


Winter 12-15-2017

Analysis of Argonaute-Small RNA-Transcription Factor Circuits Controlling Leaf Development

John Steen Hoyer

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

 Part of the [Agriculture Commons](#), [Developmental Biology Commons](#), [Genetics Commons](#), and the [Plant Sciences Commons](#)

Recommended Citation

Hoyer, John Steen, "Analysis of Argonaute-Small RNA-Transcription Factor Circuits Controlling Leaf Development" (2017). *Arts & Sciences Electronic Theses and Dissertations*. 1188.
https://openscholarship.wustl.edu/art_sci_etds/1188

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

Washington University in St. Louis
Division of Biology and Biomedical Sciences
Computational and Systems Biology

Dissertation Examination Committee:
James C. Carrington, Chair
Douglas L. Chalker
Barbara Kunkel
Blake C. Meyers

Analysis of Argonaute-Small RNA-Transcription Factor Circuits Controlling Leaf Development

by

John Steen Hoyer

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2017
Saint Louis, Missouri

© 2017, John Steen Hoyer

Contents

List of Figures	v
List of Tables	vii
Acknowledgments	viii
Abstract	x
1 Introduction	1
1.1 Silencing mechanisms in eukaryotes	2
1.1.1 Diverse small RNAs are processed from structured RNA	4
1.1.2 AGO-small RNA complexes have specialized molecular functions	6
1.1.3 Many plant miRNA cleave conserved TF mRNA targets	8
1.2 AGO-small RNA-TF modules control plant form	10
1.2.1 The miR166-HD-ZIP module controls shoot identity and polarity	11
1.2.2 The AGO7-miR390-TAS3-ARF module controls polarity and timing	13
1.2.3 The miR156-SPL module controls heteroblasty and flowering	14
1.2.4 The miR319-TCP module controls organ margins and maturation	16
1.3 Gaps in our knowledge of gene regulation in plants	17
1.3.1 AGO genes as key control points for polarity and other processes	18
1.3.2 New resources for promoter analysis	19
1.3.3 New tools for automated measurement of growth and development	19
1.3.4 Overview of contributions	20
2 Direct regulation of ARGONAUTE7 by miRNA-targeted transcription factors	22
2.1 Abstract	23
2.2 Introduction	23
2.3 Results	25
2.3.1 Multiple SPLs and TCPs bind the AGO7 promoter	25
2.3.2 SPL binding sites are not required for polar AGO7 transcription	35
2.3.3 SPL and TCP binding sites are not strictly required for AGO7 function	40
2.4 Discussion	48
2.5 Follow-up experiments	50
2.6 Methods	52
2.6.1 Plasmid construction	52

2.6.2	Y1H screens	52
2.6.3	Plant materials and growth conditions	55
2.6.4	<i>ago7</i> mutant complementation tests	55
2.6.5	GUS assays	56
2.7	Acknowledgments	56
3	Improved methods for time-lapse imaging of vegetative growth	57
3.1	Abstract	57
3.2	Introduction	58
3.2.1	Biological context and approach	59
3.3	Methods	60
3.3.1	Growth conditions	60
3.3.2	Hardware and software configuration	60
3.4	Results and Discussion	64
3.4.1	Reliable acquisition and transfer of images	64
3.4.2	Image processing	65
3.4.3	Rosette size and shape dynamics in <i>ago7</i> mutants	70
3.5	Possible improvements and related work	75
3.5.1	Phenotiki resources: comparison and priorities for benchmarking	75
3.5.2	Stereo photography, distortion correction, and color accuracy	77
3.5.3	Metadata standards, monitoring, and experimental design	77
3.5.4	Similar datasets, video analysis, and leaf phyllotactic order	79
3.5.5	Historical note and personal reflections (part 1)	82
4	Prospects and conclusion	85
4.1	How do <i>TAS3</i> tasiRNAs function in patterning?	86
4.2	Inputs and outputs of SPL and ARF circuits	89
4.3	Methods for dissecting enhancer regions in plants	91
4.4	Molecular functions of AGOs in development	94
4.5	Personal reflections (part 2) and coda	97
Appendix A	Roles and programming of <i>Arabidopsis</i> ARGONAUTE proteins during <i>Turnip mosaic virus</i> infection	100
A.1	Preface	100
Appendix B	Raspberry Pi powered imaging for plant phenotyping	131
B.1	Preface	131
Appendix C	Online documentation: Configuration for imaging plants with Raspi- Cams using <i>raspistill</i>, <i>cron</i>, and <i>rsync</i>	185
C.1	Purpose	185
C.1.1	Links	186

