, S., Hohn, T., and Heinlein, M. (2007). Modification of small RNAs associated with suppression of RNA silencing by Tobamovirus replicase protein. J. Virol. *81*, 10379–10388.

Voinnet, O. (2009). Origin, Biogenesis, and Activity of Plant MicroRNAs. Cell *136*, 669–687.

Wang, H., Hao, L., Shung, C.Y., Sunter, G., and Bisaro, D.M. (2003). Adenosine Kinase Is Inactivated by Geminivirus AL2 and L2 Proteins. Plant Cell Online *15*, 3020–3032.

Wang, H., Buckley, K.J., Yang, X., Buchmann, R.C., and Bisaro, D.M. (2005). Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. J. Virol. *79*, 7410–7418.

Wang, L., Mai, Y.X., Zhang, Y.C., Luo, Q., and Yang, H.Q. (2010a). MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in arabidopsis. Mol. Plant *3*, 794–806.

Wang, L., Song, X., Gu, L., Li, X., Cao, S., Chu, C., Cui, X., Chen, X., and Cao, X. (2013). NOT2 proteins promote polymerase II-dependent transcription and interact with multiple MicroRNA biogenesis factors in Arabidopsis. Plant Cell *25*, 715–727.

Wang, Q., Hasson, A., Rossmann, S., and Theres, K. (2016). Divide et impera: Boundaries shape the plant body and initiate new meristems. New Phytol. *209*, 485–498.

Wang, X., Xu, Y., Han, Y., Bao, S., Du, J., Yuan, M., Xu, Z., and Chong, K. (2006). Overexpression of RAN1 in Rice and Arabidopsis Alters Primordial Meristem, Mitotic Progress, and Sensitivity to Auxin 1. Plant Physiol. *140*, 91–101.

Wang, X.-B., Wu, Q., Ito, T., Cillo, F., Li, W.-X., Chen, X., Yu, J.-L., and Ding, S.-W. (2010b). RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. *107*, 484–489.

Wang, X.-B., Jovel, J., Udomporn, P., Wang, Y., Wu, Q., Li, W., Gasciolli, V., Vaucheret, H., and Ding, S. (2011). The 21-Nucleotide , but Not 22-Nucleotide , Viral Secondary Small Interfering RNAs Direct Potent Antiviral Defense by Two Cooperative Argonautes in Arabidopsis thaliana. Plant Cell.

Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F., Tian, Y., Jiang, Q., Ferguson, B.J., et al. (2014). Soybean miR172c Targets the Repressive AP2 Transcription Factor NNC1 to Activate ENOD40 Expression and Regulate Nodule Initiation. Plant Cell *26*, 4782–4801.

Wang, Y., Li, K., Chen, L., Zou, Y., Liu, H., Tian, Y., Li, D., Wang, R., Zhao, F., Ferguson, B.J., et al. (2015a). MicroRNA167-Directed Regulation of the Auxin Response Factors GmARF8a and GmARF8b Is Required for Soybean Nodulation and Lateral Root Development 1 [OPEN]. Plant Physiol. *168*, 101–116.

Wang, Z., Li, X., Jiang, Y., Shao, Q., Liu, Q., Chen, B., and Huang, D. (2015b). swDMR: A Sliding Window Approach to Identify Differentially Methylated Regions Based on Whole Genome Bisulfite Sequencing. PLoS One *10*, e0132866.

Weiberg, A., Wang, M., Lin, F., Zhao, H., Kaloshian, I., and Jin, H. (2013). Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. Science (80-.). *342*, 118–123.

Weretilnyk, E. a, Alexander, K.J., Drebenstedt, M., Snider, J.D., Summers, P.S., and Moffatt, B. a (2001). Maintaining methylation activities during salt stress. The involvement of adenosine kinase. Plant Physiol. *125*, 856–865.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Noncoding transcription by RNA Polymerase IVb/PolV mediates trasncriptional silencing of overlapping and adjacent genes. Cell *135*, 635–648.

Williams, L., Grigg, S.P., Xie, M., Christensen, S., and Fletcher, J.C. (2005). Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. Development *132*, 3657–3668.

Wilson, R.C., and Doudna, J.A. (2013). Molecular Mechanisms of RNA Interference. Annu. Rev. Biophys. *42*, 217–239.

Wlodaver, A.M., and Staley, J.P. (2014). The DExD/H-box ATPase Prp2p destabilizes and proofreads the catalytic RNA core of the spliceosome. RNA *20*, 282–294.

Wu, G. (2013). Plant MicroRNAs and Development. J. Genet. Genomics *40*, 217–230.

Wu, G., Park, M.Y., Conway, S.R., Wang, J.W., Weigel, D., and Poethig, R.S. (2009). The Sequential Action of miR156 and miR172 Regulates Developmental Timing in Arabidopsis. Cell *138*, 750–759.

Wu, J., Yang, Z., Wang, Y., Zheng, L., Ye, R., Ji, Y., Zhang, X., Cao, X., Xie, L., Wu, Z., et al. (2015). Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA. Elife 1–19.

Xie, K. (2006). Genomic Organization, Differential Expression, and Interaction of SQUAMOSA Promoter-Binding-Like Transcription Factors and microRNA156 in Rice. Plant Physiol. *142*, 280–293.

Xin, M., Wang, Y., Yao, Y., Xie, C., Peng, H., Ni, Z., and Sun, Q. (2010). Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (Triticum aestivum L .). BMC Plant Biol. *10*, 1–11.

Xue, X.Y., Zhao, B., Chao, L.M., Chen, D.Y., Cui, W.R., Mao, Y.B., Wang, L.J., and Chen, X.Y. (2014). Interaction between Two Timing MicroRNAs Controls Trichome Distribution in Arabidopsis. PLoS Genet. *10*.

Yadava, P., Suyal, G., and Mukherjee, S.K. (2010). Begomovirus DNA replication and pathogenicity. Curr. Sci. *98*, 360–368.

Yanai, O., Shani, E., Dolezal, K., Tarkowski, P., Sablowski, R., Sandberg, G., Samach, A., and Ori, N. (2005). Arabidopsis KNOXI proteins activate cytokinin biosynthesis. Curr. Biol. *15*, 1566–1571.

Yang, H., Chang, F., You, C., Cui, J., Zhu, G., Wang, L., Zheng, Y., Qi, J., and Ma, H. (2015). Whole-genome DNA methylation patterns and complex associations with gene structure and expression during flower development in Arabidopsis. Plant J *81*, 268–281.

Yang, L., Liu, Z., Lu, F., Dong, A., and Huang, H. (2006a). SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. Plant J. *47*, 841–850.

Yang, L., Xu, M., Koo, Y., He, J., and Scott Poethig, R. (2013). Sugar promotes vegetative phase change in Arabidopsis thaliana by repressing the expression of MIR156A

and MIR156C. Elife *2013*, 1–15.

Yang, X., Baliji, S., Buchmann, R.C., Wang, H., Lindbo, J.A., Sunter, G., and Bisaro, D.M. (2007). Functional modulation of the geminivirus AL2 transcription factor and silencing suppressor by self-interaction. J. Virol. *81*, 11972–11981.

Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006b). HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2 OH of the 3 0 terminal nucleotide. Nucleic Acids Res. *34*, 667–675.

Ye, K., Malinina, L., and Patel, D.J. (2003). Recognition of small interfering RNA by a viral suppressor ofRNAsilencing. Nature *426*, 874–878.

Yoshikawa, M., Iki, T., Numa, H., Miyashita, K., Meshi, T., and Ishikawa, M. (2016). A role of a short open reading frame encompassing the microRNA173 target site of the TAS2 transcript in trans-acting small interfering RNA biogenesis. Plant Physiol. *171*, pp.00148.2016.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a Crucial Step in Plant microRNA Biogenesis. Science. *307*, 932–935.

Yu, B., Chapman, E.J., Yang, Z., Carrington, J.C., and Chen, X. (2006). Transgenically expressed viral RNA silencing suppressors interfere with microRNA methylation in Arabidopsis. FEBS Lett. *580*, 3117–3120.

Yu, S., Li, C., Zhou, C.M., Zhang, T.Q., Lian, H., Sun, Y., Wu, J., Huang, J., Wang, G., and Wang, J.W. (2013). Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. Elife *2013*, 1–17.

Zahid, K., Zhao, J., Smith, N.A., Schumann, U., Fang, Y., Dennis, E.S., Zhang, R., Guo, H., and Wang, M. (2015). Nicotiana Small RNA Sequences Support a Host Genome Origin of Cucumber Mosaic Virus Satellite RNA. PLoS Genet. *11*, 1–13.

Zhang, C., Li, G., Wang, J., Zhu, S., and Li, H. (2013a). Cascading cis-cleavage on transcript from trans-acting siRNA-producing locus 3. Int. J. Mol. Sci. *14*, 14689–14699.

Zhang, H., and Zhu, J.K. (2011). RNA-directed DNA methylation. Curr. Opin. Plant Biol. *14*, 142–147.

Zhang, H., Huang, L., Dai, Y., Liu, S., Hong, Y., Tian, L., Huang, L., Cao, Z., Li, D., and Song, F. (2015). Arabidopsis AtERF15 positively regulates immunity against *Pseudomonas syringae* pv. tomato DC3000 and *Botrytis cinerea*. Front. Plant Sci. *6*, 686.

Zhang, Z., and Zhang, X. (2012). Argonautes compete for miR165/166 to regulate shoot apical meristem development. Curr. Opin. Plant Biol. *15*, 652–658.

Zhang, S., Xie, M., Ren, G., and Yu, B. (2013b). CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. Proc. Natl. Acad. Sci. U. S. A. *110*, 17588–17593.

Zhang, X., Garreton, V., and Chua, N.-H. (2005). The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. Genes Dev. *19*, 1532–1543.

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

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006b). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat. Protoc. *1*, 641–646.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E. (2007a). Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biol. *5*, 1026–1035.

Zhang, X., Germann, S., Blus, B.J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S.E. (2007b). The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nat. Struct. Mol. Biol. *14*, 869–871.

Zhang, X., Zou, Z., Zhang, J., Zhang, Y., Han, Q., Hu, T., Xu, X., Liu, H., Li, H., and Ye, Z. (2011a). Over-expression of sly-miR156a in tomato results in multiple vegetative and reproductive trait alterations and partial phenocopy of the sft mutant. FEBS Lett. *585*, 435–439.

Zhang, X., Zhao, H., Gao, S., Wang, W.-C., Katiyar-Agarwal, S., Huang, H.-D., Raikhel, N., and Jin, H. (2011b). Arabidopsis Argonaute 2 regulates innate immunity via miRNA393* -mediated silencing of a Golgi-localized SNARE gene, MEMB12. Mol. Cell *42*, 356–366.

Zhang, X., Zhang, X., Singh, J., Li, D., and Qu, F. (2012). Temperature-dependent survival of Turnip crinkle virus-infected arabidopsis plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. J. Virol. *86*, 6847–6854.

Zhang, Z., Chen, H., Huang, X., Xia, R., Zhao, Q., Lai, J., Teng, K., Li, Y., Liang, L., Du, Q., et al. (2011c). BSCTV C2 attenuates the degradation of SAMDC1 to suppress DNA methylation-mediated gene silencing in Arabidopsis. Plant Cell *23*, 273–

288.

Zhang, Z., Liu, X., Guo, X., Wang, X.-J., and Zhang, X. (2016). Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing. Nat. Plants *2*, 16049.

Zheng, B., and Chen, X. (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. Curr. Opin. Plant Biol. *14*, 123–129.

Zhou, X. (2013). Advances in understanding begomovirus satellites. Annu. Rev. Phytopathol. *51*, 357–381.

Zhou, Y., Liu, X., Engstrom, E.M., Nimchuk, Z.L., Pruneda-Paz, J.L., Tarr, P.T., Yan, A., Kay, S.A., and Meyerowitz, E.M. (2015a). Control of plant stem cell function by conserved interacting transcriptional regulators. Nature *517*, 377–380.

Zhou, Y., Honda, M., Zhu, H., Zhang, Z., Guo, X., Li, T., Li, Z., Peng, X., Nakajima, K., Duan, L., et al. (2015b). Spatiotemporal sequestration of miR165/166 by arabidopsis argonaute10 promotes shoot apical meristem maintenance. Cell Rep. *10*, 1819–1827.

Zhu, Q.H., and Helliwell, C.A. (2011). Regulation of flowering time and floral patterning by miR172. J. Exp. Bot. *62*, 487–495.

Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. (2011). Arabidopsis argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell *145*, 242–256.

Zhu, H., Zhou, Y., Castillo-González, C., Lu, A., Ge, C., Zhao, Y.-T., Duan, L., Li, Z., Axtell, M.J., Wang, X.-J., et al. (2013). Bidirectional processing of pri-miRNAs with branched terminal loops by Arabidopsis Dicer-like1. Nat. Struct. Mol. Biol. *20*, 1106–1115.

Zhu, Q.-H., Upadhyaya, N.M., Gubler, F., and Helliwell, C. a (2009). Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (Oryza sativa). BMC Plant Biol. *9*, 149.

Zrachya, A., Glick, E., Levy, Y., Arazi, T., Citovsky, V., and Gafni, Y. (2007). Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology *358*, 159–165.

Zuo, J., Niu, Q.W., and Chua, N.H. (2000). Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. *24*, 265–273.

APPENDIX A. SUPPLEMENTARY FILES

Large data sets were produced during the development of the research consigned in this manuscript. All the raw and processed data can be found in the supplementary files as follows:

- Supplementary file 1. Microarray analysis of *35S-TrAP* transgenic plants.
- Supplementary file 2. Transcriptome comparison of the genes deregulated by *TrAP* overexpression and *lhp1* loss-of-function mutant.
- Supplementary file 3. Expression levels of genes encoding for TGS components in TrAP transgenic plants.
- Supplementary file 4. List of proteins tested for interaction with TrAP.
- Supplementary file 5. CG methylation analysis of Col-0 wild type, *kyp* mutant, and TrAP transgenic plants.
- Supplementary file 6. CHG methylation analysis of Col-0 wild type, *kyp* mutant, and TrAP transgenic plants.
- Supplementary file 7. CHH methylation analysis of Col-0 wild type, *kyp* mutant, and TrAP transgenic plants.
- Supplementary file 8. List of primers used in this article.

APPENDIX B. PROTEOMICS ANALYSES OF TRAP COMPLEXES *IN PLANTA*

To further characterize TrAP function in the plant cell, we purified tomato golden mosaic virus TrAP protein complexes using a two-step immunoprecipitation assay, followed by mass spectrometry. For this, the FM-TrAP tagged protein was expressed from the beta-estradiol inducible promoter, XVE. Two different expression methods were assessed: transient expression in leaves of the host plant *N. benthamiana*, and induction of expression in seedling of *Arabidopsis thaliana* (ecotype Col-0) stable transgenic plants. The negative controls were *N. benthamiana* leaves inoculated with agrobacteria carrying an empty vector, and beta-estradiol treated *A. thaliana* Col-0 seedlings. The potential TrAP-interacting proteins were identified by comparison with the negative controls. For our analyses we only included proteins identified by more than one peptide that were not recovered in the negative control. Below are the detailed tables of the TrAP co-immunoprecipitated proteins.

Table 2. Identification of TrAP interacting proteins by IP-MS in *Nicotiana benthamiana* seedlings.