C.2	cron tables	187
C.2.1	Image capture schedule	187
C.2.2	Lights and wifi	188
C.2.3	Image transfer	188
C.2.4	Helper scripts	189
C.3	Procedures	189
C.3.1	Starting imaging	189
C.3.2	Ending imaging	190
C.3.3	Pitfalls	190
C.3.4	Transfer schedule	192
C.3.5	Monitoring	193
C.4	Plans	194
Appendix D	PlantCV v2: Image analysis software for high-throughput plant pheno-	
	typing	195
D.1	Preface	195
Bibliography	235

List of Figures

1.1	AGO-small RNA complexes control cellular and morphological processes by repressing master regulator TFs	2
1.2	Core steps of RNA silencing: Dicing, slicing, and amplification	3
1.3	Main silencing pathways in <i>A. thaliana</i>	5
1.4	Cladogram of the ten <i>AGO</i> genes of <i>A. thaliana</i>	7
1.5	AGO7-triggered biogenesis of <i>TAS3</i> tasiRNAs	8
1.6	Specification of leaf polarity by mutually antagonistic TFs and small RNAs	12
1.7	Decrease in miR156 levels and concomitant increase in SPL levels triggers leaf shape changes as plants age	15
2.1	SPL and TCP TFs bind the <i>AGO7</i> promoter in yeast.	26
2.2	Scatterplots of β -gal activities with likely nonspecific activations indicated for <i>AGO1</i> promoter fragment screens	29
2.3	Scatterplots of β -gal activities with likely nonspecific activations indicated for <i>AGO1</i> promoter fragment screens	30
2.4	Scatterplots of β -gal activities with likely nonspecific activations indicated for <i>AGO10</i> promoter fragment screens	31
2.5	Retesting binding by multiple SPLs and TCPs with alternate Y1H reporter	32
2.6	Identification of SPL11 binding sites	34
2.7	Identification of TCP binding sites	36
2.8	Testing of TCP binding sites in yeast	37
2.9	SPL binding sites are not required for polar <i>AGO7</i> transcription	38
2.10	Preliminary evidence that a short TSS-proximal region is sufficient for polar <i>AGO7</i> transcription	39
2.11	GUS reporter gene activity driven by truncated <i>AGO7</i> promoter in seedlings. . . .	40
2.12	Transgenic complementation of <i>ago7</i> downward-curved leaf phenotype	42
2.13	Leaf shape defects in transformed <i>ago7</i> mutants 1: leaf blade length to width ratio .	43
2.14	Alternate quantification of leaf shape defects in transformed <i>ago7</i> mutants with calipers	44
2.15	Leaf shape defects in transformed <i>ago7</i> mutants 2: leaf blade length to petiole length ratio	45
2.16	Transgenic complementation of <i>ago7</i> early abaxial trichome appearance	47
2.17	Model for structure of the <i>AGO7</i> proximal promoter	49
3.1	Photographs of mounted Raspberry Pi Camera imaging rigs	61

3.2	Schematic of plant arrangement illustrating overlap between camera fields of view.	63
3.3	Flow diagram illustrating steps in image processing	66
3.4	Top-down view of the course of <i>ago7</i> mutant development	68
3.5	Examples of two types of rosette missegmentation	69
3.6	Shape and size of wild-type and <i>ago7</i> mutant rosettes over time	71
3.7	Correlation between leaf-level measurements (blade and petiole lengths) in flatbed scans and top-down photos	74
3.8	Example illustrating leaf-level segmentation and misannotation	80
4.1	Revised model for control of heteroblasty and likely activation of <i>AGO7</i> by SPLs .	90
4.2	Schematic illustrating our changing understanding of small RNA targeting outcomes	95
A.1	Summary of polarity phenotypes resulting from altered <i>AGO10</i> levels	102

List of Tables

1.1	<i>MIRNA</i> gene families conserved in plants and the developmental functions of their conserved TF targets	9
2.1	Transcription factors included in Y1H prey library: counts by family.	27
2.2	Highest-scoring PWM scan match positions for <i>AGO7</i> , within first 425 bp.	50
2.3	Oligonucleotide sequences used for <i>AGO</i> promoter TOPO cloning	53
2.4	Oligonucleotide sequences directly cloned for <i>AGO7</i> promoter mutation analysis in yeast	54
2.5	Forward primer sequences used for TOPO cloning of truncated versions of the <i>AGO7</i> promoter	54
3.1	Components for an individual Raspberry Pi/Camera imaging rig	62

Acknowledgments

Many people have provided advice and encouragement over the years. I am particularly indebted to J.C. Carrington and members of the Carrington, Nusinow, Meyers, Mockler, Kellogg, Gehan, Chitwood, and Bart labs. Input from Dmitri Nusinow, Noah Fahlgren, Malia Gehan, César Lizárraga, Leonardo Chavez, Howard Berg, and others was essential for developing the imaging methodology described in Chapter 3. See also Appendices B and D. I thank Arash Abbasi, Sona Pandey, Swarup Roy Chaudhury, and Melinda Wilson for their patient collaborative help. I thank the Danforth Center Facilities and Support Services group for maintenance and help with growth chamber configuration, and Plant Growth Facility staff for excellent plant care. I am grateful for the advice provided by the members of my original research advisory committee (B.A. Cohen, G.D. Stormo, L.C. Strader, S.J. Swamidass, and D. Wang), and I thank the members of my subsequent committee for advice and helpful discussions.

The work described in Chapter 2 was enabled by a collaboration with Steve A. Kay and his postdocs Jose Pruneda-Paz and Ghislain Breton, who have gone on to establish their own laboratories. Many students assisted with the experimentation and other support tasks for that manuscript, including Mariah A. Hassert, Emily E. Holcomb, Halley Fowler, Kaylyn Bauer, and Jacob Mreen. See also section 2.7.

Figures 1.2 and 1.3 were adapted from diagrams originally made by J.C. Carrington. The first iterations of Figures 1.4 and 1.5 were made by N. Fahlgren and T.A. Montgomery, respectively. I created the other figures and am the primary author of all text in the body of this document. My contributions to the three papers included as appendices (A, B, and D) are detailed in the corresponding preface for each.

I am grateful to Danforth Center donors and US taxpayers for their support, including via National Science Foundation award 1330562 (to J.C. Carrington and M.F. Yanofsky) and a Graduate Research Fellowship (award 1143954).

John Steen Hoyer

Washington University in Saint Louis

December 2017

ABSTRACT OF THE DISSERTATION

Analysis of Argonaute-Small RNA-Transcription Factor Circuits Controlling Leaf Development

by

John Steen Hoyer

Doctor of Philosophy in Biology and Biomedical Sciences

Computational and Systems Biology

Washington University in St. Louis, December 2017

James C. Carrington, Chair

Experimental studies of plant development have yielded many insights into gene regulation, revealing interactions between core transcriptional and post-transcriptional regulatory pathways present in all land plants. This work describes a direct connection between the three main small RNA-transcription factor circuits controlling leaf shape dynamics in the reference plant *Arabidopsis thaliana*. We used a high-throughput yeast 1-hybrid platform to identify factors directly binding the promoter of the highly specialized *ARGONAUTE7* silencing factor. Two groups of developmentally significant microRNA-targeted transcription factors were the clearest hits from these screens, but transgenic complementation analysis indicated that their binding sites make only a small contribution to *ARGONAUTE7* function, possibly indicating a role in fine tuning. Timelapse imaging methodology developed to quantify these small differences may have broad utility for plant biologists. Our analysis also clarified requirements for polar transcription of *ARGONAUTE7*. This work has implications for our understanding of patterning in land plants.

Chapter 1

Introduction

Biological form and function require spatial and temporal control of gene expression. This control is achieved using combinatorial regulation at several different steps, including transcription, RNA maturation, translation, post-translational modification, and turnover. Some of these elaborated mechanisms may have facilitated the emergence of multicellularity.

RNA silencing processes operate along a continuum from transcriptional to post-transcriptional regulation in nearly all eukaryotes. In plants and animals, one important function of RNA silencing is post-transcriptional regulation of sequence-specific transcription factors (TFs). Diversification of TF molecular complexes and chromatin remodeling factors appears correlated with organismal complexity [1]. The DNA and RNA binding sites of sequence-specific regulators often control the expression of nearby genes, that is, they act *in cis*. Changes in *cis*-regulatory elements such as the binding sites of TFs and microRNAs (miRNAs, described below) provide one mechanism for morphological evolution [2, 3], via changes in the timing and location of gene expression.