Unique peptides	Gene Identifier	Description	N.bentha sequence	PROTEIN SEQUENCE, UNIQUE REFERENCE	Protein Coverage
14	gil239819394	Cell division control protein	Niben101Scf01911g15015.1	MSHQAESSDSKGAKKDPSTAILERKKSPNRLVVDEAVNDDNSVVALNPATMEKLQLFRGDTILIKGKKRKDTVVIALADETCDEPKIRMNK VVRSNLRVRLGDVVSVHQCPDVKYGKRVHILPIDDTIEGLTGDLFDAFLKPYFLEAYRPLRKGDNFLVRGGMRSVEFKVIETDPGEYCVVA PDTEIFCEGEPVKRKDEERLDEVGYDDVGGVRKQMAQIRELVELPLRHPQLFKSIGVKPPKGILLYGPPGSGKTLIARAVANETGAFFFCI NGPEIMSKLAGESESNLRKAFEEAEKNAPSIIFIDEIDSIAPKREKTHGEVERRIVSQLLTLMDGLKSRAHVIVMGATNRPNSIDPALRRF GRFDREIDIGVPDEVGRLEVLRIHTKNMKLAREVDLERISKDTHGYVGADLAALCTEAALQCIREKMDVIDLEDDSIDAEIINSMAVTNEH FQTALGTSNPSALRETVVEVPNVSWEDIGGLENVKRELQETVQYPVEHPEKFEKFGMSPSKGVLFYGPPGCGKTLLAKAIANECQANFISV KGPELLTMWFGESEANVREIFDKARQSAPCVLFFDELDSIATQRGSSVGDAGGAADRVLNQLLTEMDGMNAKKTVFIIGATNRPDIIDPAL LRPGRLDQLIYIPLPDEDSRHQIFKACLRKSPLSKDVDLRALAKYTQGFSGADITEICQRACKYAIRENIEKDIEKERKRSENPDSMDEDA DDEIAEITPSHFEESMKYARRSVSDADIRKYQAFAQTLQQSRGFGTEFRFAEASGGADATDPFATSNAGADDDDLYS	21.90%
13	gil84688912	AGO4-2	Niben101Scf05519g01007.1	MAEEDKNGAAEGLPPPPPVPPDFTPAISEPEPVKKKAALRLPMARRGLGNKGQKIQILTNHFKVNVTNVDGHFFHYSVALFYEDGRPVDGK GVGRKVLDRVHETYDTELAGKDFAYDGEKSLPTIGALPRNKMEFTVVLEDVTSNRNNGNSSPAADEGPNESDRKRLRRPYQSKSFKVEISF AAKIPMQAIANALRGQETENSQEALRVLDIILRQHAAKQGCLLVRQSFFHNDPKNFVDVGGGVLGCRGFHSSFRTTQSGLSLNIDVSTIMI IQPGPVVDFLIANQNAKDPYTLDWAKAKRMLKNLRVKTSPTNQEFKITGLSDRPCREQTFYLKQKGKDGEGDEITVYDYFVNHRNIDLRYS ADLPCINVGKPKRPTYFPIELCNLVSLQRYTKSLSTFQRSSLVEKSRQKPQEPMQVLSNALKINKYDAEPLLRACGISISSNFTQVEGRVL SPPKLKTGGDDFVPRNGRWNFNNKRLVDPTKIERWAVVNFSARCNIQGLISDLIKCGKMKGIMVEDPPDVPEESPQFRRAPPLVRVEKMPE EVQSKLPGAPKFLLCLLPERKNCDIYGPWKRKNLAEFGIVTQCIAPTRVNDQYITNVLLKTNAKLGGLNSMLTVEHAPAIPMVSKVPTIIL GMDVSHGSPGQSDVPSIAAVVSSRQWPSISRYRASVRTQSPKVEMIDNLFKPASDTEDEGIMRERALLDFYVSSGKRKPEHIIIFRDGVSES QFNQVLNIELDQIIEACKFLDEKWSPKFTVIIAQKNHHTKFFQPGDPNNVPPGTIIDNKVCHPRNYDFYLCAHAGMIGTTRPTHYHVLHDE IGFSPDDLQELVHNLSYVYQRSTTAISVVAPICYAHLAATQMGQWMKFEDTSETSSSRGGVTNAGPVTVPQLPKLEEKVSSSMFFC	16.10%
9	gil347453972	NADH-plastoquinone oxidoreductase subunit 7	Niben101Scf07823g00002.1	MTAPTTRKDLMIVNMGPQHPSMHGVLRLIVTLDGEDVVDCEPILGYLHRGMEKIAENRTIIQYLPYVTRWDYLATMFTEAITINGPEQLGN IQVPKRASYIRVIMLELSRIASHLLWLGPFMADIGAQTPFFYIFRERELIYDLFEAATGMRMMHNFFRIGGVAADLPYGWIDKCLDFCDYF LTGVAEYQKLITRNPVFLERVEGVGIIGGDEALNWGLSGPMLRASGIEWDLRKVDHYESYDEFDWQVQWQREGDSLARYLVRIGEMTESIK IIQQALEGIPGGPYENLEIRRFDRLKDPEWNDFEYRFISKKPSPTFELSKQELYVRVEAPKGELGIFLIGDQSVFPWRWKIRPPGFINLQI LPQLVKRMKLADIMTILGSIDIIMGEVDR	25.40%
8	gil88604736	leucine-rich repeat protein	Niben101Scf04386g04007.1	MMGSFLYITTLLLGIISAVHSCPPSDRAALLAFRAALNESSLGIFNTWQGYECCHGWYGVSCDQLTHRVADINLRGESEDDLPQKAHKTGY MTGTISPAICKLERLSSLTIADWKGITGPIPSCVTSLPFLRIIDLIGNKLTGEIPSEIGRLSRLTVLNVADNCLSGRIPRSLTNLSSLMHL DLRNNRIPGTLPTNFGKLRMLSRALLSGNRLIGKIPYSISYIYRLSLDLDLSLNKLSGTIPPSLGKMHVLATLNLDGNNISGTIPPTLINSR INILNLSKNSIEGYIPDSFDERSYFMVMDLSYNKLRGKIPKSISSATFIGHFDVSHNHLCGQIPAGSPFDHLEASSFTYNDCLCGKPLKPC R	23.60%
7	gil156763844	N-acetyl-gamma-glutamyl-phosphate reductase, chloroplastic	Niben101Scf03080g03001.1	MSSSLAALSSSHFNGASSVWKIQGELKVSKVGKFRIRATSMATSSPAQGLQFGEGKANKLDKLRIGVLGASGYTGTEIIRLLANHPHFQIT LMTADRKAGQSIESVFPHLVTQDLPSLVAVKDADFSAVDAVFCCLPHGTTQEIIKSLPTNLKIVDLSADFRLCDIAEYEEWYGQPHKAMAL QQEAVYGLTEISREIQSARLVANPGCYPTSVQLPLIPLIKSNLIEVRNIIIDSKSGVSGAGRAAKEANLFTEIAEGIHSYGITRHRHVPE IEQGLSEAANSKVTVSFTPHLMPMSRGMQSTIYVEMAPGVTTQDLKDRLVNYYEHEKFVVLLKNKEVPYTRHVTGSNYCLMNVFPDRIPGR AIIVSVIDNLVKGASGQALQNLNVMMGIPENLGLSSLPLFP	23.70%
7	gil365189292	Heat shock protein 90	Niben101Scf04331g09018.1	MRKWTIPSVLFLLCLLFLLPDQGRRIQANAEAESDVPVDPPKVEEKFGAIPHGLSTDSDVVKRESESMSRKTLRADAEKFEFQAEVSRLMD IIINSLYSNKDIFLRELISNASDALDKIRFLSLTDKEVLGEGDNAKLEIQIKLDKEKKILSIRDRGIGMTKEDLIKNLGTIAKSGTSAFVE KMQTSGDLNLIGQFGVGFYSVYLVADYVEVISKHNDDKQYVWESKADGAFAISEDVWNEPLGRGTEIRLHLRDDAGEYLDEYK LKDLVKKY SEFINFPIYLWASKEVEKEVPADEDETSDEEETSETSPSEEEGDEDDSEKAEDEKKPKTKKVKETTYEWELLNDVKAIWLRNPKEVTEEEY TKFYHSLAKDFSDEKPLAWSHFNAEGDVEFKAVLFVPPKAPHDLYESYYNSKKSNLKLYVRRVFISDEFDELLPKYLSFLMGLVDSDTLPL NVSREMLQQHSSLKTIKKKLIRKALDMIRKIADEDPDESNDKDKKDVEESGADNEKKGQYAKFWNEFGKSIKLGIIEDATNRNRLAKLLRF ETSKSDGKLTSLDQYISRMKAGQKDIFYITGASKEQLEKSPFLERLTKKNYEVIFFTDPVDEYLMQYLMDYEDNKFQNVSKEGLKLGKDSK AKELKESFKELTKWWKGALASDNVDDVKISNRLADTPCVVVTSKYGWSANMERIMQSQTLSDANKQAYMRGKRVLEINPRHPIIKELRDRV IKDPEDESVKQTAQLMYQTALLESGFLLNDPKDFASRIYSSVKSSLNVSPDATVEEEEDIEEPEAETDDKEAAAKDDYDAKDEL	9.60%
6	gil40457328	Glutamine synthetase GS58	Niben101Scf02330g00012.1	MAQILAPSAQWQMRMTKSSTDANPLTSKMWSSVVLKQNKRLAVKSSAKFRVFALQSDSGTVNRVEQLLNLDVTPYTDKIIAEYIWIGGSGI DMRSKSRTISKPVKHASELPKWNYDGSSTGQAPGEDSEVILYPQAIFKDPFRGGNNILVICDAYTPAGEPIPTNKRHKAAQIFSDSKVVSE VPWFGIEQEYTLLQQNVKWPLGWPVGGYPGPQGPYYCGAGADKSFGRDISDAHYKACLYAGININSGTNGEVMPGQWEFQVGPSVGIEAGDH IWCARYILERITEQAGVVLSLDPKPIEGDWNGAGCHTNYSTLSMREEGGFEVIKKAILNLSLRHKEHISAYGEGNERRLTGKHETASIDKF SWGVANRGASIRVGRDTEKQGKGYLEDRRPASNMDPYVVTGLLAETTILWEPTLEAEALAAQKLALNV	11.10%
5	gil2493810	Oxygen-dependent coproporphyrinogen-III oxidase, chloroplastic	Niben101Scf03281g01014.1	MLTPILSSASCSWTPTSQFPHSWHSSPSFLTKPLNLPFTESYKTAKRPTPNYSFKVQAMIEKEVAVSHKPDAFLRRSDMGSNVTSNSSSVR GRFEKMRREAQDSVCLAIEKADGGAKFKEDVWSRPGGGGGHSSVLQDGAVFEKAGVNVSVVYGVMPPEAYRAARPTDNGNVKPGPIPFFAA GVSSVLHPKNPFAPTLHFNYRYFETDAPKDAPGAPRQWWFGGGTDFTPAYIFEEDVKHFHSVQKAACDKFDASFYPRFKKWCVDYFYIKHR DERRGLGGIFFDDFNDYDQEMLLSFSTECANSVIPAYIPIVEKRKDTPFTDKHKAWQQLRRGRYVVEFNLVYDRGTTFGLKTGGRIESILVS LPLTARWEYDHKPEEGTEEWKLLDACINPKEWI	15.60%

Table 2 Continued.

Unique peptides	Gene Identifier	Description	N.bentha sequence	PROTEIN SEQUENCE, UNIQUE REFERENCE	Protein Coverage
5	gil296512847	SERINE HYDROXYMETHYLTRANSFERASE 1	Niben101Scf01048g00005.1	MAMATALRRLSSSVDKPIKRLYNGGSLYYMSSLPNEAVYEKEKNGVTWPKQLNAPLEEVDPEIADITIELEKARQWKGLELIPSENFTSVSV MQAVGSVMTNKYSEGYPGARYYGGNEYIDMAETLCQKRALEAFRLDPAKWGVNVQPLSGSPANFHVYTALLKPHERIMALDLPHGGHLSHG YQTDTKKISAVSIFFETMPYRLNESTGYIDYDQLEKSATLFRPKLIVAGASAYARLYDYARIRKVCDKQKAIMLADMAHISGLVAAGVIPS PFDYADVVTTTTHKSLRGPRGAMIFFRKGVKEVNKQGKEVLYDYEDKINQAVFPGLQGGPHNHTITGLAVALKQAMTPEYRAYQEQCLSNC SKFAQALAGMGYELVSGGTENHLVLVNLKNKGIDGSRVEKVLEAVHIAANKNTVPGDVSAMVPGGIRMGTPALTSRGFIEEDFVKVAEFFD AAVKIAVKIKTEAQGTKLKDFVTTLQSSASIQSEIAKLRHGVEEYAKQFPTIGFEKETMKYKN	9.70%
5	gil4827251	aldolase NPALDP1	Niben101Scf02864g04008.1	MASASLLKTSPAIGKTDFIKGQALRQPSVSVRCHPAPPSGLTVRASSYADELVKTAKTVASPGRGILAMDESNATCGKRLASIGLENTEAN RQAYRTLLVSAPGLGQYISGAILFEETLYQSTVDGKKIVDVLHEQNIVPGIKVDKGLVPLAGSNDESWCQGLDGLASRSAAYYQQGARFAK WRTVVSIPNGPSALAVKEAAWGLARYAAISQDNGLVPIVEPEILLDGEHNIDRTFEVAQKVWAEVFFYLAENNVMFEGILLKPSMVTPGAE CKEPATPDQVADYTLKLLQRRIPPAVPGIMFLSGGQSEVEATLNLNAMNQSPNPWHVSFSYARALQNTCLKTWGGRPENVQAAQEALLIRA NANSLAQLGKYTGEGESEEAKKGMFVKGYVY	10.10%
5	gil315258233	mitogen-activated protein kinase kinase2	Niben101Scf01283g02011.1	MRPLQPPPPAAAATTTSSSTTASPMPPPPSRNRPRRRTDLTLPLPQRDPALAVPLPLPPTSAPSSSSSSSSSPLPTPLNFSELERINRIGS GAGGTVYKVLHRPTGRLYALKVIYGNHEDSVRLQMCREIEILRDVDNPNVVRCHDMFDHNGEIQVLLEFMDKGSLEGIHIPKESALSDLTR QVLSGLYYLHRRKIVHRDIKPSNLLINSRREVKIADFGVSRVLAQTMDPCNSSVGTIAYMSPERINTDLNHGQYDGYAGDIWSLGVSILEF YLGRFPFSVGRSGDWASLMCAICMSQPPEAPANASREFRDFIACCLQRDPARRWTAVQLLRHPFITQNSPATTTTGNMMPLPNQVHQPAH QLLPPPPHFSS	12.50%
5	gil354549249	S-adenosyl-L-methionine synthase	Niben101Scf00285g00001.1	CETCTKTNMVMVFGEITTKANVDYEEIVRDTCRGIGFTSPDVGLDADNCKVLVNIEQQSPDIAQGVHGHLTKKPEEIGAGDQGHMFGYATD ETPELMPLTHVLATQLGAKLTEVRKNKTCPWLRPDGKTQVTVEYKKNENGAMVPIRVHTVLISTQHDETVTNEQIAKDLKEHVIKPVIPAKY LDEKTIPHLNPSGRFVIGGPHGDAGLTGRKIIIDTYGGWGAHGGGAFSGKDPTKVDRSGAYVVRQAAKSVVAAGLARRCIVQVSYAIGVAE PLSVFVDTYKTGTIPDKDILALIKENFDFRPGMI	22.10%
5	gil7682432	MAP kinase kinase	Niben101Scf02790g03012.1	MKKGSLAPNLKLSLPPPDEVNLSKFLTESGTFKDGDLLVNRDGVRIVSQSEVEAPSVIQPSDNQLCLADFEAVKVIGKGNGGIVRLVQHKW TGQFFALKAIQMNIEESMRKHIAQELRINQSSQVPYVVISYQSFFDNGAISIILEYMDGGSLADFLKKVKTIPERYLAAICKQVLKGLWYL HHEKHIIHRDLKPSNLLINHIGDVKITDFGVSAVLASTSGLANTFVGTYNYMSPERILGGAYGYRSDIWSLGLVLLECATGVFPYSPPQAD EGWVNVYELMETIVDQPAPSAPPDQFSPQFCSFISACVQKDQKDRLSANELMRHPFITMYDDLDIDLGSYFTSAGPPLATLTEL	11.50%
5	gil257670406	Unnamed protein product	Niben101Scf25430g00015.1	MIQAAFRTDEIRRTPPTPQDEMRAGMSYFHETIWKGVPKFLRRVDTALKNIGIDERVPYNAPLIQFSSWMGGDRDGNPRVTPEVTRDVCLL ARMMAANLYYSQIEDLMFALSMWRCNDDCRVRSDEVLHSSSRRDAKHYIEFWKQVPPNEPYRVILGDVRDKLYQTPERARQLLAHGYSEIP EEATYTNIEQFLEPLELCYRSLCACGDRSIADGSLLDFLRQVSTPGLSLVRLDIRQESDRHTDVLDAITQHLEIGSYREWSEERKQEWLLS ELSGKRPLFGPDLPKTEEIADVLDTFHVIAELPADCFGAYIISMATAPSDVLAVELLQRECRVKQPLRVVPLFEKLADLDAAPAAVARLFS VEWYRNRINGKQEVMIGYSDSGKDAGRLSAAWQLYKAQEELVKVAKEYGVKLTMFHGRGGTVGRGGGPTHLAILSQPPDTIHGNLRVTVQG EVIEQCFGEEEHLCFRTLQRFTAATLEHGMHPPVSPKPEWRALMDEIAVVATEKYRSIVFKEPRFVEYFRLATPELEYGRMNIGSRPSKRKP SGGIESLRAIPWIFAWTQTRFHLPVWLGFGAAFKYAINKDIKNLRMLQEMYNAWPFFRVTIDLVEMVFAKGNPSIAALYDKLLVSEDLWSF GEQLRSNYEETKSLLLQIAGHKDLLEGDPYLRQRLRLRDSYITTLNVCQAYTLKRIRDPNYSVTPRPHISKEYMESKPAAELVKLNPTSEY APGLEDTLILTMKGIAAGMQNTG	6.90%
5	gil584794	Proton pump 1	Niben101Scf00593g01002.1	MGEEKPEVLDAVLKEAVDLENIPIEEVFENLRCTKEGLTATAAQERLAIFGYNKLEEKKDSKLLKFLGFMWNPLSWVMEAAAIMAIALANG GGKPPDWQDFVGIITLLIINSTISFIEENNAGNAAAALMARLAPKAKVLRDGRWKEEDAAVLVPGDIISIKLGDIIPADARLLEGDPLKID QSALTGESLPVTKGPGDGVYSGSTCKQGEIEAIVIATGVHTFFGKAAHLVDSTNQVGHFQKVLTAIGNFCICSIAVGMIIEIIVMYPIQHR AYRPGIDNLLVLLIGGIPIAMPTVLSVTMAIGSHRLAQQGAITKRMTAIEEMAGMDVLCSDKTGTLTLNKLTVDKNLIEVFAKGVDADMVV LMAARASRTENQDAIDAAIVGMLADPKEARAGIREIHFLPFNPTDKRTALTYLDGEGKMHRVSKGAPEQILNLAHNKSDIERRVHAVIDKF AERGLRSLGVAYQEVPEGRKESAGGPWQFIGLLPLFDPPRHDSAETIRRALNLGVNVKMVTGDQLAIGKETGRRLGMGTNMYPSSALLGQT KDESISALPIDELIEKADGFAGVFPEHKYEIVKRLQARKHICGMTGDGVNDAPALKKADIGIAVDDATDAARSASDIVLTEPGLSVIISAV LTSRAIFQRMKNYTIYAVSITIRIVLGFMLLALIWKFDFPPFMVLIIAILNDGTIMTISKDRVKPSPLPDSWKLAEIFTTGIVLGGYLAMM TVIFFWAAYKTNFFPHVFGVSTLEKTATDDFRKLASAIYLQVSIISQALIFVTRSRSWSFVERPGFLLVIAFVIAQLVATLIAVYANWSFA AIEGIGWGWAGVIWIYNLVFYIPLDIIKFFIRYALSGRAWDLVFERRIAFTRKKDFGKEQRELQWAHAQRTLHGLQVPDTKLFSEATNFNE LNQLAEEAKRRAEIARLRELHTLKGHVESVVKLKGLDIETIQQAYTV	5.30%
4	gil76556492	putative chloroplast cysteine synthase 1 precursor	Niben101Scf05270g01002.1	DEEEIEVHFICPFLLNYVDVSLAYCCRGPIKLCCSQNQPLKVEGLNIAEDVTQLIGNTPMVYLNTIVKGCVANIAAKLEIMEPCCSVKDRIG FSMISDAEEKGLISPGKTVLVEPTSGNTGIGLAFIAASRGYKLILTMPASMSLERRVLLKAFGAELVLTDPAKGMKGAVSKAEEILNNTP DAYILQQFDNPANPKIHYETTGPEMWEDTKGKIDILVAGIGTGGTISGAGRFLKEQNPNIKIIGVEPTESNVLSGGKPGPHKIQGIGAGFI PGNLDQDVMDEVIEISSDEAVETAKQLALQEGLLVGISSGAAALAAIQVGKRPENAGKLIAVVFPSFGERYLSTILFQSIREECEKMQPES	15.40%
4	gil7672161	FtsZ1-1	Niben101Scf03107g01011.1	MATISNPAEIAASSPSFAFYHSSFIPKQCCFTKARRKSLCKPQRFSISSSFTPFDSAKIKVIGVGGGGNNAVNRMIGSGLQGVDFYAINTD AQALLQSAAENPLQIGELLTRGLGTGGNPLLGEQAAEESKEAIANSLKGSDMVFITAGMGGGTGSGAAPVVAQIAKEAGYLTVGVVTYPFS FEGRKRSVQALRAIEKLQKNVDTLIVIPNDRLLDIADEQTPLQDAFLLADDVLRQGVQGISDIITIPGLVNVDFADVKAVMKDSGTAMLGV GVSSSKNRAEEAAEQATLAPLIGSSIQSATGVVYNITGGKDITLQEVNRVSQVVTSLADPSANIIFGAVVDERYNGEIHVTIIATGFTQSF QKTLLSDPRGAKLADKGPVIQESMASPVTLRSSTSPSTTSRTPTRRLFF	12.60%
4	gil298548969	Unnamed protein product	Niben101Scf10132g01021.1	MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNASIRSFASTTASSKIRVENPIVEMDGDEMTRVI WTMIKEKLIYPYLELDTKYYDLGILNRDATDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMWRSPNATIRNILNGTVFREPIL CKNVPRIVPGWKKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYDFKGPGVALAMYNVDQSIRAFAESSMSMAFSKKW PLYLSTKNTILKKYDGRFKDIFEEVYEEKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMTSVLL SSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSVASIFAWARGLGHRAQLDGNQKLSEFVHALEAACVGTIESGKMTKDLAILVHGPKVSR EHYLNTEEFIDPVAQKLQEKLGACAVV	9.50%

Table 2 Continued.

Unique peptides	Gene Identifier	Description	N.bentha sequence	PROTEIN SEQUENCE, UNIQUE REFERENCE	Protein Coverage
4	gil304273262	Mitochondrial glycine decarboxylase complex P-protein-like protein	Niben101Scf03839g04019.1	MCGMKIVAVGTDAKGNINIEELRKAAEANKDXLXALMVTYPSTHGVYEEGIDEICKIIHDNGGQVYMDGANMNAQVGLTSPGFIGADVCHL NLHKTFCIPHGGGGPGMGPIGVKKHLAPFXPSHPVVPTGGIPSPDKSEPLGTISAAPWGSALILPISYTYIAMMGSKGLTDASKIAILNAN YMAKRLEKHYPVLFRGVNGTCAHEFIIDLRGFKNTAGIEPEDVAKRLIDYGFHGPTMSWPVPGTLMIEPTESETLISIREEIAEIEKGKAD IHNNVLKGAPHPPSLLMGDAWTKPYSREYAAFPASWLRVAKFWPSTGRVDNVYGDRNLTCTLLSVSQTVEEQAAATA	11.40%
3	gil153793260	chilling-responsive protein	Niben101Scf14210g00020.1	MSSSENPEIVERVFGDKEKEEKEDKKDEQKGGFIEKVKDFIQDIGEKIEETIGFGKPTADVTEIHIPHINLKKAEIVVDVLVKNPNPVPIP LIDINYLIDSDGRKLVSGLIPDAGTIHAHGSETVKIPVNLVYDDIKNTYHDIQPGSIIPYRIKVDLIVDVPVFGRLTLPLEKTGEIPIPYK PDIDLEKIHFERFSFEETVAVLKLKLENKNNFDLALNSLDYDLWLSDVNVGGAELEKSAKLEKNGISYIDIPITFRPKDFGSALWDMIRGR GTGYTMKGNINVDTPFGAMKLPISKGGGGTTRLKKNKEDGGDDDEDED	13.10%
3	gil189306755	epoxide hydrolase	Niben101Scf00640g04023.1	VADAGFRAIAPDFRGYGLSELPAEPEKATFRDLVDDLLDMLDSLGIHQVFLVGKDFGARVAYHFALVHSDRVSAVVTLGVPFLLTGPETFP RDLLPNGFYMLRWQEAGRAEKDFGRFDTKTVVKNIYIMFSGSELPVAKDDEEIMDLVDPSAPLPDWFTEQDLANYASLYEKSSFRTALQVP YRAWLEEYGVKDIKVKVPCLLVMGEKDYSLKFGGLBQYVKSGMVKEYVPNLETIFLPEGSHFVQEQFPEQVNQLIIIFLKKLI	11.30%
3	gil307940738	G-strand specific single-stranded telomere-binding protein 1	Niben101Scf02797g00005.1	MGSRKNDFGDGASPGKIFIGGLAKETTLEQFVKYFEKYGEIIDSVIMKDRHTGRPRGFGFITYSDPTVVDTVIAETHIINDKQVEIKRTIP KGSAESKDFKTKKIFVGGIPTSMNEDEFKGFFSKYGKVVDCEIIRDHVSKRSRGFGFIVFDNELVVDNVLSEGNMIDMLGTQVEIKKAEPK KPSNPASSGHGYGSESSRGRGYSDSYGGFGNSYSSFGSGGFGPASYRSSGGMGGRFGDYGYGGGEYGGRYGDFGGGEFGGYRGDPSLGYSSR YGSYAGGFGGGAGGYGGSGLMAYGRGASYGGYGGAGSGAGYDSGPGAGYGGPGGLYGSRAGYSGSGRYHPYAR	16.50%
3	gil6179940	DnaJ-like protein	Niben101Scf11137g01016.1	MGVDYYKVLGVDKNATDDDLKKAYRKLAMKWHPDKNPNNKKAAEAKFKQISEAYDVLSDSQKRAVYDQYGRDGLKGGVPPPGAGGPGGGSP FFSTGEGPQSFRFNTRSADDIFAEFFGFSSPFGGAGGRGPRFGGMFGDDMFASFGEGGGGGGASMYQSAPRKEAPIQQNLPCNLEDLYKGT TKKMKISREVADASGKRMQVEEILTINIKPGWKKGTKITFQEKGNEQPGVIPADLVFIIDEKPHRVFSRDGNDLIVTQKISLVEALTGTTV QLTTLDGRNLTIPVNSVIQPNYEHVVPGEGMPLPKDPTKKGNLRIKFDIKFPVRLTTTQKAGIKELLGS	12%
3	gil7417008	cell death associated protein	Niben101Scf05283g00016.1	MVHEKQVIEEVSGWLRVFEDGSVDRTWTGPPEVKFMAEPVPPHDYFIDGVAVKDVVADEKSGSRLRIYLPERNDNSANKLPVILHFHGGGF CVSHADWFMYYTVYTRLARAAKAIIVSVFLPLAPEHRLPAACDAGFAALLWLRDLSRQQGHEPWLNDYADFNRVFLIGDSSGGNIVHQVAV KAGEENLSPMRLAGAIPIHPGFVRSYRSKSELEQEQTPFLTLDMVDKFLGLALPVGSNKDHQITCPMGEAAPAVEELKLPPYLYCVAEKDL IKDTEMEFYEAMKKGEKDVELFINNGVGHSFYLNKIAVRMDPVTGSETEKLCEAVAEFINKH	9.60%
3	gil320148816	S-nitrosoglutathione reductase	Niben101Scf03263g08013.1	EPNKPLVIEDVQVAPPQAGEVRVKVLYTALCHTDAYTWSGKDPEGLFPCVLGHEAAGIVESVGEGVTEVQPGDHVIPCYQAECRECKFCKS GKTNLCGKVRAATGVGVMMNDRQSRFSINGKPIYHFMGTSTFSQYTVVHDVSVAKIDPVAPLEKVCLLGCGVPTGLGAVWNTAKVESGSIV AVFGLGTVGLAVAEGAKAAGASRIIGIDIDSKKFDRAKNFGVTEFINPKEHEKPIQQVIVDLTDGGVDYSFECIGNVSVMRAALECCHKGW GTSVIVGVAASGQEISTRPFQLVTGRVWKGTAFGGFKSRSQVPWLVDKY	9.90%
3	gil53854350	acetoacetyl-coenzyme A thiolase	Niben101Scf04727g03006.1	MAPAAADSIKPRDVCIVGVARTPMGGFLGSLSSLSATKLGSTAIEGALKKANVDPSLVEEVFFGNVLSANLGQAPARQAALGAGIPNTVVC TAVNKVCASGLKATMLAAQSIQLGINDVVVAGGMESMSNVPKYIAEARKGSRLGHDSLVDGMLKDGLTDVYKDCGMGVCAEICAENHKITR EEQDDYAVQSFERGIAAQEAGAFAWEIVPVEVPGGRGKPSIIVDKDDGLGKFDGAKLRKLRPSFKEKDGTVTAGNASSISDGAAALVLVSG EKAIKLGLNVIAKISGYADAAQAPELFTTAPALAIPKAIKSASLEASQIDYYEINEAFAVVSLANQKLLGLNPEKVNVHGGAVSLGHPLGC SGARILVTLLGVLRQKNGKYGAAGVCNGGGGASALVLELV	8.70%
3	gil10798636	Elongation factor 2	Niben101Scf01848g04005.1	DAIHRGGGQVIPTARRVIYASQLTAKPRLLEPVYLVEIQAPEQALGGIYGVLNQKRGHVFEEMQRPGTPLYNIKAYLPVIESFGFSGTLRA ATSGQAFPQCVFDHWDMMSSDPLDAGTQAHQLVLDIRKRKGLKQAVTPLSEFEDKL	31.30%
3	gil115610	Phosphoenolpyruvate carboxylase	Niben101Scf04036g04008.1	MATRSLEKLASIDAQLRALVPGKVSEDDKLVEYDALLLDRFLDILQDLHGEDLKETVQECYELSAEYEGKHDPKKLEELGNVLTSLDPGDS IVIAKAFSHMLNLANLAEEVQIAYRRRQKLKRGDFADENNATTESDIEETFKKLVGDLKKSPQEVFDALKNQTVDLVLTAHPTQSVRRSLL QKHGRIRDCLAQLYAKDITPDDKQELDEALQREIQAAFRTDEIRRTAPTPQDEMRAGMSYFHETIWKGVPKFLRRVDTALKNIGINERLPY NAPLIQFSSWMGGDRDGNPRVTLEVTRDVCLLARMMAANLYYSQIEELMFELSMWRCNDDLRIRAAELYRSSRRDTKHYIEEFWKTIPPSEP YRVILGDVRDKLYQTRERTRQMLAHGISDIPEDATYNNVEQFLEPLELCYRSLCECGDRPIADGSLLDFLRQVSTFGLSFVRLDIRQESDR HTDVLDAITQHLEIGSYREWSEERRQEWLLSELSGKRPLFGPDLPRTEEIADVLDTLHVIAELPSDCFGAYIISMATAPSDVLAVELLQRE CHVKQPLRVVVPLFEKLDDLESASAAVARLFSIEWYRNRINGKQEVMVGYSDSGKDAGRFSAAWQLYKAQEELIKVAKEHGVKLTMFHGRGG TVGRGGGPTHLAILSQPPDTIQGSLRVTVQGEVIEQSFGEEHLCFRTLQRFTAATLEHGMHPPVSPKPEWRALMDEIAVIATEKYRSIVFK EPRFVEYSALATPELEYGRMNIGSRPSKRKPSGGIESLRAIPWIFAWTQTRFHLPVWLGFGAAFKYAIDKDIKNLRMFHEMYNEWPFFRVTI IDLVEMVFAKGNPGIAALYDKLLVSEDLLPFGELLRSNYEETRSLLLQIAGHKDLLEGDPYLKQRLRLRDSYITTLNLLQAYTLKRIRDPN YHVTLRPHISKDYMESKSAAELVQLNPTSEYAPGLEDTLILTMKGIAAGLQNTG	3.90%
3	gil75252690	DNA gyrase subunit A	Niben101Scf17442g00002.1	MKLHTLNPQTPLTQSKPMAFSTGITPSRFSGLRKTSSELRFLSSVTPPPRKQLRPVSARRKEEEVGDEGNGSVILRDRGENEDRNGGERVV LTELHKEATEAYMSYAMSVLLGRALPDVRDGLKPVHRRILYAMHELGLSSKKPYKKCARVVGEVLGKFHPHGDTAVYDSLVRMAQDFSLRS PLIRGHGNFGSIDADPPAAMRYTECRLEALTESMLLADLEQNTVDFVPNFDNSQKEPSLLPARVPNLLLNGASGIAVGMATNIPPHNLGEL VDALSALIHNPEATLQELLEYMPGPDFPTGGIIMGNIGILEAFRTGRGRVVIRGKTDIELLDSKTKRAAIIIQEIPYQTNKASLVEKIADL VENKILEGVSDIRDESDRSGMRIVIELKRGSDPAIVLNNLYRLTALQSSFSCNMVGILNGQPKLMGLKELLQAFLDFRCSVVERRARFKLS QAQERNHIVEGIIVGLDNLDEVINTIRKASSNALAAASLRKEFELSEKQAEAILDISLRRLTALERNKFVEEGKSLRTQISKLEELLSSKK QILQLIEEEAIEIKNKFFNPRRSMLEDTDSGDLEDIDVIPNEEMLLAISEKGYVKRMKPDTFNLQNRGTIGKSVGKLRVNDAMSDFLVCRA HDKVLYFSDKGTVYSSPAYKIPECSRTAAGTPLVQILSLSDGERITSIIPVSEFAADQYLVMLTVNGYIKKVSLNYFASIRCTGIIAIQLV PDDELKWVKCCSNNDFVAMASQNGMVILTPCANIRALGRNTRGSVAMRLKEGDKVASMDIIPDALQKELDKTLEVQQRQYRSMKGPWLLFV SESGYGKRVPVSRFRTSPLNRVGLFGYKFSSEDCLAAVFVVGFSLGEDGESDEQVVLVSQSGTVNRIKVRDISIQSRYARGVILMRLEHAG KIQSASLISAADADPEDEDATAVAA	4.10%

Table 2 Continued.

Unique peptides	Gene Identifier	Description	N.bentha sequence	PROTEIN SEQUENCE, UNIQUE REFERENCE	Protein Coverage
3	gi|298676333	ARGONAUTE 1	Niben101Scf05146g06007.1	MVRKKRTDVPGGAESFESHETGGGRGGAQRPSQQQQHQHQHQQGGGRGWAPQHGGHGGRGGGGAPRGGMGPQQSYGGPPEYYQQGRGTQQY QRGGGQPQRRGGMGGRGAPSGPPRSPVPELHQATQTPHQPVPYGRPSETYSEAGSSSQPPEPTTQQVTQQFQQLVVLPEAAATQAIQPASS KSMRFPLRPGKGSTGIRCIVKANHFFAELPDKDLHQYDVSITPVVSSRGVNRAVMEQLVKLYRESHLGKRLPAYDGRK**SLYTAGPLPFVQK** DFKITLIDDDDGPGGASCRREREFKVVIKLAARADLHHLGMFLQGRQADAPQEALQVLDIVLRELPTSRYCPVGRSFYSPHLGRRQPLGEG LESWRGFYQSIRPTQMGLSLNIDMSSTAFIEPLPIIDFVSQLLNRDISSRPLSDADRVKIKKALRGVKVEVTHRGNMRRKYR**ISGLTSQAT** **R**ELTFPVDERGTMKAVVEYFRETYGFVIRHTQLPCLQVGNTQRPNYLPMEVCKIVEGQRYSKRLNERQITALLKVTCQRPQEREHDILQTV HHNAYADDPYAKEFGIKISEK**LAQVEAR**VLPAPWLKYHDTGREKDCLPQVGQWNMMNKKMVNGGTVNNWICVNFSRNVQDTVARGFCSELA QMCMISGMNFNPNPVLPPVSARPDQVERVLKTRFHDAMTNLQPHGRELDLLIVILPDNNGSLYGDLKRICETELGIVSQCCLTKHVFKMSK QYLANVSLKINVKVGGRNTVLVDALSRRIPLVSDRPTIIFGADVTHPHPGEDSSPSIAAVVASQDWPEITKYAGLVSAQAHRQELIQDLYK TWQDPVRGPVTGGMIKELLISFRRATGQKPQRIIFYRDGVSEGQFYQVLLFELDAIRKACASLEPNYQPPVTFVVVQKRHHTRLFANNHRD RNAVDRSGNILPGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLWDENNFTADALQSLTNNLCYTYARCTRSVSIVPPAYYAHLAAF RARFYMEPETSDSGSVTSAAASNRGGVGAMGRSTRAPGAGAAVRPLPALKENVKRVMFYC	2.80%
3	gi|84688908	AGO1-2	Niben101Scf08137g02022.1	MVRKRRTELPGSGESSGSQETGGQGRGQHPQQLHQATSQTPYQTAMTTQPIPYARPTETSSEAGSSSQPPEQAALQVTQQFQQLALQQEAA TTQAVPPASSKLLRFPLRPGKGSNGMRCIVKANHFFAELPDKDLHQYDVTISPEVSSRGVNR**AVMAQLVK**LYQESHLGKRLPAYDGRKSLY [illegible] MEPETSDGGSVTSGAAGGRGGGAGAAGRNTRAPSAGAAVRPLPALKDNVKRVMFYC	3.40%
2	gi|111162651	chloroplast aldolase	Niben101Scf04664g01003.1	**SAAYYQQGAR**FAKWRTVVSIPNGPSALAVKEAAWGLARYAAISQDSGLVPIVEPEILLDGEHGIDRTFEVAQKVWAEVFFYLAENNVMFEG ILLKPSMVTPGAECKDR**ATPQQVADYTLSLLK**RRIPPPVPGIMFLSGGQSEVEATLNLNAMNQAPNPWHVSFSYARALQNTCLKTWGGQPE NVKAAQDALLTRAKANSLAQLGKYTGEGES	11.80%
2	gi|1172836	GTP-binding nuclear protein Ran-B1	Niben101Scf08341g01001.1	MALPNQQTVDYPSFKLVIVGDGGTGKTTFVKR**HLTGEFEK**KYEPTIGVEVHPLDFFTNCGKIRFYCWDTAGQEKFGGLRDGYYIHGQCAII MFDVTARLTYKNVPTWHRDLCRVCENIPIVLCGNKVDVKNRQVKAKQVTFHRKKNLQYYEISAK**SNYNFEKPFLYLAR**KLAGDANLHFVES PALAPPEVQIDLAAQQLHEQELLQAAAHALPDDDDEAFE	10%
2	gi|170277	O-methyltransferase	Niben101Scf00225g00008.1	PHVIGDAPAYPGVEHVGGDMFASVPK**ADAIFMK**WICHDWSDEHCLKFLKNCYEALPANGKVIIAECILPEAPDTSLATKNTVHVDIVMLAH NPGGKER**TEKEFEALAK**GAGFTGFARLVALTTLGSWNSTNN	12.90%
2	gi|40549128	beta-cyanoalanine synthase	Niben101Scf05118g08013.1	KQEMMQPTSSIKDRPAFAMINDAEKKGLISPGETTLIEPTSGNMGISMAFMAAMKGYK**MVLTMPSYTSLER**RVCMR**AFGADLVITDPTK**GM GGTVKKAYDLLESTPNAFMLQQFSNPANTQV	21.30%
2	gi|48375044	putative mitochondrial malate dehydrogenase	Niben101Scf13540g04030.1	TPGVAADVSHINTRSQVSGFAGDEQLRQALEGADVVIIPAGVPRKPGMTR**DDLFNINAGIVK**SLCTAIAKYCPHALVNMISNPVNSTVPIA AEVFKKAGTYDEKRLFGVTTLDVVRAKTFYAGKAKVNVADVIVPVVGGHAGITILPLFSQATPKANLGDEEIEALTKR**TQDGGTEVVEAK**A GKGSATLSMAYAGAIFADACLKGLNGVQML	11.30%
2	gi|6685068	FtsZ-like protein 2	Niben101Scf07163g00020.1	MATMLGLSSNTGIDILSSSSNSLSFYHSTRFTQCFSPKSLCKRQRRRFSICSSLSSAKIKVVGVGGGGNNAVNRMIGSGLQGVDFYAVNTD AQALLQSTVENPIQIGELLTRGLGTGGNPLLGEQAAEESKEHIANALKGSDMVFITAGMGGGTGSGAAPVVAQIAKEAGYLTVGVVTYPFS FEGRKRSLQALEAIEKLQK**NVDTLIVIPNDR**LLDIADEQTPLQNAFLLADDVLCQGVQGISDIITIPGLVNVDFADVKAIMKDSGTAMLGV GVSSSRNRAEEAAEQATLAPLIGSSIQSATGDVYNITGGK**DITLQEVNK**VSQVVTSLADPSANIIFGAVVDERYNGEIQVTLIATGFAQSF QNSLLTDPRGAKLVDKSKGTTERTVSPDTLRSSESPSTKPRPATRRLFF	5.10%
2	gi|76262913	cyc07	Niben101Scf01249g06003.1	MAVGKNKRISKGKKGGKKAADPYAKKDWYDIKAPSVFDIKNVGKTLVTRTQGTKIASEGLKHRVFEVSLADLQKDEDQSFRKIRLRAEDV QGKNVLTNFWGMDFTTDKLRSLVKKWQTLIEAHVDVK**TTDSYTLR**MFCIAFTKKRPNQQKRTCYAQSSQIRQIRRKMVEIMRNQASSCDLK ELVAKFIPESIGREIEK**ATSSIFPLQNVYIR**KVKILKAPKFDIGKLMEVHGDYSEDVGVKLDRPADETVAEAEPEIPGA	8.40%
2	gi|78059504	RPN8	Niben101Scf04077g02006.1	YAVPFEEDDRDPSIWFLDHNYHESMFSMFRRINAKEHVVGWYSTGPKLRENDLNIHGLFNDYVPTPVLVIIDVQPKELGIPTK**AYYAVEEV** **K**ENATQKSQKVFAHVPSEIAAHEVEVIGVEHLLRDVK**DTTISTLATEVTGK**LAALKGLDARLQEIRSCLDLVIDGKLPLNHEILYHLQDVF NLLPNLNVAELVKAFSVKTNDMMLVIYLSSLIRSVIALHNLINNKML	10%
2	gi|12643806	Glutamate dehydrogenase B	Niben101Scf09000g00009.1	MNALAATNRNFKLASRLLGLDSKLEQCLLIPFREIKVECTIPKDDGSLATFIGFRVQHDNARGPMKGGIRYHPEVDPDEVNALAQLMTWKT AVANIPYGGAKGGIGCSPSDLSISELERLTRVFTQKIHDLIGVHTDVPAPDMGTNPQTMAWILDEYSKFHGYSPAVVTGKPIDLGGSLGRD AATGRGVLFAAEALLRDHGKSIAGQRFVVQGFGNVGSWAAQLITEQGGK**IVAVSDITGAIK**NKNGIDIASLLKHVKENRGVKGFHGADSID PNSILVEDCDVLIPAALGGVINRDNAKDIKAKFIVEAANHPTDPEADEILAKKGVVILPDIYANSGGVTVSYFEWVQNIQGFMWDEERVNT ELKAYMNRGFKDVKDMCKTHNCDLR**MGAPTLGVNR**VARATTLRGWEA	5.40%
2	gi|2500521	ATP-dependent RNA helicase eIF4A-15	Niben101Scf02031g00026.1	MAGAAPEGSQFDARQFDTKMNELLSADGQDFFTSYDEVYDSFDAMGLQENLLRGIYAYGFEKPSAIQQRGIVPFCKGLDVIQQAQSGTGKT ATFCSGILQQLAYELLECQALVLAPTRELAQQIEKVMRALGDYLGVKVTACVGGTSVREDQRILSSGVHVVVGTPGRVFDMLRRQSLRPDH IK**MFVLDEADEMLSR**GFKDQIYDIFQLLPPKIQVGVFSATMPPEALEITRKFMNKPVRILVKRDELTLEGIKQFYVNVDKEEWKLETLCDL YETLAITQSVIFVNTRRKVDWLTDKMRSRDHTVSATHGDMDQNTRDIIMREFRSGSSR**VLITTDLLAR**GIDVQQVSLVINFDLPTQPENYL HRIGRSGRFGRKGVAINFVTKDDERMLSDIQRFYNVVIEELPANVADLL	5.60%

Table 2 Continued.