This chapter introduces key RNA-guided regulatory pathways and describes the developmental context in which they operate. I describe several conserved small RNA-TF modules that act near the top of regulatory hierarchies controlling several aspects of plant growth (Figure 1.1). These

examples show the importance of understanding the transcriptional basis for specialization among the ARGONAUTES (AGOs) that effect small RNA function, and also highlight the importance and the utility of studying vegetative growth using the model plant *Arabidopsis thaliana*. Building on several recent reviews [4–6], I make the case that densely connected sets of conserved miRNA-TF pairs define core networks controlling the development of leaves and other organs. I highlight the involvement of the *AGO7* gene in both timing and patterning, and outline why *AGO7* was chosen for study in this work described here.

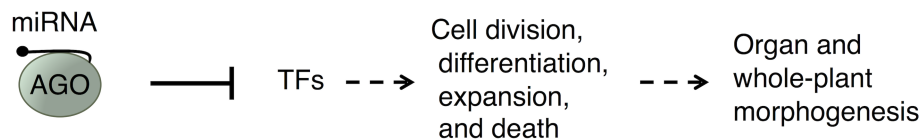


Figure 1.1: Many AGO-small RNA complexes control cellular and morphological processes by repressing master regulator TFs.

1.1 Silencing mechanisms in eukaryotes

Small RNA molecules guide a number of sequence-specific responses; the core molecular mechanisms common to silencing pathways are schematically illustrated in Figure 1.2. At the highest level, small RNAs can be classified based on the type of structured RNA from which they are processed: short interfering RNAs (siRNA) are produced from double-stranded RNA, whereas microRNAs (miRNA) are excised from imperfect hairpin foldbacks [7, 8]. miRNA were initially identified as regulators of developmental timing in the nematode *Caenorhabditis elegans* [9, 10], and subsequently shown to regulate a very broad array of processes. Chemically similar siRNA molecules were identified during analyses of transgenic and virus-infected plants [11], which was an important indication that these RNAs function in diverse eukaryotes.¹

¹ There are several good historical reviews of this subject, including references 12–15.

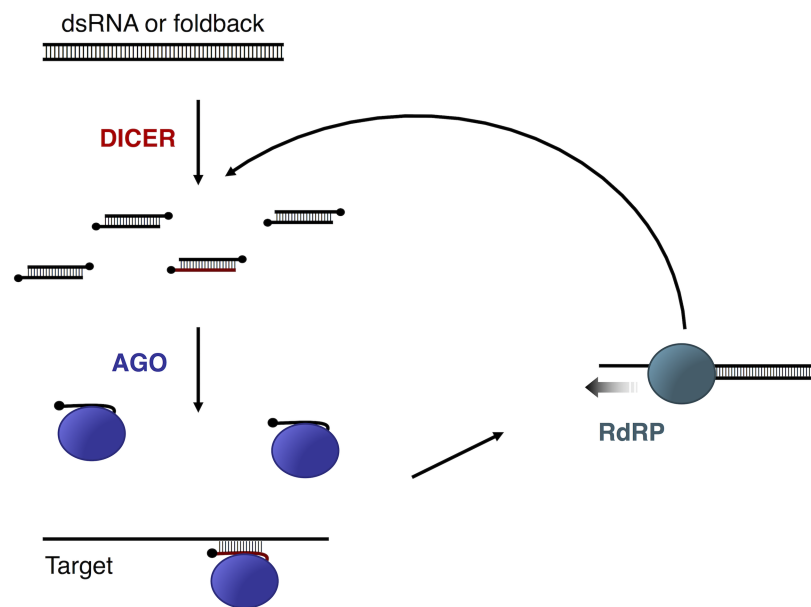


Figure 1.2: Core steps of RNA silencing: DICER proteins cleave double-stranded (ds)RNA and self-complementary RNA foldbacks into small RNA duplexes. These small RNAs are loaded into ARGONAUTE (AGO) proteins and guide targeting of other RNA molecules via base-pairing. AGO targeting can lead to endonucleolytic cleavage and/or other molecular outcomes. RNA-dependent RNA polymerase (RDR/RdRP) proteins convert some small RNA targets into dsRNA, yielding new substrates for DICER proteins.