Unique peptides	Gene Identifier	Description	N.bentha sequence	PROTEIN SEQUENCE, UNIQUE REFERENCE	Protein Coverage
2	gil6715512	Vacuolar H+-ATPase B subunit	Niben101Scf01486g03005.1	MGVAPNNIEMEEGNLEVGMEYRTVSGVAGPLVILDKVKGPKYQEIVNIRLGDGTTRRGQVLEVDGEKAVVQVFEGTSGIDNKYTTVQFTGE VLKTPVSLDMLGRIFNGSGKPIDNGPPILPEAYRDISGSSINPSERTYPEEMIQTGISTVDVMNSIARGQKIPLFSAAGLPHNEIAAQICR QAGLVKRLEKSDNLLEGGEGDNFAIVFAAMGVNMETAQFFKRDFEENRSMERVTLFLNLANDPTIERIITPRIALTTAEYLAYECGKHVLV ILTDMSSYADALREVSAAREEVPGSGGYPGYMYTDLATIYERAGRIEGRTGSITQIPILTMPNDDITHPTPDLTGYITGRQIYIDRQLHNR QIYPPINVLPSLSRLMKSAIGEGMTRRDHSDVSNQLYANYAIGKDVQAMKAVVGEEALSSEDLLYLEFLDKFERKFVAQGAYDTRNIFQSL DLAWTLLRIFPRLHRIPAKTLDQYYSRDASN	3.70%
2	gil119866037	Impa2	Niben101Scf13164g00010.1	MSLRPSARTEVRRNRYKVAVDAEEGRRRREDNMVEIRKSKREESLLKKRREGLQPQQQPFPANLHTSTVEKKLESLPSMVAGVWSNDNNLQ LEATTQFRKLLSIERSPPIEEVIQSGVVPRFVEFLMREDFPQLQFEAAWALTNIASGTSDNTRVVIDHGAVPIFVKLLGSPSDDVREQAVW ALGNVAGDSPRCRDLVLSNGALIPLLAQLNEHTKLSMLRNATWTLSNFCRGKPQPPFEQVRPALSALQRLVHSNDEEVLTDACWALSYLSD GTNDKIQAVIEAGVCPRLVELLMHPSPSVLIPALRTVGNIVTGDDLQTQCIIEHGALTCLLSLLTHNHKKSIKKEACWTISNITAGNKEQI QAVIEAGLIAPLVNLLQTAEFDIKKEAAWAISNATSGGTHEQIKFLVSQGCIKPLCDLLVCPDPRIVTVCLEGLENILKVGEAEKANTGGI NYYAQLTDDAEGLEKIENLQSHDNNEIYEKAVKILETYWLEEEDETLPAGDEAQAGFNFGGNDIQLPSGGFKFG	4.70%
2	gil1841462	Elongation factor 2	Niben101Scf01848g04005.1	NEKEVEAHPIRAMKFSVSPVVRVAVQCKLASDLPKLVEGLKRLAKSDPMVLCSIEESGEHIIAGAGELHLEICLKDLQDDFMGGAEIIVSD PVVSFRETVLEKSCRTVMSKSPNKHNRLYMEARPMEEGLARAIDDGRIGPRDDPKVRSKILSEEFGWDKDLAKKIWCFGPETTGPNMV	19%
2	gil296529814	Chaperone protein ClpB IPR001270	Niben101Scf04410g09016.1	MNPEKFTHKTNEALAGALELALSAGHAQFTPLHMAVALISDHNGIFRQAIVNAGGNEEVANSVERVLNQAMKKLPSQTPAPDEIPPSTSLI KVLRRAQSSQKSRGDSHLAVDQLILGLLEDSQIGDLLKEAGVSASRVKSEVEKLRGKEGRKVESASGDTTFQALNTYGRDLVEQAGKLDPV IGRDEEIRRVVRILSRRTKNNPVLIGEPGVGKTAVVEGLAQRIVRGDVPSNLADVRLIALDMGALVAGAKYRGEFEERLKAVLKEVEEAEG KVILFIDEIHLVLGAGRTEGSMDAANLFKPMLARGQLRCIGATTLEEYRKYVEKDAAFERRFQQVYVAEPSVADTISILRGLKERYEGHHG VKIQDRALVVAAQLSSRYITGRHLPDKAIDLVDEACANVRVQLDSQPEEIDNLERKRIQLEVELHALEKEKDKASKARLIEVRKELDDLRD KLQPLMMRYKKEKERIDELRRLKQKRDELIYALQEAERRYDLARAADLRYGAIQEVETAIANLESTSAESTMLTETVGPDQIAEVVSRWTG IPVSRLGQNEKEKLIGLGDRLHQRVVGQDHAVRAVAEAVLRSRAGLGRPQQPTGSFLFLGPTGVGKTELAKALAEQLFDDDKLMIRIDMSE YMEQHSVSRLIGAPPGYVGHDEGGQLTEAVRRRPYSVVLFDEVEKAHPAVFNTLLQVLDDGRLTDGQGRTVDFTNSVIIMTSNLGAEYLLS GLMGKCTMETAREMVMQEVRKQFKPELLNRLDEIVVFDPLSHKQLRQVCRYQMKDVALRLAERGIALGVTEAAPDVILTESYDPVYGARPI RRWLERKVVTELSKMLVKEEIDENSTVYIDAGVSGKDLTYRVEKNGGLVNAATGQKSDILIQLPNGPRSDAVQAVKKMRIEEIEDDEMED	2.30%
2	gil3914361	Phosphatidylcholine-hydrolyzing phospholipase D1	Niben101Scf12266g06004.1	MAQILLHGTLHVTIYEVDNLQKEGGGHFFSKIKEHVEETIGFGKGTPAIYATVDLEKARVGRTRKIKNEPNNPRWYESFHIYCAHMASNVI FTVKDDNPIGATLIGRAYVPVEELLEGEEIDKWVEILDREMNPIAEGSKIHVKLQFFDVSRDPNWERGIRSSKYPGVPYTFFAQRTGCRVS LYQDAHVPDNFIPKIPLSGGKYYEPHRCWEDIFDAIINAKHLIYITGWSVYTEITLVRDSRRQKPGGDITLGELLKKKASEGVKVLMLVWD DRTSVGLLKKDGLMATHDQETEQFFQGTEVNCVLCPRNPDDGGSIVQSLQIGTMFTHHQKIVVVDSELPSGESEKRRILSFVGGIDLCDGR YDTPFHSLFRTLDTAHHDDFHQPNFPDGAITKGGPREPWHDIHSRLEGPIAWDVLFNFEQRWRKQGGKDVLVNFRELDDIIIPPSPVMHLLD DSETWNVQLFRSIDEGAAFGFPETPEDAAKAGLVSGKDNIIDRSIQDAYIHAIRRAKNFIYIENQYFLGSSYDWQSDDIKVEDIGALHVIP KELALKIVSKIEAGERFTVYVVVPMWPEGIPESASVQAILDWQRRTMEMMYKHIVQALNAKGIEEDPRNYLTFFCIGNREVKKSGAYEPSE TPEPDSDYIRAQEARRFMIYVHSKMMIVDDEYIIVGSANINQRSMDGARDSEIAMGAYQPHHLATREPARGQIHGFRMALWYEHLGMLDET FLHPESEECVSKVNRMADKYWDLYSSESLERDLPGHLLRYPIGVASEGDVTELPGAEHFPDTKARVLGTKSDYLPPILTT	3%
2	gil122063400	S-adenosyl-L-homocysteine hydrolase	Niben101Scf10608g02009.1	MALLVEKTTSGREYKVKDMSQADFGRLEIELAEVEMPGLMACRTEFGPSQPFKGAKITGSLHMTIQTAVLIETLTALGAEVRWCSCNIFST QDHAAAAIARDSAAVFAWKGETLQEYWWCTERALDWGPGGGPDLIVDDGGDATLLIHEGVKAEEEFAKNGTIPDPNSTDNAEFQLVLTIIK ESLKTDPLKYTKMKERLVGVSEETTTGVKRLYQMQANGTLLFPAINVNDSVTKSKFDNLYGCRHSLPDGLMRATDVMIAGKVALVAGYGDV GKGCAAALKQAGARVIVTEIDPICALQATMEGLQVLTLEDVVSDVDIFVTTTGNKDIIMVDHMRKMKNNAIVCNIGHFDNEIDMLGLETYP GVKRITIKPQTDRWVFPDTNSGIIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLELWNEKSSGKYEKKVYVLPKHLDEKVAALHLGKL GAKLTKLSKDQADYISVPVEGPYKPAHYRY	3.70%
2	gil1705613	Salicylic acid-binding protein	Niben101Scf14996g00009.1	MDLSKFRPSSAYDSPFLTTNAGGPVYNNVSSLTVGPRGPVLLEDYHLIEKLATFDRERIPERVVHARGASAKGFFEVTHDISHLTCADFLR APGVQTPVICRFSTVVHERGSPESLRDIRGFAVKFYTREGNFDLVGNNVPVFFNRDAKSFPDTIRALKPNPKSHIQEYWKILDFFSFLPES LHTFAWFFDDVCLPTDYRHMEGYGVHAYQLINKAGKAHYVKFHWKPTCGVKCMSEEEAIRVGGTNHSHATKDLYDSIAAGNYPEWKLFIQI MDTEDVDKFDFDPLDVTKTWPEDILPLMPVGRLVLNRNIDNFFAENEQLAFNPGHIVPGLYYSEDKLLQTRIFAYADTQRHRIGPNYMQLP VNAPKCAHNNHRDGAMNFMHRDEEVDYLPSRFDPCRHAEQYPIPSRVLTGRREMCVIEKENNFKQAGERYRSWEPDRQDRYVSKWVEHLS DPRVTYEIRSIWISYLSQADKSCGQKVASRLTLKPTM	4.90%
2	gil347453937	PSII 47kD protein	Niben101Scf01208g01014.1	MGLPWYRVHTVVLNDPGRLLSVHIMHTALVAGWAGSMALYELAVFDPSDPVLDPMWRQGMFVIPFMTRLGITNSWGGWSITGGTVTNPGIW SYEGVAGAHIVFSGLCFLAAIWHWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLSGVACFGFGAFHVTGLYGPGIWVSDPYGLTGKVQPVN PAWGVEGFDPFVPGGIASHHIAAGTLGILAGLFHLSVRPPQRLYKGLRMGNIETVLSSSIAAVFFAAFVVAGTMWYGSATTPIELFGPTRY QWDQGYFQQEIYRRVSAGLAENQSLSEAWSKIPEKLAFYDYIGNNPAKGGLFRAGSMDNGDGIAVGWLGHPIFRDKEGRELFVRRMPTFFE TFPVVLVDGDGIVRADVPFRRAESKYSVEQVGVTVEFYGGELNGVSYSDPATVKKYARPAQLGEIFELDRATLKSDGVFRSSPRGWFTFGH ASFALLFFFGHIWHGARTLFRDVFAGIDPDLDAQVEFGAFQKLGDPTTKRQAA	4.70%
2	gil347453944	Ribosomal protein S11	Niben101Scf00568g04020.1	MAKAIPKISSRRNGRIGSRKGARRIPKGVIHVQASFNNTIVTVTDVRGRVVSWSSAGTSGFKGTRRGTPFAAQTAAANAIRTVVDQGMQRA EVMIKGPGLGRDAALRAIRRSGILLTFVRDVTPMPHNGCRPPKKRRV	15.90%
2	gil38488580	heat shock protein 70	Niben101Scf04364g01014.1	WFEKIASKLPEADKKKIEDAIESAIQWLDANQLAESDEFEDKMKGLESICNPIIAKMYQGAGGDMGGAMDDDAPAPSGGSGAGPKIEEVD	40%

Table 3. Identification of TrAP interacting proteins by IP-MS in *Arabidopsis thaliana* seedlings.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
31	156	AT5G13960	KRYPTONITE, KYP, SDG33, SET DOMAIN PROTEIN 33, SU(VAR)3-9 HOMOLOG 4, SUVH4	MAGKRKRANA PDQTERRSSV RVQKVRQKAL DEKARLVQER VKLLSDRKSE ICVDDTELHE KEEENVDGSP KRRSPPKLTA MQKGKQKLSV SLNGKDVNLE PHLKVTKCLR LFNKQYLLCV QAKLSRPDLK GVTEMIKAKA ILYPRKIIGD LPGIDVGHRF FSRAEMCAVG FHNHWLNGID YMSMEYEKEY SNYKLPLAVS IVMSGQYEDD LDNADTVTYT GQGGHNLTGN KRQIKDQLLE RGNLALKHCC EYNVPVRVTR GHNCKSSYTK RVYTYDGLYK VEKFWAQKGV SGFTVYKYRL KRLEGQPELT TDQVNFVAGR IPTSTSEIEG LVCEDISGGL EFKGIPATNR VDDSPVSPTS GFTYIKSLII EPNVIIPKSS TGCNCRGSCT DSKKCACAKL NGGNFPYVDL NDGRLIESRD VVFECGPHCG CGPKCVNRTS QKRLRFNLEV FRSAKKGWAV RSWEYIPAGS PVCEYIGVVR RTADVDTISD NEYIFEIDCQ QTMQGLGGRQ RRLRDVAVPM NNGVSQSSED ENAPEFCIDA GSTGNFARFI NHSCEPNLFV QCVLSSHQDI RLARVVLFAA DNISPMQELT YDYGYALDSV MGPDGKVKQL ACYCGALNCR KRLY	37.02%
13	50	TGMV-AL2	Flag-Myc4-TrAP	MDYKDDDDKH EQKLISEEDL EQKLISEEDL EQKLISEEDL EQKLISEEDL GSRRAKLSTS LYKKAGSAAA VLEENLYFQG SFTHAMRNSS SSTPPSIKAQ MRAAKRRAIR RRRIDLNCGC SIYIHIDCRN NGFTHRGTYH CASSREWRLY LGDNKSPLFQ DNQRRGSPLH QHQDIPLTNQ VQPQPEESIG SPQGISQLPS MDDIDDRFWE NLFK	36.92%
11	17	AT5G11710	EPSIN1	MDFMKVFDQT VREIKREVNL KVLKVPEMEQ KVLDATDNEP WGPHGTALAE IAQATKKFSE CQMVMSVLWT RLSETGKDWR YVYKALAVID YLISNGSERA VDEIIEHTYQ ISSLTSFEYV EPNGKDVGIN VRKKAENIVA LLNNKEKISE IRDKAVANRN KYVGLSSTGI TYKSGSSASF GGSFQSGSSN FDSYKDRDSR EDKNDYESFQ KSRRGVKTEE QSYTSKKSFS RYGSTDHDNL SSGKKSPDSA KHRSYVSAAP SNNDDDFDDF DPRGTSSNKP STGSANQVDL FGGDLIGDFL DSGPTETSST NNNENFQEAD LFADAAFVSA SAQGAEPGSQ TQKQVDLFSA SEPSVTVSSA PPTVDLFASS ESVVSPEAKI SIPESMATPN IVDPFAAVPM DNFDGSDPFG AFTSHSSASVS TGPQAPSVHG SATNTTSPLS FADSKPQHLQ KKDPFQVKSG IWADSLSRGL IDLNITAPKK ASLADVGVVG DLSNEDGNKA SAASYYSGWS MGAGSGLGKT GLYSTQQQQQ QQQQQQAPDI SDDFFSSLSN QRVQSGGFKQ	26.61%
9	19	AT4G20360	ATRAB8D, ATRABE1B, RAB GTPASE HOMOLOG E1B, RABE1B	MAISAFPAAC5 SSSRILCSYS SPSPSLCPAI STSGKLKTLT LSSSFLPSYS LTTTSASQST RRSFTVRAAR GKFERKKPMV NIGTIGHVDH GKTTLTAALT MALASIGSSV AKKYDEIDAA PEERARGITI NTATVEYETE NRHYAHVDCP GHADYVKNMI TGAAQMDGAI LVVSGADGPM PQTKEHILLA KQVGVPDMVV FLNKEDQVDD AELLELVELE VRELLSSYEF NGDDIPIISG SALLAVETLT ENPKVKRGDN KWVDKIYELM DAVDDYIPIP QRQTELPFLL AVEDVFSITG RGTVATGRVE RGTVKVGETV DLVGLRETRS YTVTGVEMFQ KILDEALAGD NVGLLLRGIQ KADIQRGMVL AKPGSITPHT KFEAIIYVLK KEEGGRHSPF FAGYRPQFYM RTTDVTGKVT KIMNDKDEES KMVMPGDRVK IVVELIVPVA CEQGMRFAIR EGGKTVGAGV IGTILE	20.17%
7	9	AT3G26650	GAPA-1, GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE A SUBUNIT, GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE A SUBUNIT 1	MASVTFSVPK GFTEFSGLRS SSASLPFGKK LSSDEFVSIV SFQTSAMGSS GGYRKGVTEA KLKVAINGFG RIGRNFLRCW HGRKDSPLDI IAINDTGGVK QASHLLKYDS TLGIFDADVK PSGETAISVD GKIIQVVSNR NPSLLPWKEL GIDIVIEGTG VFVDREGAGK NIEAGAKKVI ITAPGKGDIP TVVVGVNADA YSHDEPIISN ASCTTNCLAP FVKVLDQKFG IIKGTMTTTH SYTGDQRLLD ASHRDLRRAR AAALNIVPTS TGAAKAVALV LPNLKGKLNG IALRVPTPNV SVVDLVVQVS KKTFAEEVNA AFRDSAEKEL KGILDVCDEP LVSVDFRCSD FSTTIDSSLT MVMGDDMVKV IAWYDNEWGY SQRVVDLADI VANNWK	24.75%
7	10	ATCG00490	RBCL	MSPQTETKAS VGFKAGVKEY KLTYYTPEYE TKDTDILAAF RVTPQPGVPP EEAGAAVAAE SSTGTWTTVW TDGLTSLDRY KGRCYHIEPV PGEETQFIAY VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT KDDENVNSQP FMRWRDRFLF CAEAIYKSQA ETGEIKGHYL NATAGTCEEM IKRAVFARELGVPIVMHDYL TGGFTANTSL SHYCRDNGLL LHIHRAMHAV IDRQKNHGMH FRVLAKALRL SGGDHIHAGT VVGKLEGDRE STLGFVDLLR DDYVEKDRSR GIFFTQDWVS LPGVLPVASG GIHVWHMPAL TEIFGDDSVL QFGGGTLGHP WGNAPGAVAN RVALEACVQA RNEGRDLAVE GNEIIREACK WSPELAAACE VWKEITFNFP TIDKLDGQE	18.79%
6	9	AT4G13940	ATSAHH1, EMB1395, EMBRYO DEFECTIVE 1395, HOG1, HOMOLOGY-DEPENDENT GENE SILENCING 1, MATERNAL EFFECT EMBRYO ARREST 58, MEE58, S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE 1, SAH1, SAHH1	MALLVEKTSS GREYKVKDMS QADFGRLELE LAEVEMPGLM ACRTEFGPSQ PFKGARITGS LMMTIQTAVL IETLTALGAE VRWCSCNIFS TQDHAAAAIA RDSAAVFAWK GETLQEYWWC TERALDWGPG GGPDLIVDDG GDATLLIHEG VKAEEIFEKT GQVPDPTSTD NPEFQIVLSI IKEGLQVDPK KYHKMKERLV GVSEETTIGV KRLYQMQQNG TLLFPAINVN DSVTKSKFDN LYGCRHSLPD GLMRATDVMI AGKVAVICGY GDVGKGCAAA MKTAGARVIV TEIDPICALQ ALMEGLQVLT LEDVVSEADI FVTITGNKDI IMVDHMRKMK NNAIVCNIGH FDNEIDMLGL ETYPGVKRIT IKPQTDRWVF PETKAGIIVL AEGRLMNLGC ATGHPSFVMS CSFTNQVIAQ LELWNEKASG KYEKKVYVLP KHLDEKVALL HLGKLGARLT KLSKDQSDYV SIPIEGPYKP PHYRY	9.90%
6	8	AT1G80410	EMB2753, EMBRYO DEFECTIVE 2753, NAA15, OMA, OMISHA	MGASLPPKEA NLFKLIVKSY ETKQYKKGLK AADAILKKFP DHGETLSMKG LTLNCMDRKT EAVELVRLGV KNDIKSHVCW HVLGLLYRSD REYREAIKCY RNALRIDPDN LEILRDLSLL QAQMRDLSGF VETRQQLLTL KPNHRMNWIG FAVSQHLNAN ASKAVEILEA YEGTLEDDYP PENELIEMTE MILYKVSLLE ESGSFDKALE ELMKKEPKIV DKLSYKEQEV SLLSKVGRLE EANKLYRVLL SMNPDNYRYM EGLQKCLGLY SESGQYSSDQ IEKLNALYQS LSEQYTRSSA VKRIPLDFLQ DENFKEAVAK YIKPLLTKGV PSLFSDLSSL YDHPRKPDIL EQLVVEMKHS IGTTGSFPGS DVKEPPSTLL WTLFFLAQHY DRRGQYDVAL CKIDEAIAHT PTVIDLYSVK SRIMKHAGDL TAAAALADEA RGMDLADRYI NSECVKRMLQ ADQVPLAEKT AVLFTKEGDQ LNNLHDMQCM WYDLASGDSY FRQGDLGRAL KKFLAVEKHY ADISEDQFDF HSYCLRKMTL RSYVDMLKFQ DRLHSFPYFH KAAIRAIRCY LKLHDSPKST AGEDEMSKLA PAQKKKIKKQ KKAEARAKKE AESKSEESTA SGASKSGKRN VKPVDPDPHG QKLIQVEEPM AEASKYLRLL QKHSPNSLET HLLSFEVHMR KQKFLLAFQA VKQLLKLGAE NPDSHRSLVK FFLMTESISA PTTEAEKLRW RVLEAERPSI SQLQNKSLME ANKEFLGRHE DSLVHRAAYA EMLYILDPSK KTEAIKIIED STNKVVQTNE ALGQAREWKL KDCIAVHTLL DTVLLDSQAA SRWKSRCAEY FPCSTHFEGK HCSLMPDSVV NSSRKSNENG DTPNMFMGQT ELSDGQLEAF KSLSVAT	6.91%
5	9	AT4G35090	CAT2, CATALASE 2	MDPVKYRPAS SYNSPFFTTN SGAPVWNNNS SMTVGPRGPI LLEDYHLVEK LNFDRERIP ERVVHARGAS AKGFFEVTKD ISNLTCADFL RAPGVQTPVI VREFTVIHER GSPETLRDPR GFAVKFYTRE GNFDLVGNNF PVFFIRDGMK FPDMVHALKP NPKSMIQENW RILDFFSMHP ESLHWFTTLF DDIGIPQDYR HMDGSGVNTY MLINKAGKAH VVKFHWKPTC GVKSLLEEDA IAVGGTNHSH ATQDLYDSIA AGMYPEWKLF IQIIDPADED KFDFDPLDVT KTWPEDILPL QPVGRMVLNK NIDNFFAENE QLAFCPAIIV PGIHYSDDKL LQTRVFSYAD TQRHRLGPNV LQLPVNAPKC AHHNNHNEGF MNFMHRDEEV NYFPSRYDQV RHAEKYPTPP AVCSGKRERC IIEKENNFKE PGERYRTFTP ERQERFIQRW IDALSDPRIT HEIRSIWISY WSQADKSLGQ KLASRLNVRP SI	15.04%
5	8	AT1G56070	LOS1, LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1	MVKFTADELR RIMDYKHNIR NMSVIAHVDH GKSTLTDSLV AAAGIIAQEV AGDVRMTDTR ADEAERGITI KSTGISLYYE MTDESLKSFT GARDGNEYLI NLIDSPGHVD FSSEVTAALR ITDGALVVVD CIEGVCVQTE TVLRQALGER IRPVLTVNKM DRCFLELQVD GEEAYQTFSR VIENANVIMA TYEDPLLGDV QVYPEKGTVA FSAGLHGWAF TLTNFAKMYA SKFGVVESKM MERLWGENFF DPATRKWSGK NTGSPTCKRG FVQFCYEPIK QIIATCHNDQ KDKLWPMLAK LGVSMKNDEK ELMGKPLMKR VMQTWLPAST ALLEMMIFHL PSPHTAQRYR VENLYEGPLD DQYANAIRNC DPNGPLMLYV SKMIPASDKG RFFAFGRVFA GKVSTGMKVR IMGPNYIPGE KKDLYTKSVQ RTVIWMGKRQ ETVEDVPCGN TVAMVGLDQF ITKNATLTNE KEVDAHPIRA MKFSVSPVVR VAVQCKVASD LPKLVEGLKR LAKSDPMVVC TMEESGEHIV AGAGELHLEI CLKDLQDDFM GGAEIIKSDP VVSFRETVCD RSTRTVMSKS PNKHNRLYME ARPMEEGLAE AIDDGRIGPR DDPKIRSKIL AEEFGWDKDL AKKIWAFGPE TTGPNMVVDM CKGVQYLNEI KDSVVAGFQW ASKEGPLAEE NMRGICFEVC DVVLHSDAIH RGGGQVIPTA RRVIYASQIT AKPRLLEPVY MVEIQAPEGA LGGIYSVLNQ KRGHVFEEMQ RPGTPLYNIK AYLPVVESFG FSSQLRAATS GQAFPQCVFD HWEMMSSDPL EPGTQASVLV ADIRKRKGLK EAMTPLSEFE DKL	6.29%
4	6	ATCG00120	ATP SYNTHASE SUBUNIT ALPHA, ATPA	MVTIRADEIS NIIRERIEQY NREVTIVNTG TVLQVGDGIA RIYGLDEVMA GELVEFEEGT IGIALNLESN NVGVVLMGDG LMIQEGSSVK ATGKIAQIPV SEAYLGRVIN ALANPIDGRG KISASESRLI ESPAPGIISR RSVYEPLQTG LIAIDSMIPI GRGQRELIIG DRQTGKTAVA TDTILNQQGQ NVICVIVAIG QRASSVAQVV TSLQERGAME YTIVVAETAD SPATLQYLAP YTGAALAEYF MYREQHTLII YDDLSKQAQA YRQMSLLLRR PPGREAYPGD VFYLHSRLLE RAAKLSSQLG EGSMTALPIV ETQSGDVSAY IPTNVISITD GQIFLSADLF NAGIRPAINV GISVSRVGSA AQIKAMKQVA GKLKLELAQF AELEAFSQFS SDLDKATQNQ LARGQRLREL LKQSQSAPLT VEEQIMTIYT GTNGYLDGLE IGQVRKFLVQ LRTYLKTNKP QFQEIIASTK TLTAEAESFL KEGIQEQLER FLLQEKV	8.09%