1.1.1 Diverse small RNAs are processed from structured RNA

Distinct biogenesis pathways generate small RNAs that guide specific AGO molecular functions (Figure 1.3). Plants have several specialized DICER-like (DCL) proteins that produced small RNAs from structured RNA [16] whereas many animal lineages have only a single DICER [17]. siRNAs often function in sequence-specific defense against nucleic acid parasites (via RNA cleavage or repressive DNA methylation), whereas miRNA generally regulate endogenous genes.² No *MIRNA* gene families are conserved between eukaryote kingdoms, and there are a number of differences between miRNA biogenesis in plants and animals [18]. This pattern suggests that antiviral or antitransposon silencing pathways were co-opted for endogenous gene regulation independently in both animal and plant lineages.³ Silencing can be conceptualized as having three main functions.

1. Transposon control: RDR2 and DCL3 act with the AGO4 clade to control *de novo* DNA methylation, which generally represses transcription [21].⁴ Specialized plant-specific polymerases (Pol IV and Pol V) act in a large multisubunit complex to synthesize transcripts, for both targeting and dsRNA synthesis [24]. Plants lack proteins from the PIWI subfamily of AGOs, but 24-nt siRNAs function somewhat similarly to PIWI-interacting (pi)RNAs to control transposons in plant reproductive cells [25–27]. Transcriptional silencing can act on transgenes and sequence-similar endogenous loci, possibly because of their resemblance to highly-transcribed transposons or virus

² The term “RNA interference” is often used to explicitly refer to siRNA but not miRNA pathways; usage is not consistent. I use the broader term “RNA silencing” [14] to emphasize the commonality and connections between transcriptional and *post*-transcriptional silencing (Figure 4.2). Plant phased siRNAs often have regulatory functions that are not defensive *per se*, as described in sections 1.1.2 and 1.2.2.

³ Another possibility is that miRNA regulation was functional in a common ancestor of plants, animals, and possibly other eukaryotes, and that any ancestral *MIRNA* gene families have either been lost or obscured by sequence divergence over time [18]. Some fungi and oömycetes have small-RNA-producing loci that can be reasonably classified as *MIRNA* [19, 20], but little is known about the function of these genes.

⁴ A single report has indicated that *A. thaliana* AGO3 also binds 24-nt heterochromatic siRNAs and functions in DNA methylation [22]. *A. thaliana* AGO2 and AGO3 are products of a recent tandem duplication, and independent duplications of AGO2 occurred in other plant clades [23], so it remains to be seen if this function is conserved.

RNA.⁵ RNA-directed DNA methylation may also contribute to control of chromatin structure at loci other than transposons, with possible functions in development, intercellular communication, and large-scale chromatin packing [21].

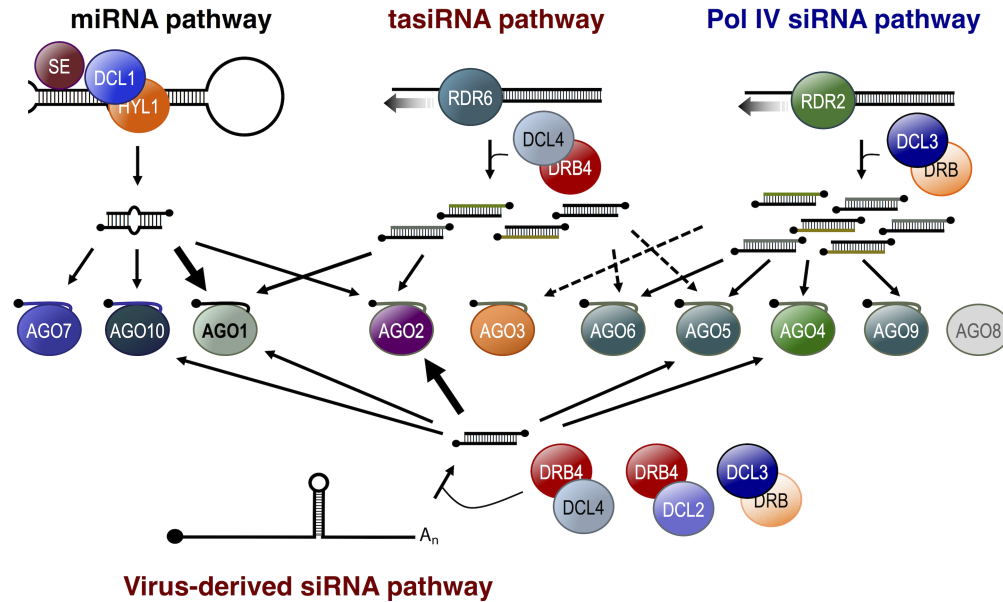


Figure 1.3: Specialized silencing pathways in *A. thaliana*. Specific DICER-like (DCL) proteins produce small RNAs, either from imperfect foldbacks (miRNA pathway) or from double-stranded RNA (siRNA pathways). The resulting small RNAs are loaded into different AGO proteins depending on their length, 5' nucleotide, foldback structure, and possibly other unknown factors; see also Figure 1.4. Viral substrates for small RNA production are thought to include dsRNA replication intermediates, self-complementary folded regions of genomic RNA, and/or dsRNA synthesized by host RNA-dependent RNA polymerases (RDRs) during the silencing amplification phase.

2. Antiviral defense: Direct cleavage of viral RNA by DCL2 and DCL4 makes a large contribution to defense [16, 31, 32]. The small RNA duplexes produced by such “dicing” are loaded into

⁵ A few early examples were particularly important in the recognition of silencing phenomena. One dramatic example was photobleached flowers in transgene-overexpressing petunia plants, due to silencing of endogenous pigment biosynthesis genes [28, 29]. Similar silencing of selectable markers and reporters provided important tools for genetic analysis of these pathways [13, 30].

AGO proteins which cleave virus RNA and can likely trigger amplification of silencing and immunization of distal tissues [33]. Nearly all plant viruses encode at least one protein that can suppress silencing responses at any number of steps [32], indicating that RNA silencing is a natural defense mechanism. Viral suppressors interfere with miRNA function [34, 35], providing one overarching explanation for the variety of developmental defects induced by virus infection, including total loss of polarity in “shoestring” leaves [36].

3. Endogenous gene regulation: Plant miRNA act by repressing endogenous transcripts, including many TF mRNAs. Almost all plant miRNA are produced by DCL1, together with a group of other biogenesis factors [37], and are loaded into AGO1. *ago1* and *dcl1* mutants were identified based on their pleiotropic developmental defects [38, 39], caused by perturbation of the regulatory circuits described below. Many plant miRNA repress their targets by endonucleolytic cleavage (“slicing”), but other outcomes are possible, as discussed below and in section 4.4.

1.1.2 AGO-small RNA complexes have specialized molecular functions

Plant *AGO* genes can be grouped into three main clades, named after their *A. thaliana* representatives (Figure 1.4). Similar AGO proteins generally have similar molecular functions: AGO4/6/9 proteins function in RNA-directed DNA methylation [40], while AGO1/10/5 and AGO2/3/7 clade AGOs function primarily in post-transcriptional regulation [23, 41]. Several different AGOs function in defense against RNA and DNA viruses [33].⁶

⁶ AGO2 has the most prominent antiviral role, for mechanistically unclear reasons—see Appendix A. AGO2 also functions in antibacterial defense signaling with the complementary “star” strand of miR393 (miR393*; reference 42) and in DNA double-strand break repair [43, 44].