Table 3 Continued.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
3	6	AT5G62690	TUB2, TUBULIN BETA CHAIN 2	MREILMIQGG QCGNQIGAKF WEVVCAEHGI DPTGRYTGDS DLQLERINVY YNEASCGRFV PRAVLMDLEP GTMDSLRSGP YGQTFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDSVLDVV RKEAENCDCL QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMLTFSVF PSPKVSDTVV EPYNATLSVH QLVENADECM VLDNEALYDI CFRTLKLTTP SFGDLNHLIS ATMSGVTCCL RFPGQLNSDL RKLAVNLIPF PRLHFFMVGF APLTSRGSQQ YRSLTVPELT QQMWDSKNMM CAADPRHGRY LTASAMFRGK MSTKEVDEQM LNVQNKNSSY FVEWIPNNVK STVCDIPPTG LKMASTFIGN STSIQEMFRR VSEQFTAMFR RKAFLHWYTG EGMDEMEFTE AESNMNDLVS EYQQYQDATA DEEGDYEDEE EGEYQQEEEY	21.11%
3	3	AT5G44340	TUB4, TUBULIN BETA CHAIN 4	MREILMIQGG QCGNQIGAKF WEVICDEHGI DNTGQYVGDS PLQLERIDVY FNEASGGKYV PRAVLMDLEP GTMDSLRSGP FGQIFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDSVLDVV RKEAENSDCL QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMGTFSVF PSPKVSDTVV EPYNATLSVH QLVENADECM VLDNEALYDI CFRTLKLANP TFGDLNHLIS ATMSGVTCCL RFPGQLNSDL RKLAVNLIPF PRLHFFMVGF APLTSRGSQQ YRALSVPELT QQMWDAKNMM CAADPRHGRY LTASAVFRGK LSTKEVDEQM MNIQNKNSSY FVEWIPNNVK SSVCDIAPKG LKMASTFIGN STSIQEMFRR VSEQFTAMFR RKAFLHWYTG EGMDEMEFTE AESNMNDLVA EYQQYQDATA GEEEYEEEEE EYET	19.59%
3	3	AT3G05530	ATS6A.2, REGULATORY PARTICLE TRIPLE-A ATPASE 5A, RPT5A	MATPMVEDTS SFEEDQLASM STEDITRATR LLDNEIRILK EDAQRTNLEC DSYKEKIKEN QEKIKLNKQL PYLVGNIVEI LEMNPEDDAE EDGANIDLDS QRKGKCVVLK TSTRQTIFLP VVGLVDPDSL KPGDLVGVNK DSYLILDTLP SEYDSRVKAM EVDEKPTEDY NDIGGLEKQI QELVEAIVLP MTHKERFEKL GVRPPKGVLL YGPPGTGKTL MARACAAQTN ATFLKLAGPQ LVQMFIGDGA KLVRDAFQLA KEKAPCIIFI DEIDAIGTKR FDSEVSGDRE VQRTMLELLN QLDGFSSDER IKVIAATNRA DILDPALMRS GRLDRKIEFP MPTEEARARI LQIHSRKMNV MPDVNFEELA RSTDDFNGAQ LKAVCVEAGM LALRRDATEV NHEDFNEGII QVQAKKKASL NYYA	7.08%
3	4	AT1G20630	CAT1, CATALASE 1	MDPYRVRPSS AHDSPFFTTN SGAPVWNNNS SLTVGTRGPI LLEDYHLLEK LANFDRERIP ERVVHARGAS AKGFFEVTHD ITQLTSADFL RGPGVQTPVI VRFSTVIHER GSPETLRDPR GFAVKFYTRE GNFDLVGNNF PVFFVRDGMK FPDMVHALKP NPKSHIQENW RILDFFSHHP ESLHMFSFLF DDLGIPQDYR HMEGAGVNTY MLINKAGKAH YVKFHWKPTC GIKCLSDEEA IRVGGAMHSH ATKDLYDSIA AGNYPQWNLF VQVMDPAHED KFDFDPLDVT KIWPEDILPL QFVGRLVLNK NIDNFFNENE QIAFCPALVV PGIHYSDDKL LQTRIFSYAD SQRHRLGPNY LQLPVNAPKC AHHNNHHDGF MNFMHRDEEV NYFPSRLDPV RHAEKYPTTP IVCSGNREKC FIGKENNFKQ PGERYRSWDS DRQERFVKRF VEALSEPRVT HEIRSIWISY WSQADKSLGQ KLATRLNVRP NF	5.69%
3	4	AT1G06950	ARABIDOPSIS THALIANA TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS 110, ATTIC110, TIC110, TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS 110	MNPSLVTAIN APISPSPRSP LLSMFLPTLP MRFSKSECLS RRRYRVSFPR SSAASSDQLS VSTQAKNPGI HGNKKELTGL QPIVEKMTPP VRLATSAVVL AASLATGYGL GLRLAGSRNI AFGGAAVAGA AGGAVVYALN SAVPEVAAIS LHNYVAEFED PASVTKDDVE KIADRYGVNK GDEAFQAEIC DIYCRYVTSV LPTEGQSLKG DEVAKIVKFK NALGIDEPDA AAMHMEIGRR IFRQRLETGE REGDAEQRRA FMRLVYVSAL VFGDASSFLL PWKRVLKVTD AQVEIAIREN AKQLYAERLK LVGRDINVEN LVDLRKSQLS FKLSDELAED LFREHTRKVV VENISSALSI LKSRTRAAKS LASVVEELEK VLEFNNLLVS LKSHSEADQF ARGVGPISLI GDESDFERRM DDLKLLYRAY VTDALSGGRL EENKLVAMSQ LRNILGLGKR EAEAISVDVT SKSYRKRLAN AVSSGDLEAQ DSKAKYLQKL CEELHFDAQK AGAIHEEIYR QKLQQCVTDG ELSDDNVAAL LRLRVMLCIP QQTVDTAHAE ICGTIFEKVV RDAISSGVDG YDAETRKSVR KAAMGLRLSR ETAMSIASKA VRRVFTNYIR RARAAENRTD SAKELKKMIA FNTLVVTEMV ADIKGESSDK APEEDPVQEK EEDDEDEEWG SLESLRKTRP DKELAEKMGK PGQTEITLKD DLPDRDRIDL YKTYLLYCVT GEVTRIPFGA QITTKRDDSE YLLLNQLGGI LGLSSKEIVN IHVGLAEQAF RQQAEVILAD GQLTKARVEQ LDELQKQVGL PQPQAEKVIK NITTTKMANA IETAVNQGRL NIKQIRELKE ANVSLDSMIA VSLREKLFKK TVSDIFSSGT GEFDETEVYQ TIPSDLSIDV EKAKRVVHDL AQSRLSNSLV QAVALLRQRN SKGVVLSLND LLACDKAVPA EPMSWEVSEE LSDLYAIYSK SDPKPAPEKV LRLQYLLGID DSTATALREM EDGALSSAAE EGNFVF	3.94%
2	3	AT5G14740	BETA CA2, BETA CARBONIC ANHYDRASE 2, CA18, CA2, CARBONIC ANHYDRASE 18, CARBONIC ANHYDRASE 2	MGNESYEDAI EALKKLLIEK DDLKDVAAAK VKKITAELQA ASSSDSKSFD PVERIKEGFV TFKKEKYETN PALYGELARG QSPKYMVFAC SDSRVCPSHV LDFHPGDAFV VRNIANMVPP FDKVKYAGVG AAIEYAVLHL KVENIVVIGH SACGGIKGLM SFPLDGNNST DFIEDWVKIC LPAKSKVLAE SESSAFEDQC GRCEREAVNV SLANLLTYPF VREGVVKGTL ALKGGYYDFV NGSFELWELQ FGISPVHSI	11.58%
2	2	AT1G20620	ATCAT3, CAT3, CATALASE 3, SEN2, SENESCENCE 2	MDPYKYRPSS AYNAPFYTTN GGAPVSNNIS SLTIGERGPV LLEDYHLIEK VANFTRERIP ERVVHARGIS AKGFFEVTHD ISNLTCADFL RAPGVQTPVI VRFSTVVHER ASPETMRDIR GFAVKFYTRE GNFDLVGNNT PVFFIRDGIQ FPDVVHALKP NPKTNIQEYW RILDYMSHLP ESLLTWCWMF DDVGIPQDYR HMEGFGVHTY TLIAKSGKVL FVKFHWKPTC GIKNLTDEEA KVVGGAMHSH ATKDLHDAIA SGNYPEWKLF IQTMDPADED KFDFDPLDVT KIWPEDILPL QFVGRLVLNR TIDNFFNETE QLAFNPGLVV PGIYYSDDKL LQCRIFAYGD TQRHRLGPNY LQLPVNAPKC AHHNNHHEGF MNFMHRDEEI NYYPSKFDPV RCAEKVPTPT NSYTGIRTKC VIKKENNFKQ AGDRYRSWAP DRQADRFVKRW VEILSEPRLT HEIRGIWISY WSQADRSLGQ KLASRLNVRP SI	10.98%
2	2	AT5G28840	"GDP-D-MANNOSE 3',5'-EPIMERASE"	MGTTNGTDYG AYTYKELERE QYWPSENLKI SITGAGGFIA SHIARRLKHE GHYVIASDWK KNEHMTEDMF CDEFHLVDLR VMENCLKVTE GVDHVFNLAA DMGGMGFIQS NHSVIMYNNT MISFNMIEAA RINGIKRFFY ASSACIYPEF KQLETTNVSL KESDAWPAEP QDAYGLEKLA TEELCKHYNK DFGIECRIGR FHNIYGPFGT WKGGREKAPA AFCRKAQTST DRFEMWGDGL QTRSFTFIDE CVEGVLRLTK SDFREPVNIG SDEMVSMNEM AEMVLSFEEK KLPIHHIPGP EGVRGRNSDN NLIKEKLGWA PNMRLKEGLR ITYFWIKEQI EKEKAKGSDV SLYGSSKVVG TQAPVQLGSL RAADGKE	7.16%
2	2	AT2G28000	CH-CPN60A, CHAPERONIN-60ALPHA, CHAPERONIN-60ALPHA1, CHLOROPLAST CHAPERONIN 60ALPHA, CPN60A, CPN60ALPHA1, SCHLEPPERLESS, SLP	MASANALSSA SVLCSSRQSK LGGGNQQQQQ RVSYNKRTIR RFSVRANVKE IAFDQHSRAA LQAGIDKLAD CVGLTLGPRG RNVVLDEFGS PKVVNDGVTI ARAIELPNAM ENAGAALIRE VASKTNDSAG DGTTTASILA REIIKHGLLS VTSGANPVSL KRGIDKTVQG LIEELQKKAR PVKGRDDIRA VASISAGNDD LIGSMIADAI DKVGPDGVLS IESSSSFETT VEVEEGMEID RGYISPQFVT NPEKLLAEFE NARVLITDQK ITAIKDIIPI LEKTTQLRAP LLIIAEDVTG EALATLVVNK LRGVLNVVAV KAPGFGERRK AMLQDIAILT GAEYLAMDMS LLVENATIDQ LGIARKVTIS KDSTTLIADA ASKDELQARI AQLKKELFET DSVYDSEKLA ERIAKLSGGV AVIKVGAATE TELEDRKLRI EDAKNATFAA IEEGIVPGGG AALVHLSTVI PAIKETFEDA DERLGADIVQ KALLSPAALI AQNAGVEGEV VVEKIMFSDW ENGYNAMTDT YENLFEAGVI DPAKVTRCAL QNAASVAGMV LTTQAIVVDK PKPKAPAAAA PEGLMV	6.66%
2	3	AT4G16143	IMPA-2	MSLRPNAKTE VRRNRYKVAV DAEEGRRRRE DNMVEIRKSK REESLQKKRR EGLQANQLPQ FAPSPVPASS TVEKKLESLP AMVGGVWSDD RSLQLEATTQ FRKLLSIERS PPIEEVIDAG VVPRFVEFLT REDYPQLQFE AAWALTNIAS GTSENTKVVI EHGAVPIFVQ LLASQSDDVR EQAVWALGNV AGDSPRCRDL VLGQGALIPL LSQLNEHAKL SMLRNATWTL SNFCRGKPQP PFDQVRPALP ALERLIHSTD EEVLTDACWA LSYLSDGTND KIQSVIEAGV VPRLVELLQH QSPSVLIPAL RSIGNIVTGD DLQTQCVISH GALLSLLSLL THNHKKSIKK EACWTISNIT AGNRDQIQAV CEAGLICPLV NLLQNAEFDI KKEAAWAISN ATSGGSPDQI KYMVEQGVVK PLCDLLVCPD PRIITVCLEG LENILKVGEA EKVTGNTGDV NFYAQLIDDA EGLEKIENLQ SHDNSEIYEK AVKILETYWL EEEDETLPPG DPSAQGFQFG GGNDAAVPPG GFNFQ	6.36%
2	2	AT5G43960	Nuclear transport factor 2 (NTF2)	MATPVPGATQ VGSVFVGQVY QVLQQQPDLI HQFYSEPSRA IRIDGDSTET ANSLLHIHNM VMSLNFTAIE VKTINSVESW EGGVLVVVSG SVKTKEFSNR RSFVQTFFLA PQEKGYFVLS DVFLFVDEGT VYYHQPSYLS EIKHEAQLNP PTRHPDPQVS DYVLEEEASD YVNAVQIKDD LVDKYSLQED QHQPQHEDYE DEVAIEETPR EEVAVDVVHE HRAAPVEEPV GEKSKMSYAS ILKVAKEAAT VPVAATQPSY NKSSQDINEW DQPMRTPSPQ LAAPLAPIQQ SNSSTYVSDY GAEAEDGSGF EDFEFKSVYV RNLPSDISAS EIEEEFKNFG TIKPDGVFLR TRKDVMGVCY AFVEFEDMTS VENAIKASPI YLGGRQVYIE ERRPNPAGVR GARRGGGRGR GGYPTEAPRG RFGGRGSGRG NQDGGDYRPR GNGYYRGGGR	5.33%
2	3	AT5G22060	ARABIDOPSIS THALIANA DNAJ HOMOLOGUE 2, ATJ2	MFGRGPSRKS DNTKFYEILG VPKTAAPEDL KKAYKKAAIK NHPDKGGDPE KFKELAQAYE VLSDPEKREI YDQYGEDALK EGMGGGGGGH DPFDIFSSFF GSGGHPFGSH SRGRRQRRGE DVVHPLKVSL EDVYLGTTKK LSLSRKALCS KCNGKGSKSG SGMKISIRQF GPGMMQQVQH ACNDCKGTGE TINDRDRCPQ CKGEKVVSEK KVLEVNVEKG MQHNQKITFS GQADEAPDTV TGDIVFVIQQ KEHPKFKRKG EDLFVEHTIS LTEALCGFQF VLTHLDKRQL LIKSKPGEVV KPDSYKAISD EGMPIYQRPF MKGKLYIHFT VEFPESLSPD QTKAIEAVLP KPTKAAISDM EIDDCEETTL HDVNIEDEMK RKAQAQREAY DDDEEDHPGG AQRVQCAQQ	5.01%