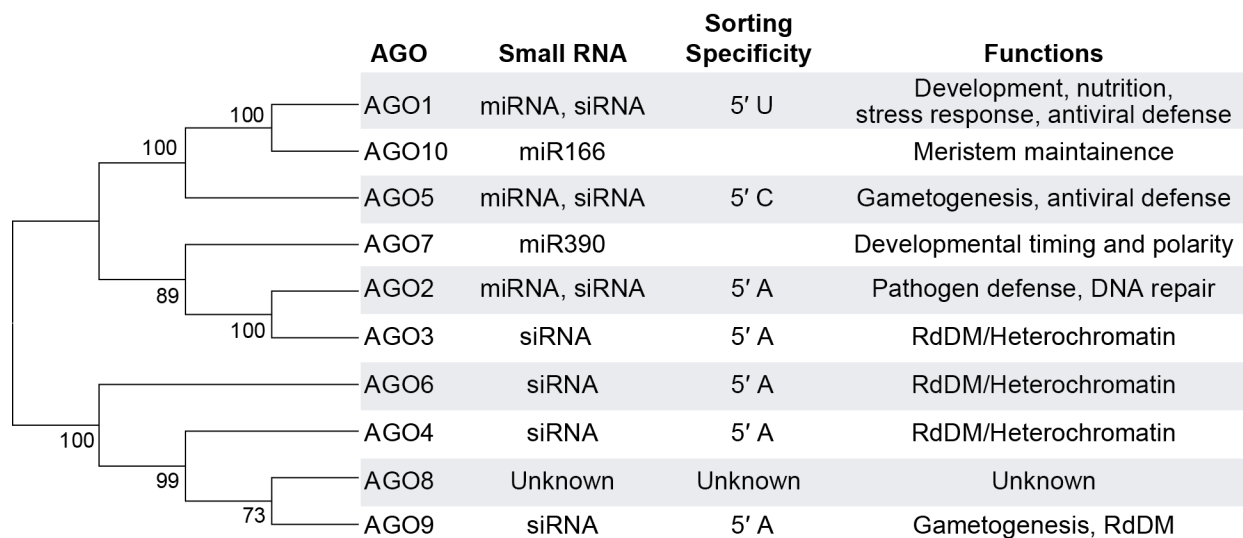


Figure 1.4: Cladogram illustrating the relation of the ten *AGO* genes of *A. thaliana*, with 5' nucleotide loading determinants and biological functions indicated. For AGO10 and AGO7, no 5' determinant is listed, because each binds essentially just one miRNA species. (Sections 1.2.1 and 1.2.2 describe the functions of AGO10-miR166 and AGO7-miR390 complexes, and Appendix A provides additional references and information on antiviral roles). Tree was inferred with the Molecular Evolutionary Genetic Analysis software tool; bootstrap values from 1,000 replicates are indicated. See reference 23 for a phylogenetic tree of *AGO* genes across the plant kingdom.

AGO-small RNA complex binding can trigger production of additional small RNAs from target transcripts. The resulting siRNAs are *trans*-acting (tasiRNA), but can also act in *cis* on other RNA from the same locus [45–47]. The most deeply conserved example is the action of the AGO7-miR390 complex at two sites on *TAS3* transcripts [48–50], as illustrated in Figure 1.5. AGO7 cleaves one site, setting the register for “phased” dicing at precisely spaced positions by DCL4 after dsRNA synthesis by RDR6. AGO7 may recruit the SGS3 protein⁷ to stabilize the initial cleavage product and/or recruit RDR6 (Figure 1.5). The function of the second noncleaved miR390 binding site is unclear, and other single-site targeting events can also trigger phased siRNA production, including by a small subset of AGO1-small RNA complexes [47]. *TAS3* tasiRNA are bound by AGO1 and function in timing and patterning, as described below.

⁷ The *SGS3* gene was identified (together with RDR6) based on *suppressor of gene silencing* mutants [51, 52].

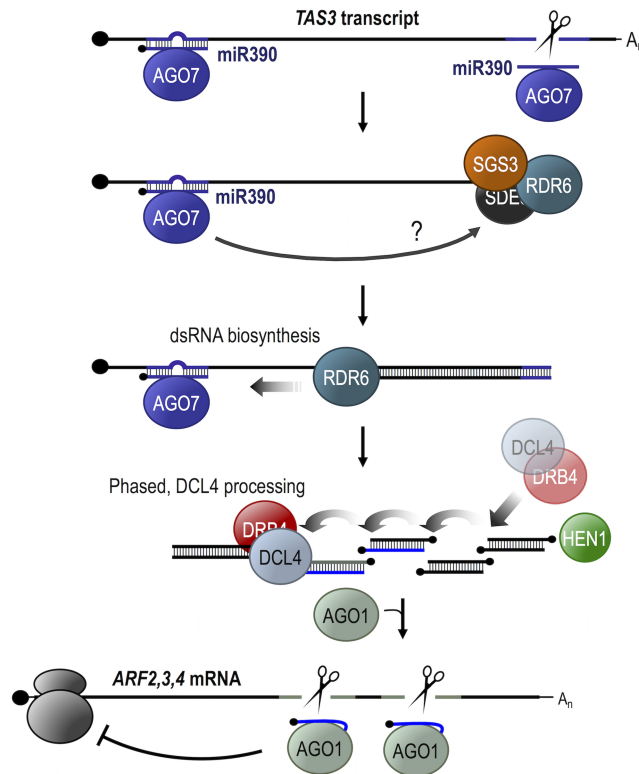


Figure 1.5: AGO7-triggered biogenesis of *TAS3* tasiRNAs via RNA-directed RNA synthesis and processive dicing. AGO7 specifically binds miR390 and acts at two sites on noncoding *TAS3* transcripts, as described in the text. *TAS3* tasiRNAs are loaded into AGO1 and cleave *ARF* target RNAs. Recent progress in understanding these molecular events is discussed further in section 4.4.⁸