Table 3 Continued.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
2	2	AT2G38040	ACETYL CO-ENZYME A CARBOXYLASE CARBOXYLTRANSFERASE ALPHA SUBUNIT, CAC3	MASISMSSLA LGGASSASAS DYLRSSSNGV NGVPLKTLGR AVFTTIRRKD LAVTSRLKKG KKFEHPWPAN PDPNVKGGVL SYLAEFKPLG DTQKPVTLDF EKPLVELEKK IVDVRKQANE TGLDFTEQII TLENKYRQAL KDLYTMLTPI QRVNIARMPN RPTFLDMIHN ITDKFMELHG DRAGYDDPAI VTGIGTIDGK RYMFIGHQKG RNTKENIMRN FGMPTPHGYR KALRMMYYAD HNGFPIVTFI DTPGAYADLK SEELGQGEAI ANNLRTMFGL KVPILSIVIG EGGSGGALAI GCANKMLMLE NAVFYVASPE ACAAILWKTS KAAPEAAEKL RITSKELVKL NVADGIIPEP LGGAHADPSW TSQQIKIAIN ENMNEFGKMS GEELLKHRMA KYRKIGVFIE GEPIEPSRKI NMKKREAVFS DSRKLQGEVD KLKEQILKAK ETSTEAEPSS EVLNEMIEKL KSEIDDEYTE AAIAVGLEER LTAMREEFSK ASSEEHLMNP VLIEKIEKLK EEFNTRLTDA PNYESLKSKL NMLRDFSRAK AASEATSLKK EINKRFQEAV DRPEIREKVE AIKAEVASSG ASSFDELPDA LKEKVLKTKG EVEAEMAGVL KSMGLELDAV KQNQKDTAEQ IYAANENLQE KLEKLNQEIT SKIEEVVRTP EIKSMVELLK VETAKASKTP GVTEAYQKIE ALEQQIKQKI AEALNTSGLQ EKQDELEKEL AAARELAAEE SDGSVKEDDD DDEDSSESGK SEMVNPSFA	4.55%
2	2	AT3G23990	CPN60 HEAT SHOCK PROTEIN 60, HEAT SHOCK PROTEIN 60-3B, HSP60	MYRFASNLAS KARIAQNARQ VSSRMSWSRN YAAKEIKFGV EARALMLKGV EDLADAVKVT MGPKGRNVVI EQSWGAPKVT KDGVTVAKSI EFKDKIKNVG ASLVKQVANA TNDVAGDGTT CATVLTRAIF AEGCKSVAAG MNAMDLRRGI SMAVDAVVTN LKSKARMIST SEEIAQVGTI SANGEREIGE LIAKAMEKVG KEGVITIQDG KTLFNELEVV EGMKLDRGYT SPYFITNQKT QKCELDDPLI LIHEKKISSI NSIVKVLELA LKRQRPLLIV SEDVESDALA TLILNKLRAG IKVCAIKAPG FGENRKANLQ DLAALTGGEV ITDELGMNLE KVDLSMLGTC KKVTVSKDDT VILDGAGDKK GIEERCEQIR SAIELSTSDY DKEKLQERLA KLSGGVAVLK IGGASEAEVG EKKDRVTDAL NATKAAVEEG ILPGGGVALL YAARELEKLP TANFDQKIGV QIIQNALKTP VYTIASNAGV EGAVIVGKLL EQDNPDLGYD AAKGEYVDMV KAGIIDPLKV IRTALVDAAS VSSLLTTTEA VVVDLPKDES ESGAAGAGMG GMGGMDY	4.16%
2	3	AT2G05920	SBT1.8, SUBTILASE 1.8	MASSSSSSSS ITIITTFLFL LLHTTARKTY IIRVNHSDKP ESFLTHHDWY TSQLNSESSL LYTYTTSFHG FSAYLDSTEA DSLLSSSNSI LDIFEDPLYT LHTTRTPEFL GLNSEFGVKD LGSSSNGVII GVLDTGVWPE SRSFDDTDMP EIPSKWKGEC ESGSDFDSKL CNKKLIGARS FSKGFQMASG GGFSSKRESV SPRDVDGHGT HTSTTAAGSA VRNASFLGYA AGTARGMATR ARVATYKVCW STGCFGSDIL AAMDRAILDG VDVLSLSLGG GSAPYYRDTI AIGAFSAMER GVFVSCSAGN SGPTRASVAN VAPWVMTVGA GTLDRDFPAF ANLGNGKRLT GVSLYSGVGM GTKPLELVYN KGNSSSSNLC LPGSLDSSIV RGKIVVCDRG VNARVEKGAV VRDAGGLGMI MANTAASGEE LVADSHLLPA IAVGKKTGDL LREYVKSDSK PTALLVFKGT VLDVKPSPVV AAFSSRGPNT VTPEILKPDV IGPGVNILAG WSDAIGPTGL DKDSRRTQFN IMSGTSMSCP HISGLAGLLK AAHPEWSPSA IKSALMTTAY VLDNTNAPLM DAADNSLSNP YAHGSGHVDP QKALSPGLVY DISTEEYIRF LCSLDYTVDH IVAIVKRPSV NCSKKFSDPG QLNYPSFSVL FGGKRVVRYT REVTNVGAAS SVYKVTVNGA PSVGISVKPS KLSFKSVGEK KRYTVTFVSK KGVSMTNKAE FGSITWSNPQ MEVRSPVAFS WNRF	3.45%
2	2	AT5G26742	ATRH3, EMB1138, EMBRYO DEFECTIVE 1138, RH3	MASTVGVPSL YQVPHLEISK PNSKKRSNCL SLSLDKPFFT PLSLVRRTRR IHSSSLLVFS AVATPNSVLS EEAFKSLGLS DHDEYDLDGD NNNVEADDGE ELAISKLSLP QRLEESLEKR GITHLFPIQR AVLVPALQGR DIIARAKTGT GKTLAFGIPI IKRLTEEAGD YTAFRRSGRL PKFLVLAPTR ELAKQVEKEI KESAPYLSTV CVYGGVSYTI QQSALTRGVD VVVGTPGRII DLIEGRSLKL GEVEYLVLDE ADQMLAVGFE EAVESILENL PTKRQSMLFS ATMPTWVKKL ARKYLDNPLN IDLVGDQDEK LAEGIKLYAI ATTSTSKRTI LSDLITVYAK GGKTIVFTQT KRDADEVSLA LSNSIATEAL HGDISQHQRE RTLNAFRQGK FTVLVATDVA SRGLDIPNVD LVIHYELPND PETFVHRSGR TGRAGKEGSA ILMHTSSQRR TVRSLERDVG CHFEFISPPT VGDLLESSAD QVVATLNGVH PDSIKFFSAT AQKLVEEKGT DALAAALAML SGFSQPPSSR SLLSHEKGWV TLQLIRDPTN ARGFLSARSV TGFLSDLVRT AADEVGKIFL IADDRIQGAV FDLPEEIAKE LLEKDVPEGN SLSMITKLPP LQDDGPSSDN YGRFSSRDRM PRGGGGSRGS RGGRGGSSRG RDSWGGDDDR GSRRSSGGGS SWSRGGSSSR GSSDDWLIGG RSSSSSRAPS RERSFGGSCF ICGKSGHRAT DCPDKRGF	2.54%
2	2	AT1G69830	ALPHA-AMYLASE-LIKE 3, AMY3, ATAMY3	MSTVPIESLL HHSYLRHNSK VNRGNRSFIP ISLNLRSHFT SNKLLMSIGK SVGVSSMNKS PVAIRATSSD TAVVETAQSD DVIFKEIFPV QRIEKAEGKI YVRLKEVKEK NWELSVGCSI PGKWILHWGV SYVGDTGSEW DQPPEDMRPP GSIAIKDYAI ETPLKKLSEG DSFFEVAINL NLESSVAALN FVLKDEETGA WYQHKGRDFK VPLVDDVPDN GNLIGAKKGF GALGQLSNIP LKQDKSSAET DSIEERKGLQ EFYEEMPISK RVADDNSVSV TARKCPETSK NIVSIETDLP GDVIVHHGVC KNGTKKWEIP SEPYPEETSL FKNKALRTRL QRKDDGNGSF GLFSLDGKLE GLCFVLKLNE NTWLNYRGED FYVPFLTSSS SPVETEAAQV SKPKRKTDKE VSASGFTKEI ITEIRNLAID ISSHKNQKTN VKEVQENILQ EIEKLAAEAY SIFRSTTPAF SEEGVLEAEA DKPDIKISSG TGSGFEILCQ GFNWESNKSG RWYLELQEKA DELASLGFTV LWLPPPTESV SPEGYMPKDL YNLNSRYGTI DELKDTVKKF HKVGIKVLGD AVLNHRCAHF KNQNGVWNLF GGRLNWDDRA VVADDPHFQG RGNKSSGDNF KAAPNIDHSQ DFVRKDIKEW LCWMMEEVGY DGWRLDFVRG FWGGYVKDYM DASKPYFAVG EYWDSLSYTY GEMDVNQDAH RQRIVDWINA TSGAAGAFDV TTKGILHTAL QKCEYWRLSD PKGKPPGVVG WWPSRAVTFI ENHDTGSTQG HWRFPEGKEM QGYAYILTHP GTPAVFFDHI FSDYHSEIAA LLSLRNRQKL HCRSEVNIDK SERDVYAAII DEKVAMKIGP GHYEPPNGSQ NWSVAVEGRD YKVWETS	2.25%
2	2	AT5G63840	PRIORITY IN SWEET LIFE 5, PSL5, RADIAL SWELLING 3, RSW3	MRSLLFVLSL ICFCSQTALS WKKEEFRSCD QTPFCKRARS RTPGACSLIV GDVSITDGDL VAKLLPKAPN QGDGDQIKPL ILSLSVYKDG IVRLKIDEDH SLNPPKKRFQ VPDVVVSEFE EKKIWLQKVA TETISGDTSP SSVVYVSDGY EAVVRHDPFE VYVREKSGDR RRVVSLNSHG LFDFEQLGRK TEGDNWEEKF RTHTDSRPSG PQSISFDVSF YDSSFVYGIP EHATSFALKP TKGPGVEESE PYRLFNLDVF EYDHESPFGL YGSIPFMVSH GKSGKTSGFF WLNAAEMQID VLAHGWDAES GISLPSSMSR IDTFWMSEAG IVDTFTFVGP EPKDVVKQYA SVTGTSAMPQ LFATGYHQCR WNYKDEEDVA QVDSKFDEMD IPYDVLWLDI EMTDGKRYFT WDSVLFPHPE EMQKKLAAKG RKMVTIVDPH IKRDDSYFLH KEATQMGYYV KDSSGKDFDG WCWPGSSSYI DMLSPEIRKW WGGRFSYNYV VGSTPSLYTW NDMNEPSVFN GPEVTMPRDA LNVGGVEHRE VHNAYGYYFH MATSDGLVMR EEGKDRPFVL SRAIFPGTQR YGAIWTGDNT AEWEHLRVSI PMILTLGLTG ITFSGADIGG FFGNPEPELL VRWYQVGAYY PFFRGHAHHD TKRREPWLFG ERNTELMRDA IHTRYTLLPY FYTLFREANV TGVPVVRPLW MEFPQDEATF SNDEAFMVGS GLLVQGVYTK GTTQASVYLP GKESWYDLRN GKTYVGGKTH KMDAPEESIP AFQKAGTIIP RKDRFRRSSS QMDNDPYTLV VALNSSQEAE GELYIDDGKS FEFRRGSYIH RRFVFSKGVL TSTNLAPPEA RLSSQCLIDR IILLGHSSGP KSALVEPLNQ KAEIEMGPLR MGGLVASSGT KVLTIRKPGV RVDQDWTVKI L	2.17%
2	2	AT2G21390	Coatomer, alpha subunit	MLTKFETKSN RVKGLSFHPK RPWILASLHS GVIQLWDYRM GTLIDRFDEH EGPVRGVMFH NSQPLFVSGG DDYKIKVWNY KTHRCLFTLL GHLDYIRTVQ FHHENPWIVS ASDDQTIRIW NWQSRTCISV LTGHNHYVMC ASFHPKEDLV VSASLDQTVR VWDIGALKKK SASPADDLMR FSQMNSDLFG GVDAIVKYVL EGHDRGVNWA SFHPTLPLIV SGADDRQVKL WRMNETKAWE VDTLRGHMNN VSSVMFHAKQ DIIVSNSEDK SIRVWDATKR TGIQTFRREH DRFWILAVHP EINLLAAGHD NGMIVFKLER ERPAFALSGD SLFYAKDRFL RYYEYSTQKD SQVIPIRRPG TPSLNQSPRT LSYSPTENAV LICSDLDGGS YELYIIPKDS VGRSDVVQDA KRGTGGSAVF IARNRFAVLE KSTSQVLVKN LKNEVVKKSS LPIPTDAIFY AGTGNLLCRS EDKVVIFDLQ QRLVLGELQT PFVRVVVWSN DMESVALLSK HTIIIASKKL VLQCTLHETI RVKSGAMDDN GVFIYTTLNH IKYCLPNGDS GIIRTLDVPI YITKVSGNTI FCLDRDGKNR AITINATEYI FKLALLRKKY DHVMSMIKNS QLCGQAMIAY LQQKGFPEVA LHFVEDERIR FNLALESGNI SVAVASATEI NEKDHWYRLG VEALRQGNSR IVEFAYQQTK NFERLSFLYL ITGNLDKLSK LMKIAEVKNN VMGQFHNALY LGDVKERVKI LENAGHLPLA YITASVKGLT DIAERLAIEL GDNVPSLPEG KTPSLLMPPS PIMCGGDWPL LRVMKGIFEG GLESANRGAV DEEEEDVEGD WGEGLDKFDV DGMENTDIEA ILDGAEAGEE EDDEEGGWGL DLDLPFELDT PKASANARSS TFVTPPQGMP VSQIWSQKSS LAAEQAAAGS FDTAMRLLHR QLGIKNFAPL KSMFLDLFSG SHSYLRAFSS SPVVPLAIER GWSESSSPNV RGPPALVFDF SQLEAKLKSG YKATTAGKLS EALRVFLSIL QTIPLVVVES RREVDEVKEL VIIVKEYVLG LQLELKRREM KDDPVRQQEL AAYFTHCKLQ TPHLRLAYFS AMTVCYKSKN MATAAHFARS LLDTNPTIES QARTARQVMQ AAERNMTDAT TLNVFDFNPF VICGSTVVPI YKGQKDVACP YCTARFVPSQ EGNICSVCDL AVIGADASGL LCSASQVR	1.97%
1	1	AT5G08315	Encodes a defensin-like (DEFL) family protein	MAGLKVFSFA LLLILTFSLI DVEGYNVENG GSLCCNNHPK FGKCNTKNDD QRCNSWCLNG CDNGKGGYCK SMSHGGQCHC YC	26.83%
1	3	AT2G29550	TUB7, TUBULIN BETA-7 CHAIN	MREILHIQGG QCGNQIGSKF WEVVNLEHGI DQTGRYVGDS ELQLERVNVY YNEASCGRYV PRAVLMDLEP GTMDSVRSGP YGQIFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDSVLDVV RKEAENCDCL QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMNTFSVF PSPKVSDTVV EPYNATLSVH QLVENADECM VLDNEALYDI CFRTLKLSTP SFGDLNHLIS ATMSGVTCCL RFPGQLNSDL RKLAVNLIPF PRLHFFMVGF APLTSRGSQQ YRNLTVPELT QQMWDAKNMM CAADPRHGRY LTASAMFRGK MSTKEVDEQM LNVQNKNSSY FVEWIPNNVK STVCDIPPTG LKMASTFIGN STSIQEMFRR VSEQFTAMFR RKAFLHWYTG EGMDEMEFTE AESNMNDLVS EYQQYQDATA DEEGEYEEEE AEYEQEETY	22.05%

Table 3 Continued.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
1	1	AT5G23860	TUB8, TUBULIN BETA 8	MREILHIQGG QCGNQIGAKF WEVVCAEHGI DBTGRYQGEN DLQLERVNVY YNEASCGRFV PRAVLMDLEP GTMDSVRSGP YGQIFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDSVLDVV RKEAENCDCL QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMLTFSVF PSPKVSDTVV EPYNATLSVH QLVENADECM VLDNEALYDI CFRTLKLTTP SFGDLNHLIS ATMSGVTCCL RFPGQLNSDL RKLAVNLIPF PRLHFFMVGF APLTSRGSQQ YRALTVPELT QQMWDAKNMM CAADPRHGRY LTASAMTRGK MSTKEVDEQM INVQNKNSSY FVEWIPNNVK STVCDIPPTG LKMAATFVG NSTSIQEMFR RVSEQFTAMF RKAFLHWYTG EGMDEMEFTE AESNMNDLVS EYQQYQDATA DEEEGEYEEE DEVEVQEEQ	21.83%
1	1	AT1G20010	TUB5, TUBULIN BETA-5 CHAIN	MREILHIQGG QCGNQIGSKF WEVICDEHGI DBTGRYSGDT ADLQLERINV YYNEASGGRY VPRAVLMDLE PGTMDSIRSG PFGQIFRPDN FVFGQSGAGN NWAKGHYTEG AELIDAVLDV VRKEAENCDC LQGFQVCHSL GGGTGSGMGT LLISKIREEY PDRMMLTFSV FPSPKVSDTV VEPYNATLSV HQLVENADEC MVLDNEALYD ICFRTLKLST PSFGDLNHLI SATMSGVTCS LRFPGQLNSD LRKLAVNLIP FPRLHFFMVG FAPLTSRGSQ QYISLTVPEL TQQMWDSKNM MCAADPRHGR YLTASAIFRG QMSTKEVDEQ ILNIQNKNSS YFVEWIPNNV KSSVCDIPPK GLKMAATFVG NSTSIQEMFR RVSEQFTAMF RKAFLHWYT GEGMDEMEFT EAESNMNDLV AEYQQYQDAT ADEEGEYDVE EEEEGDYET	18.93%
1	1	AT5G12250	BETA-6 TUBULIN, TUB6	MREILHIQGG QCGNQIGSKF WEVVCDEHGI DPTGRYVGNS DLQLERVNVY YNEASCGRYV PRAILMDLEP GTMDSVRTGP YGQIFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDAVLDVV RKEAENCDCL QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMLTFSVF PSPKVSDTVV EPYNATLSVH QLVENADECM VLDNEALYDI CFRTLKLTTP SFGDLNHLIS ATMSGVTCCL RFPGQLNSDL RKLAVNLIPF PRLHFFMVGF APLTSRGSQQ YRALTVPELT QQMWDSKNMM CAADPRHGRY LTASAMTRGK MSTKEVDEQM INVQNKNSSY FVEWIPNNVK SSVCDIAPRG LSMASTFIGN STSIQEMFRR VSEQFTAMFR RKAFLHWYTG EGMDEMEFTE AESNMNDLVS EYQQYQDATA DDEGEYEEDE DEEEILDHE	18.49%
1	1	AT1G12900	GAPA-2, GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE A SUBUNIT 2	MASATFSVAK PSLQGFSEFS GLRNSSALPF AKRSSSDEFV SFVSFQTSAM RSNGGYRKGV TEAKIKVAIN GFGRIGRNFL RCWHGRKDSP LDVVVINDTG GVKQASHLLK YDSTLGIFDA DVKPSGDSAL SVDGKIIKIV SDRNPSNLPW GELGIDLVIE GTGVFVDRDG AGKHLQAGAK KVLITAPGKG DIPTYVVGVN AELYSHEDTI ISNASCTTNC LAPFVKVLDQ KFGIIKGTMT TTHSYTGDQR LLDASHRDLR RARAAALNIV PTSTGAAKAV ALVLPNLKGK LNGIALRVPT PNVSVVDLVV QVSKKTFAEE VNAAFRDAAE KELKGILDVC DEPLVSVDFR CSDVSSTIDS SLTMVMGDDM VKVIAWYDNE WGYSQRVVDL ADIVANNWK	18.05%
1	1	AT2G27730	copper ion binding	MATRNALRIV SRRFSSGKVL SEEERAAENV FIKHMEQEKL QKLARQGPGE QAAGSASEAK VAGATASASA ESGPKVSEDK NRNYAVVAGV VAIVGSIGWY LKAGGKKQPE VQE	13.27%
1	2	ATMG00280	ORF110A. Ribulose bisphosphate carboxylase large chain, catalytic domain	MNNAAKRADC WFGAKNYGRA VYECLRGGLY FTKDDENVNS QPFMRWRDRF LFCAEAVYKA QAETGGIKGH YLNATAGTCE EMIKRAVFAR ELGVPIVMHD YLNRGIHRKY	10.91%
1	2	AT1G56410	EARLY-RESPONSIVE TO DEHYDRATION 2, ERD2, HEAT SHOCK PROTEIN 70T-1, HSP70T-1	MAGKGEGPAI GIDLGTTYSC VGVWQHDRVE IIANDQGNRT TPSYVAFTDS ERLIGDAAKN QVAMNPVNTV FDAKRLIGRR FSDASVQSDM KFWPFKVTPG QADKPMIFVN YKGEEKQFAA EEISSMVLIK MREIAEAYLG SSIKNAVVTV PAYFNDSQRQ ATKDAGVIAG LNVLRIINEP TAAAIAYGLD KKATSVGIKN VLIFDLGGGT FDVSLLTIEE GIFEVKATAG DTHLGGEDFD NRMVNHFVQE FKRKNKKDIS GDARALRRLR TACERAKRTL SSTAQTTVEV DSLFEGIDFY SPITRAKFEE MNMDLFRKCM EPVMKCLRDS KMDKSMVHDV VLVGGSTRIP KVQQLLQDFF NGKELCKSIN PDEAVAYGAA VQAAILSGEG NEKVQDLLLL DVTPLSLGIE TIGGVMTTLI QRNTTIPAKK EQEFTTTVDN QPDVLIQVYE GERARTIDNN ILGQFVLSGI PPAPRGIPQF TVCFDIDSNG ILNVSAEDKA TGKKNKITIT NDKGRLSKDD IEKMVQEAEK YKSEDEEHKK KVEAKNGLEN YAYNVGNTLR DMGEKLPAAD KKKFEDSIEE VIQWLDDNQL AEADEFEHKM KELESVWSTI ITKMYQG	8.75%
1	1	AT1G26880	RPL34A Ribosomal protein L34e superfamily protein	MVQRLVYRSR HSYATKSNQH RIVKTPGGKL VYQTTKKRAS GPKCPVTGKR IQGIPHLRPS EYKRSRLSRN RRTVNRAYGG VLSGSAVRER IIRAFLVEEQ KIVKKVLKLQ KAKEKVAPKA	6.67%
1	1	AT5G13780	MXE10.5 Encodes the catalytic subunit of a N-terminal acetyltransferase.	MVCIRRATVD DLLAMQACNL MCLPENYQMK YYLYHILSWP QLLYVAEDYN GRIVGYVLAK MEEESNECHG MITSLAVLRT MRKLGLATKL MTAAQAAMEQ VYEAEYVSLH VRRSNRAAFN LYTETLGYKI NDVEAKYYAD GEDAYDMRKN LKGKQNHMKA HGHHHHHGGG CCSGDAKVVE TAQAVDGKAV SK	5.73%
1	1	AT1G58380	XW6	MAERGGEGGA ERGGDRGDFG RGFGGGRGGG RGRDRGPRGR GRRGGRASEE TKWVPVTKLG RLVADNKITK LEQIYLHSLP VKEYQIIDHL VGPTLKDEVM KIMPVQKQTR AGQRTRFKAF VVVGDGNGHV GLGVKCSKEV ATAIRGAIIL AKLSVVPVRR GYWGNKIGKP HTVPCKVTGK CGSVTVRMVP APRGSGIVAA RVPKKVLQFA GIDDVFTSSR GSTKTLGNFV KATFDCLQKT YGFLTPEFWK ETRFSRSPYQ EHTDFLSTKA VSATKVITEG EDQA	5.28%
1	1	AT4G02520	ATGSTF2, ATPM24, ATPM24.1, GLUTATHIONE S-TRANSFERASE PHI 2, GST2, GSTF2	MAGIKVFGHP ASIATRRVLI ALHEKNLDFE LVHVELKDGE HKKEPFLSRN PFGQVPAFED GDLKLFESRA ITQYIAHRYE NQGTNLLQTD SKNISQYAIM AIGMQVEDHQ FDPVASKLAF EQIFKSIYGL TTDEAVVAEE EAKLAKVLDV YEARLKEFKY LAGETFTLTD LHHIPAIQYL LGTPTKKLFT ERPAVNEWVA EITKRPASEK VQ	3.77%
1	2	ATCG00280	PHOTOSYSTEM II REACTION CENTER PROTEIN C, PSBC	MKTLYSLRRF YHVETLFNGT LALAGRDQET TGFAWWAGNA RLINLSGKLL GAHVAHAGLI VFWAGAMNLF EVAHFVPEKP MYEQGLILLP HLATLGWGVG PGGEVIDTFP YFVSGVLHLI SSAVLGFGGI YHALLGPETL EESFPFFGYV WKDRNNKMTT LGIHLILLGV GAFLLVFKAL YFGGVDTWA PGGGDVRKIT NLTLSPSVIF GYLLKSPFGG EGWIVSVDDL EDIIGGHVWL GSICIFGGIW HILTKPFAWA RRALVWSGEA YLSYSLAALS VCGFIACCFV WFNNTAYPSE FYGPTGPEAS QAQAFTFLVR DQRLGANVGS AQGPTGLGKY LMRSPTGEVI FGGETMRFWD LRAPWLEPLR GPNGLDLSRL KKDIQPWQER RSAEYMTHAP LGSLNSVGGV ATEINAVNYV SPRSWLSTSH FVLGFFLFVG HLWHAGRARA AAAGFEKGID RDFEPVLSMT PLN	3.38%
1	1	AT3G58140	phenylalanyl-tRNA synthetase class IIc family protein	MTVFSVQSTI FSRASVALLS SNGFKRFSFV SSFSSSAAYS PPKMRKRRYP IVSAVDIGGV AIARNDVVRE DDPTNNVPDS IFSKLGMQLH RRDKHPIGIL KNAIYDYFDS NYSNKFEKFE DLSPIVTTEQ NFDDVLVPAD HVSRSLNDTY YVDSQTVLRC HTSAHQAELL RKGHSRFLVT GDVYRRDSID STHYPVFHQM EGFCVTSPED WNGSGKDSTL YAAEDLKKCL EGLARHLFGS VEMRWVDTYF PFTNPSFELE IYFKEDWLEV LGCGVTEQVI LKQSGLENNV AWAFGLGLER LAMVLFDIPD IRFFWSSDER FTSQFGKGEL GVKFKPYSKY PPCYKDISFW ISDLFTENNF CEVVRGIAGD LVEEVKLIDQ FTNKKKGLTS HCYRIVFRSM ERSLTDEEVN DLQSKVRDEV QKKLNVELR	3.03%
1	1	AT3G25520	ATL5, OLI5, OLIGOCELLULA 5, PGY3, PIGGYBACK3, RIBOSOMAL PROTEIN L5, RIBOSOMAL PROTEIN L5 A, RPL5A	MVFVKSTKSN AYFKRYQVKF RRRRDGKTDY RARIRLINQD KNKYNTPKYR FVVRFTNKDI VAQIVSASIA GDIVKASAYA HELPQYGLTV GLTNYAAAYC TGLLLARRVL KMLEMDDEYE GNVEATGEDF SVEPTDSRRP FRALLDVGLI RTTTGNRVFG ALKGALDGGL DIPHSDKRFA GFHKENKQLD AEIHRNYIYG GHVSNYMKLL GEDEPEKLQT HFSAYIKKGV EAESIEELYK KVHAAIRADP KQHKRYNLKK LTYEERKNKL IERVKLNGA GGDDDDEDDE E	2.99%
1	1	AT1G74910	KJC1, KONJAC 1	MGSSMEEKVV AVIMVGGPTK GTRFRPLSLN IPKPLFPIAG QPMVHHPISA CKRIPNLAQI YLVGFYEERE FALYVSAISH ELKVPVRYLR EDKPHGSAGG LYHFRNLIME DSPSHIFLLN CDVCCSFPLP KMLEAHRGYG GIGTLLVIKV SPESASQFGE LVADPVTNEL LHYTEKPETF VSDRINCGVY VFTPEIFNAI GDVSTQRKDR ATLKRVSSFE QLPAGKRKRL YYETMDFWEQ IKSPGMSLRC SGLYLSQPRL TSPQLLASGD GTRSAIVIGD VYIHPSAKVH PTAKIGPNVS ISANARVGPG VRLMSCIILD DVEIMENAVV TNAIVGWKSS IGRWSRVQAE GVYNSKLGVT ILGDSVAVED EVVVTSSIVL PNKTLNVSVQ DEIIL	2.89%
1	1	AT3G06720	AIMP ALPHA, AT-IMP, ATKAP ALPHA, IMPA-1, IMPA1, IMPORTIN ALPHA, IMPORTIN ALPHA ISOFORM 1	MSLRPNAKTE VRRNRYKVAV DAEEGRRRRE DNMVEIRKSK REESLMKKRR EGMQALQGFP SASAASVDKK LDSLKDMVAG VWSDDPALQL ESTTQFRKLL SIERSPPIEE VISAGVVPRF VEFLKKEDYP AIQFEAAWAL TNIASGTSDH TKVVIDHNAV PIFVQLLASP SDDVREQAVW ALGNVAGDSP RCRDLVLGCG ALLPLLNQLN EHAKLSMLRN ATWTLSNFCR GKPQPHFDQV KPALPALERL IHSDDEEVLT DACWALSYLS DGTNDKIQTV IQAGVVPKLV ELLLHMSPSV LIPALRTVGN IVTGDDIQTQ CVINSGALPC LANLLTQNHK KSIKKEACWT ISNITAGNKD QIQTVVEANL ISPLVSLLQN AEFDIKKEAA WAISNATSGG SHDQIKYLVE QGCIKPLCDL LVCPDPRIIT VCLEGLENIL KVGEAEKNLG HTGDMNYYAQ LIDDAEGLEK IENLQSHDNN EIYEKAVKIL ETYWLEEEDD ETQQPPGVDG SQAGFQFGGN QAPVPSGGFN FS	2.82%

Table 3 Continued.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
1	1	AT1G43190	POLYPYRIMIDINE TRACT-BINDING PROTEIN 3, PTB3	MAESSKVVMV RNVGHEISEN DLLQLFQPFG VITKLVNLRA KNQALLQMQD VSSAVSALQF FTNVQPTIRG RNVYVQFSSM QELTTIEQNI NGREDEPNRI LLVTIHMMLY PITVDVLHQV FSPYGFVEKL VTFQKSAGFQ ALIQYQVQQC AASARTALQG RNIYDGCCQL DIQFSNLEEL QVNYNNDRSR DYTNPNLPAE QNGRSSHPCY GDTGVAYPQM ANTSAIAAAF GGGLPPGITG TNDRCTVLVS NLNADSIDED KLFNLFSLYG NIVRIKLLRN KPDHALVQMG DGFQAELAVH FLKGAMLFGK RLEVNFSKHP NITPGTDSHD YVNSNLNREN RNAAKNYRYC CSPTKMIHLS TLPQDVTEEE VMNHVQEHGA VVNTKVFEMN GKKQALVQFE NEEEAAEALV CKHATSLGGS IIRISFSQLQ TI	2.78%
1	1	AT5G35790	APG1 G6PD1, GLUCOSE-6-PHOSPHATE DEHYDROGENASE 1	MATHSMIIPS PSSSSSSLAT AASPFKETLP LFSRSLTFPR KSLFSQVRLR FFAEKHSQLD TSNGCATNFA SLQDSGDQLT EEMVTKGEST LSITVVGASG DLAKKKIFPA LFALFYEGCL PQDFSVFGYA RTKLTMEELR DMISSTLTCR IDQREKCGDK MEQFLKRCFY HSGQYNSEED FAELNKKLKE KEAGKISNRL YYLSIPPNIF VDVVRCASLR ASSENGWTRV IVEKPFGRDS ESSGELTRCL KQYLTEEQIF RIDHYLGKEL VENLSVLRFS NLVFEPLWSR NYIRNVQLIF SEDFGTEGRG GYFDQYGIIR DIMQNHLLQI LALFAMETPV SLDAEDIRSE KVKVLRSMKP LRLEDVVVGQ YKGHNKGGKT YPGYTDDPTV PNHSLTPTFA AAAMFINNAR WDGVPFLMKA GKALHTRGAE IRVQFRHVPG NLYKKSFATN LDNATNELVI RVQPDEGIYL RINNKVPGLG MRLDRSDLNL LYRSRYPREI PDAYERLLLD AIEGERRLFI RSDELDAAWD LFTPALKELE EKKIIPELYP YGSRGPVGAH YLASKYNVRW GDLGEA	2.60%
1	1	AT2G04030	CR88 ATHSP90.5, ATHSP90C, CR88, EMB1956, EMBRYO DEFECTIVE 1956, HEAT SHOCK PROTEIN 88.1, HEAT SHOCK PROTEIN 90.5, HSP88.1, HSP90.5	MAPALSRSLY TSPLTSVPIT PVSSRLSHLR SSFLPHGGAL RTGVSCSWNL EKRCNRFAVK CDAAVAEKET TEEGSGERFE YQAEVSRLLD LIVHSLYSHK EVFLRELVSN ASDALDKLRF LSVTEPSLLG DGGDLEIRIK PDPDNGTITI TDTGIGMTKE ELIDCLGTIA QSGTSKFLKA LKENKDLGAD NGLIGQFGVG FYSAFLVAEK VVVSTKSPKS DKQYVWESVA DSSSYLIREE TDPDNILRRG TQITLYLRED DKYEFAESTR IKNLVKNYSQ FVGFPIYTWQ EKSRTIEVEE DEPVKEGEEG EPKKKKTTKT EKYWDWELAN ETKPLWMRNS KEVEKGEYNE FYKKAFNEFL DPLAHTHFTT EGEVEFRSIL YIPGMGPLNN EDVTNPKTKN IRLYVKRVFI SDDFDGELFP RVLSFVKGVV DSDDLPLNVS REILQESRIV RIMRKRLIRK TFDMIQEISE SENKEDYKKF WEMFGRFLKL GCIEDTGNHK RITPLLRFFS SKNEEELTSL DDYIENMGEN QKAIYYLATD SLKSAKSAPF LEKLIQKDIE VLYLVEPIDE VAIQNLQTYK EKKFVDISKE DLELGDEDEV KDREAKQEFN LLCDWIKQQL GDKVAKVQVS NRLSSSPCVL VSGKFGWSAN MERLMKAQAL GDTSSLEFMR GRRILEINPD HPIIKDLNAA CKNAPESTEA TRVVDLLYDT AIISSGFTPD SPAELGNKIY EMMAMAVGGR WGRVEEEEES STVNEGDDKS GETEVVEPSE VRAESDPWQD	2.44%
1	1	AT1G78900	VHA-A VACUOLAR ATP SYNTHASE SUBUNIT A, VHA-A	MPAFYGGKLT TFEDDEKESE YGYVRKVSGP VVVADGMAGA AMYELVRVGH DNLIGEIIRL EGDSATIQVY EETAGLTVND PVLRTHKPLS VELGPGILGN IFDGIQRPLK TIARISGDVY IPRGVSVPAL DKDCLWEPQP NKFVEGDTIT GGDLYATVFE NTLMNHLVAL PPDAMGKITY IAPAGQYSLK DTVIELEFQG IKKSYTMLQS WPVRTPRPVA SKLAADTPLL TGQRVLDALF PSVLGGTCAI PGAFGCGKTV ISQALSKYSN SDAVVYVGCG ERGNEMAEVL MDFPQLTMTL PDGREESVMK RTTLVANTSN MPVAAREASI YTGITIAEYF RDMGYNVSMM ADSTSRWAEA LREISGRLAE MPADSGYPAY LAARLASFYE RAGKVKCLGG PERNGSVTIV GAVSPPGGDF SDPVTSATLS IVQVFWGLDK KLAQRKHFPS VNWLISYSKY STALESFYEK FDPDFINIRT KAREVLQRED DLNEIVQLVG KDALAEGDKI TLETAKLLRE DYLAQNAFTP YDKFCPFYKS VWNMRNIIHF YNLANQAVER AAGMDGQKIT YTLIKHRLGD LFYRLVSQKF EDPAEGEDTL VEKFKKLYDD LNAGFRALED ETR	2.41%
1	1	AT4G28980	CAK1AT, CDK-ACTIVATING KINASE 1AT, CDKF;1, CYCLIN-DEPENDENT KINASE F;1	MDKQPATSWS IHTRPEIIAK YEIFERVGSG AYADVYRARR LSDGLIVALK EIFDYQSAFR EIDALTILNG SPNVVVMHEY FWREEENAVL VLEFLRSDLA AVIRDGKRKK KVEGGDGFSV GEIKRWMIQI LTGVDACHRN LIVHRDLKPG NMLISDDGVL KLADFGQARI LMEHDIVASD ENQQAYKLED KDGETSEPPE VIPDYENSPR QGSDGQEREA MSKDEYFRQV EELKAKQVVR DDTDKDSNVH DGDISCLATC TVSEMDDDLG RNSFSYDADE AVDDTQGLMT SCVGTRWFRP PELLYGSTMY GLEVDLWSLG CVFAELLSLE PLFPGISDID QISRVTNVLG NLNEEVWPGC VDLPDYKSIS FAKVESPLGI EGCLPNHSGD VISLLKKLIC YDPASRATTK EMLNDKYLSE EPLPVPVSEL YVPPTMSGPD EDSPRKWNDY REMDSDSDFD GFGPMNVKPT SSGFTIEFP	2.30%
1	1	AT5G11580	Regulator of chromosome condensation (RCC1) family protein	MEDKSPTLLI SEDLSRKIIS LAAGEAMTIA LTGDGCVYSW GRGMFGRLGT GKESDELVPV LVEFPNQAEG DRIRIVGVAA GAYHSLAVSD DGSVWCWGYN IYGQLGFDGE NSLAPCLIKN LFERETSSSS LNDSGREVGS DLKVCTVKAG SMMSLAIDNV GGLWMWGNVP PQDSEPDPRL SFTSIPIPFP ILDFHGRTVL KVACGDEHVV ALVGPGDIMK DNSYDVSVLY SWGNNHGQL GLGDGESRAR PQTVETFNQK SGLTVYDIAC GAHHTALLTY RKETPKGPSI CWTFGFGENG QLGHRSNKSS SLPEPVSDLP EHAYLVSVDC GLFHTSVVSS EGVVWSWGME RGLGLCPDVN FTEVEAGDDS VPRKISGGSS RFRDPVQVSC GAAHTVLVVD GGYKLWSWGR GRNGVLGTGN VSDCYVPTLV FWPNELKPEK EEVPDDGKSA STEEIKRLES KLMVMERYAS ILHGSIFGKP FNEEEDIPYS LRVSGYFDMG KEWGEMLESA DKSQLMRLQA FYEDMIGRVK DKVLQRRIQE IMKDCLQSSA PKY	1.99%
1	1	AT3G06510	ATSFR2, SENSITIVE TO FREEZING 2, SFR2	MELFALLIKV AGLLATVTVG ANVVSYSRFR RQNLAKFRSP IDESKEVLAD FNSIEHEEGK FFFGLATAPA HAEDDLDDAW LQFAKETPCS AEEAEAADKK ARRKKVKLAV GAITKGLAKN THGKEDKNAA DKPPSKNVAA WHNAPHAEDR LKFWSDPDKE VKLAKDTGVT VFRMGVDWSR IMPVEPTKGI KEAVNYEAVE MYKWILKKVR SNGMKVMLTL FHHSLPPWAA DYGGWKMEKT VDYFMDFTRI VVDSMYDLVD SWVTFNEPHI FTMLTYMCGS WPGNNPDFLE IATSTLPMGV FHRALHWMAV AHSKAYDYIH GKISLKKPLV GVAHHVSFMR PYGLFDIGAV TISNSLTIFP YIDSICEKLD FIGINYYGDG AVCGAGLKLV ETDEYSESGR GVYPDGLYRV LLMFHERYKH LKVPFIVTEN GVSDETDVIR RPYLIEHLLA LYAAMLKGVP VLGYIFWTIS DNWEWADGYG PKFGLVAVDR SHDLARTLRQ SYHLFSKIVR SGKVTRKDRS LAWNELQKAA KAGKLRPFYR GVDNHNLMYA DGLDKPQWRP FVDRDWRFGH YQMDGLQDPL SRVARTLLIW PLIMKKRIRK VKIKHTDDAG LVLHPALASP FD	1.93%
1	1	AT1G32230	ARABIDOPSIS THALIANA P8 (INTERACTING PROTEIN), ATP8, ATRCD1, CEO, CEO1, RADICAL-INDUCED CELL DEATH1, RCD1, REDOX IMBALANCED 1, RIMB1	MEAKIVKVLD SSRCEDGFGK KRKRAASYAA YVTGVSCAKL QNVPPPNGQC QIPDKRRRLE GENKLSAYEN RSGKALVRYY TYFKKTGIAK RVMQYENGEW NDLPEHVICA IQNELEEKSA AIEFKLCGHS FILDFLHMQR LDMETGAKTP LAWIDNAGKC FFPEIYESDE RTNYCHMKCV EDPKQNAPHD IKLRLEIDVN GGETPRLNLE ECSDESGDNM MDDVPLAQRS SNEHYDEATE DSCSRKLEAA VSKWDETDAI VVSGAKLTGS EVLDKDAVKK MFAVGTASLG NVPVLDVGRF SSEIAEARLA LFQKQVEITK KHRGDANVRY AWLPAKREVL SAVHMQGLGV GGAFIRKSIY GVGIHLTAAD CPYFSARYCD VDENGVRYMV LCRVIMGNME LLRGDKAQFF SGGEEYDNGV DDIESPKNYI VWNINMNTHI FPEFVVRFKL SNLPNAEGNL IAKRDNSGVT LEGPKDLPPQ LESNQGARGS GSANSVGSST TRPKSPWMPF PTLFAAISHK VAENDMLLIN ADYQQLRDKK MTRAEFVRKL RVIVGDDLLR STITTLQNQP KSKEIPGSIR DHEEGAGGL	1.87%
1	1	AT5G65010	ASN2, ASPARAGINE SYNTHETASE 2	MCGILAVLGC IDNSQARKSR IIELSRRLRH RGPDWSGLHC YEDCYLAHER LAIIDPTSGD QPLYNEDKTV AVTVNGEIYN HKILREKLKS HQFRTGSDCE VIAHLYEEHG EEFIDMLDGM FAFVLLDTRD KSFIAARDAI GITPLYIGWG LDGSVWFASE MKALSDDCEQ FMSFPPGHIY SSKQGGLRRW YNPPWYNEQV PSTPYDPLVL RNAFEKAVIK RLMTDVPFGV LLSGGLDSSL VAAVALRMLE KSEAARQWGS QLHTFCIGLQ GSPDLKAGRE VADYLGTRHH EFQFTVQDGI DAIEEVIYHI ETYDVTTIRA STPMFLMSRK IKSLGVKMVL SGEGSDEILG GYLYFHKAPN KKEFHEETCR KIKALHQFDC LRANKSTSAW GVEARVPFLD KEFLNVAMSI DPEWKLIKPD LGRIEKWVLR NAFDDEERPY LPKHILYRQK EQFSDGVGYS WIDGLKDHAN KHVSDTMLSN ASFVFPDNTP LTKEAYYYRT IFEKFFPKSA ARATVPGGPS IACSTAKAVE WDATWSKNLD PSGRAALGVH VAAYEEDKAA AAAKAGSDLV DPLPKNGT	1.73%
1	2	AT2G31810	ACT domain-containing small subunit of acetolactate synthase protein	MAAISVSSSP SIRCLRSACS DSSPALVSST RVSFPAKISY LSGISSHRGD EMGKRMEGFV RSVDGKISDA SFSEASSATP KSKVRKHTIS VFVGDESGMI NRIAGVFARR GYNIESLAVG LNRDKALFTI VVCGTERVLQ QVIEQLQKLV NVLKVEDISS EPQVERELML VKVNAHPESR AEIMWLVDTF RARVVDIAEH ALTIEVTGDP GKMIAVERNL KKFQIREIVR TGKIALRREK MGATAPFWRF SAASYPDLKE QAPVSVLRSS KKGAIVPQKE TSAGGDVIPV EPFFDPKVHR ILDAHWGLLT DEDTSGLRSH TLSLLVNDIP GVLMIVTGVF ARRGYNIQSL AVGHAETKGI SRITTVIPAT DESVSKLVQQ LYKLVDVHEV HDLTHLPFSE RELMLIKIAV NAAARRDVLD IASIFRAKAV DVSDHTITLQ LTGDLDKMVA LQRLLEPYGI CEVARTGRVA LARESGVDSK YLRGYSFPLT G	1.67%
1	1	AT4G24280	CHLOROPLAST HEAT SHOCK PROTEIN 70-1, CPHSC70-1	MASSAAQIHV LGGIGFASSS SSKRNLNGKG GTFMPRSAFF GTRTGPFSTP TSAFLRMGTR NGGGASRYAV GPVRVVNEKV VGIDLGTTNS AVAAMEGGKP TIVTNAEGQR TTPSVVAYTK SGDRLVGQIA KRQAVVNPEN TFFSVKRFIG RKMNEVDEES KQVSYRVVRD ENNNVKLECP AINKQFAAEE ISAQVLRKLV DDASRFLNDK VTKAVITVPA YFNDSQRTAT KDAGRIAGLE VLRIINEPTA ASLAYGFDRK ANETILVFDL GGGTFDVSVL EVGDGVFEVL STSGDTHLGG DDFDRKVVDW LAAEFKKDEG IDLLKDKQAL QRLTEAAEKA KIELSSLTQT NMSLPFITAT ADGPKHIETT LTRAKFEELC SDLLDRVRTP VENSLRDAKL SFKDIDEVIL VGGSTRIPAV QELVRKVTGK EPNVTVNPDE VVALGAAVQA GVLAGDVSDI VLLDVTPLSI GLETLGGVMT KIIPRNTTLP TSKSEVFSTA ADGQTSVEIN VLQGEREFVR DNKSLGSFRL DGIPPAPRGV PQIEVKFDID ANGILSVSAV DKGTGKKQDI TITGASTLPK DEVDQMVQEA ERFAKDDKEK RDAIDTKNQA DSVVYQTEKQ LKELGEKIPG EVKEKVEAKL QELKDKIGSG STQEIKDAMA ALNQEVMQIG QSLYNQPGAG GPGAGPSPGG EGASSGDSSS SKGGDGDDVI DADFTDSQ	1.67%

Table 3 Continued.

Peptides		Gene Symbol	Description	Protein Sequence, unique reference, other references	Protein Coverage
Unique	Total				
1	1	AT3G48560	ACETOHYDROXY ACID SYNTHASE, ACETOLACTATE SYNTHASE, AHAS, ALS, CHLORSULFURON/IMIDAZOLINONE RESISTANT 1, CSR1, IMIDAZOLE RESISTANT 1, IMR1, TRIAZOLOPYRIMIDINE RESISTANT 5, TZP5	MAAATTTTTT SSSISFSTKP SPSSSKSPLP ISRFSLPFSL NPNKSSSSSR RRGIKSSSPS SISAVLNTTT NVTTTPSPTK PTKPETFISR FAPDQPRKGA DILVEALERQ GVETVFAYPG GASMEIHQAL TRSSSIRNVL PRMEQGGVFA AEGYARSSGK PGICIATSGP GATNLVSGLA DALLDSVPLV AITGQVPRRM IGTDAFQETP IVEVTRSITK HNYLVMDVED IPRIIEEAFF LATSGRPGPV LVDVPKDIQQ QLAIPNWEQA MRLPGYMSRM PKPPEDSHLE QIVRLISESK KPVLYVGGGC LNSSDELGRF VELTGIPVAS TLMGLGSYPC DDELSLHMLG MHGTVYANYA VEHSDLLLAF GVRFDDRVTG KLEAFASRAK IVHIDIDSAE IGKNKTPHVS VCGDVKLALQ GMNKVLENRA EELKLDFGVW RNELNVQKQK FPLSFKTFGE AIPPQYAIKV LDELTDGKAI ISTGVGQHQM WAAQFYNYKK PRQWLSSGGL GAMGFGLPAA IGASVANPDA IVVDIDGDGS FIMNVQELAT IRVENLPVKV LLLNNQHLGM VMQWEDRFYK ANRAHTFLGD PAQEDEIFPN MLLFAAACGI PAARVTKKAD LREAIQTMLD TPGPYLLDVI CPHQEHVLPM IPSGGTFNDV ITEGDGRIKY	1.34%
1	1	AT1G06220	CLO, CLOTHO, GAMETOPHYTE FACTOR 1, GFA1, MATERNAL EFFECT EMBRYO ARREST 5, MEE5	MESSLYDEFG NYVGPEIESD RDSDDEVEDE DLQDKHLEEN GSDGEQGPGG SNGWITTIND VEMENQIVLP EDKKYYPTAE EVYGEDVETL VMDEDEQPLE QPIIKPVRDI RFEVGVKDQA TYVSTQFLIG LMSNPALVRN VALVGHLQHG KTVFMDMLVE QTHHMSTFNA KNEKHMKYTD TRVDEQERNI SIKAVPMSLV LEDSRSKSYL CNIMDTPGHV NFSDEMTASL RLADGAVLIV DAAEGVMVNT ERAIRHAIQD HLPIVVVINK VDRLITELKL PPRDAYYKLR HTIEVINNHI SAASTTAGDL PLIDPAAGNV CFASGTAGWS FTLQSFAKMY AKLHGVAMDV DKFASRLWGD VYYHSDTRVF KRSPPVGGGE RAFVQFILEP LYKIYSQVIG EHKKSVETTL AELGVTLSNS AYKLNVRPLL RLACSSVFGS ASGFTDMLVK HIPSPREAAA RKVDHSYTGT KDSPIYESMV ECDPSGPLMV NVTKLYPKSD TSVFDVFGRV YSGRLQTGQS VRVLGEGYSP EDEEDMTIKE VTKLWIYQAR YRIPVSSAPP GSWVLIEGVD ASIMKTATLC NASYDEDVYI FRALQFNTLP VVKTATEPLN PSELPKMVEG LRKISKSYPL AITKVEESGE MTILGTGELY LDSIMKDLRE LYSEVEVKVA DPVVSFCETV VESSSMKCFA ETPNKKNKIT MIAEPLDRGL AEDIENGVVS IDWNRKQLGD FFRTKYDWDL LAARSIWAFG PDKQGPNILL DDTLPTEVDR NLMMAVKDSI VQGFQWGARE GPLCDEPIRN VKFKIVDARI APEPLHRGSG QMIPTARRVA YSAFLMATPR LMEPVYYVEI QTPIDCVTAI YTVLSRRRGH VTSDVPQPGT PAYIVKAFLP VIESFGFETD LRYHTQGQAF CLSVFDHWAI VPGDPLDKAI QLRPLEPAPI QHLAREFMVK TRRRKGM3ED VSGNKFFDEA MGVELAQQTG DLHLQMI	1.32%
1	1	AT3G20630	UBP12 ATUBP14, DA3, PER1, PHOSPHATE DEFICIENCY ROOT HAIR DEFECTIVE1, TITAN6, TTN6, UBIQUITIN-SPECIFIC PROTEASE 14, UBP14	MTHMTPPPVD QPEDEEMLVP NSDLVDGPAQ PMEVTQPETA ASTVENQPAE DPPTLKFTWT IPNFSRQNTR KHYSDVFVVG GYKWRILIFP KGNNVDHLSM YLDVSDAASL PYGWSRYAQF SLAVVNQINT RYTVRKETQH QFNARESDWG FTSFMPLSEL YDPSRGYLVN DTVLVEAEVA VRKVLDYWSY DSKKETGFVG LKNQGATCYM NSLLQTLYHI PYFRKAVYHM PTTENDAPTA SIPLALQSLF YKLQYNDTSV ATKELTKSFG WDTYDSFMQH DVQELNRVLC EKLEDKMKGT VVEGTIQQLF EGHMMNYIEC INVDFKSTRK ESFYDLQLDV KGCKDVYASF DKYVEVERLE GDNKYHAEGH GLQDAKKGVL FIDFPPVLQL QLKRFEYDFM RDTMVKINDR YEFPLELDLD REDGKYLSPD ADRSVRNLYT LHSVLVHSGG VHGGHYYAFI RPTLSDQWYK FDDERVTKED LKRALEEQYG GEEELPQTNP GFNNNPPFKF TKYSNAYMLV YIRESDKDKI ICNVDEKDIA EHLRVRLKKE QEEKEDKRRY KAQAHLYTII KVARDEDLKE QIGKDIYFDL VDHDKVRSFR IQKQTPFQQF KEEVAKEFGV PVQLQRFWIW AKRQNHTYRP NRPLTPQEEL QPVGQIREAS NKANTAELKL FLEVEHLDLR PIPPPEKSKE DILLFFKLYD PEKAVLSYAG RLMVKSSSKP MDITGKLNEM VGFAPDEEIE LFEEIKFEPC VMCEHLDKKT SFRLCQIEDG DIICFQKPLV NKEIECLYPA VPSFLEVYQN RQLVRFRALE KPKEDEFVLE LSKQHTYDDV VEKVAEKLGL DDPSKLRLTS NNCYSQQPKP QPIKYRGVDM LSDMLVHYNQ TSDILYYEVL DIPLPELQGL KTLKVAFMHA TKEEVVIHNI RLPKQSTVGD VINELKTKVE LSMPDAELRL LEVFYHKIYK IFPSTERIEN INDQYWTLRA EEIPEEEKNI GPNDRLILVY KFAKETGQNQ QVQNFGEPFF LVIHEGETLE EIKNRIQKKL HVSDEDFAKW KFAFMSMGRP EYLQDTDVYG NRFQRRDVIG AFEQYLGLEH ADTTPKRAYA ANQNRHAYEK PVKIYN	0.99%
1	1	AT1G29900	CARB, CARBAMOYL PHOSPHATE SYNTHETASE B, VEN3, VENOSA 3	MRNHCLELSS NCSSIFASSK SNPRFSPSKL SYSTFFSRSA IYYRSKPKQA SSSSSFSTFP PCLNRKSSLT HVLKPVSELA DTTTKPFSPE IVGKRTDLKK IMILGAGPIV IGQACEFDYS GTQACKALRE EGYEVILINS NPATIMTDPE TANRTYIAPM TPELVEQVIE KERPDALLPT MGGQTALNLA VALAESGALE KYGVELIGAK LGAIKKAEDR GLKTPPSGIG TTLDECFDIA EKIGEFPLII RPAFTLGGTG GGIAYNKEEF ESICKSGLAA SATSQVLVEK SLLGWKEYEL EVMRDLADNV VIICSIENID PMGVHTGDSI TVAPAQTLTD REYQRLRDYS IAIIREIGVE CGGSNVQFAV NPVDGEVMII EMNPRVSRSS ALASKATGFP IAKMAAKLSV GYTLDQIPND ITRKTPASFE PSIDYVVTKI PRFAFEKFPG SQPLLTTQMK SVGESMALGR TFQESFQKAL RSLECGFSGW GCAKIKELDW DWDQLKYSLR VPNPDRIHAI YAAMKKGMKI DEIYELSMVD KWFLTQLKEL VDVEQVLMSG TLSEITKEDL YEVKKRGFSD KQIAFATKTT EEEVRTKRIS LGVVPSVKRV DTCAAEFEAH TPYMYSSYDV ECESAPNNKK KVLILGGGPN RIGQGIEFDY CCCHTSFALQ DAGYETIMLN SNPETVSTDY DTSDRLYFEP LTIEDVLNVI DLEKPDGIIV QFGGQTPLKL ALPIKHYLDK HMPMSLSGAG PVRIWGTSPD SIDAAEDRER FNAILDELKI EQPKGGIAKS EADALAIAKE VGYPVVVRPS YVLGGRAMEI VYDDSRLITY LENAVQVDPE RPVLVDKYLS DAIEIDVDTL TDSYGNVVIG GIMEHIEQAG VHSGDSACML PTQTIPASCL QTIRTWTIKL AKKLNVCGLM NCQYAITTSG DVFLLEANPR ASRTVPFVSK AIGHPLAKYA ALVMSGKSLK DLNFEKEVIP KHVSVKEAVF PFEKFQGCDV ILGPEMRSTG EVMSISSEFS SAFAMAQIAA GQKLPLSGTV FLSLNDMTKP MLEKIAVSFL ELGFKIVATS GTAMFLELKG IPVERVLKLH EGRPHAADMV ANGQIMLMLI TSSGDALDQK DGRQLRQMAL AYKVPVITTV AGALATAEGI KSLKSSAIKM TALQDFFEVK NVSSLLV	0.84%
1	1	AT1G20960	BRR2, EMB1507, EMBRYO DEFECTIVE 1507	MANLGGGAEA HARFKQVEYR ANSSLVLTTD NRPRDTHEPT GEPETLWGKI DPRSFGDRVA KGRPQELEDK LKKSKKKERD VVDDMVNIRQ SKRRRLREES VLTDTDDAVY QPKTKETRAA YEAMLGLIQK QLGGQPPSIV SGAADEILAV LKNDAFRNPE KKMEIEKLLN KIENHEFDQL VSIGKLITDF QEGGDSGGGR ANDDEGLDDD LGVAVEFEEN EEDDEESDPD MVEEDDDEED DEPTRTGGMQ VDAGINDEDA GDANEGTNLN VQDIDAYWLQ RKISQAYEQQ IDPQQCQVLA EELLKILAEG DDRVVEDKLL MHLQYEKFSL VKFLLRNRLK VVWCTRLARA EDQEERNRIE EEMRGLGPEL TAIVEQLHAT RATAKEREEN LQKSINEEAR RLKDETGGDG GRGRRDVADR DSESGWVKGQ RQMLDLESLA FDQGGLLMAN KRCDLPPGSY RSHGKGYDEV HVPWVSKKVD RNERLVKITE MPDWAQPAFK GMQQLNRVQS KVYDTALFKA ENILLCAPTG AGKTNVAMLT ILQQLEMNRN TDGTYNHGDY KIVYVAPMKA LVAEVVGNLS NRLKDYGVIV RELSGDQSLT GREIEETQII VTTPEKWDII TRKSGDRTYT QLVRLLIIDE IHLLHDNRGP VLESIVARTL RQIETTKENI RLVGLSATLP NYEDVALFLR VDLKKGLFKF DRSYRPVPLH QQYIGISVKK PLQRFQLMND LCYQKVLAGA GKHQVLIFVH SRKETSKTAR AIRDTAMAND TLSRFLKEDS VTRDVLHSHE DIVKNSDLKD ILPYGFAIHH AGLSRGDREI VETLFSQGHV QVLVSTATLA WGVNLPAHTV IIKGTQVYNP EKGAWMELSP LDVMQMLGRA GRPQYDQHGE GIIITGYSEL QYYLSLMNEQ LPIESQFISK LADQLNAEIV LGTVQNAREA CHWLGYTYLY IRMVRNPTLY GLAPDALAKD VVLEERRADL IHSAATILDK NNLVKYDRKS GYFQVTDLGR IASYYYITHG TIATYNEHLK PTMGDIDLYR LFSLSDEFKY VTVRQDEKME LAKLLDRVPI PIKETLEEPS AKINVLLQAY ISQLKLEGLS LTSDMVYITQ SAGRLVRALY EIVLKRGWAQ LAEKALNLSK MVGKRMWSVQ TPLRQFHGLS NDILMQLEKK DLVWERYYDL SAQELGELIR SPKMGKPLHK FIHQFPKVTL SAHVQPITRT VLNVELTVTP DFLWDEKIHK YVEPFWIIVE DNDGEKILHH EYFLLKKQYI DEDHTLHFTV PIFEPLPPQY GSETVLPVSF RHLILPEKYP PPTELLDLQP LPVTALRNPN YEILYQDFKH FNPVQTQVFT VLYNTNDNVL VAAPTGSGKT ICAEFAILRN HHEGPDATMR VVYIAPLEAI AKEQFRIWEG KFGKGLGLRV VELTGETALD LKLLEKGQII ISTPEKWDAL SRSWKQRKYV QQVSLFIVDE LHLIGGQHGP VLEVIVSRMR YISSQVINKI RIVALSTSLA NAKDLGEWIG ASSHGLFNFP PGVRPVFLEI HIQGVDISSF EARMQAMTKP TYTAIVQHAK NKKPAIVFVP TRKHVRLTAV DLMAYSHMDN PQSPDFLLGK LEELDPFVEQ IREETLKETL CHGIGYLHEG LSSLDQEIVT QLFEAGRIQV CVMSSSLCWG TPLTAHLVVV MGTQYYDGRE NSHSDYFVPD LLQMMGRASR PLLDNAGKCV IFCHAPRKEY YKKFLYEAFP VESQLQHFLH DNFNAEVVAG VIENKQDAVD YLTWTFMYRR LPQNPNYYNL QGVSHRHLSD HLSELVENTL SDLEASKCIE VEDEMELSPL NLGMIASYYY ISYTTIERFS SLLSSKTKMK GLLEILTSAS EYDMIPIRPG EEDTVRRLIN HQRFSFENPK CTDPHVKANA LLQAHFSRQN IGGNLAMDQR DVLLSATRLL QAMVDVISSN GWLNLALLAM EVSQMVTQGM WERDSMLLQL PHFTKDLAKR CQENPGKNIE TVFDLVEMED EERQELLKMS DAQLLDIARF CNRFPNIDLT YEIVGSEEVN PGKEVTLQVM LERDMEGRTE VGPVDSLRYP KTKEEGWWLV VGDTKTNQLL AIKRVSLQRK VKVKLDFTAP SEPGEKSYTL YFMCDSYLGC DQEYSFSVDV KGSGAGDRME E	0.46%