1.1.3 Many plant miRNA cleave conserved TF mRNA targets

Nine *MIRNA* gene families are conserved in essentially all land plants [54–56]. The products of seven of these gene families regulate TFs that direct developmental programs (Table 1.1).⁹ The corresponding mature miRNA are produced at relatively high levels, and were therefore among the first small RNAs identified by initial cloning and sequencing efforts [59–61]. Corresponding

⁸ As noted in the acknowledgements above, this diagram and the last three were adapted from related figures made by others. See also summary figures in references 48, 50, and 53.

⁹ The other two conserved families (*MIR408* and *MIR395*) function in responses to abiotic stress [57, 58].

Table 1.1: Deeply conserved families of plant *MIRNA*, their conserved transcription factor targets, and the developmental processes they regulate. The upper section of the table lists seven of the nine *MIRNA* families likely present in the last common ancestor of all land plants. The lower section lists four additional families conserved in angiosperms. Mature miR390 does not directly target transcription factor mRNA, but is listed because it indirectly regulates *ARF* genes, as discussed in the text.

<i>MIRNA</i> gene family	TF target family	Developmental roles in <i>A. thaliana</i>
miR156/miR529	SBP/SPL	Leaf initiation and shape, flowering, lateral root growth
miR159/miR319	MYB and TCP	Leaf margin and flower patterning, senescence
miR160	ARF	Seed, leaf, root, and flower patterning
miR166	HD-ZIP	Embryo, leaf, and root patterning
miR171	GRAS	Shoot branching, light response
miR390	(ARF)	Leaf and flower/fruit shape, lateral root growth
miR396	GRF	Control of cell proliferation in leaves
miR164	NAC	Leaf and flower margin development
miR167	ARF	Root architecture and flower patterning
miR169	NF-YA	Root architecture, stress-induced flowering
miR172	AP2	Timing and patterning of flower development

loci and binding sites were quickly shown to be present in rice, suggesting an ancient origin [62], which has been confirmed by analysis of older plant lineages such as mosses [63–65]. All of these miRNA-target pairs were likely present in the last common ancestor of all land plants.

Ten additional *MIRNA* families are present in flowering plants (angiosperms), but not in mosses, spike mosses, or conifers. Four of these families have important roles in regulating development (Table 1.1, bottom half). Additionally, miR393 and miR394 target transcripts encoding auxin signaling F-box proteins (AFB1, AFB2, AFB3, and TRANSPORT INHIBITOR RESPONSE1) and the LEAF CURLING RESPONSIVENESS F-box protein [66–69]. miRNA produced from two other families conserved in flowering plants (*MIR168* and *MIR162*) target AGO1 and DCL1, providing homeostatic feedback to the core miRNA machinery [70–73].

In addition to these conserved *MIRNA* families, many *MIRNA* loci are limited to specific plant lineages, indicating frequent appearance and disappearance [74–78]. These *MIRNA* genes can drift rapidly, in part because they are not subject to reading frame constraints. New miRNA are often products of inverted duplications and retain complementarity to their parent loci [54, 79]. Other miRNA appear spontaneously [80], and may not have any function or targets at all [81].

Several *MIRNA* families are conserved across animal lineages [82, 83]. The first described miRNA is nematode-lineage-specific, but the second (*let-7*, also a regulator of timing) is broadly conserved [84, 85]. Preferential conservation analysis suggests the possibility that most transcripts in mammals are miRNA targets [86]. As noted above, no *MIRNA* gene families show clear conservation across kingdoms. Many of the miRNA-targeted TF families listed Table 1.1 are plant-kingdom-specific.¹⁰ Other notable groups such as homeodomain TFs are present in both plants and animals but have rather different developmental roles in each [87]. These differences provide an opportunity for comparative analysis.

1.2 AGO-small RNA-TF modules control plant form

As noted above, sets of miRNA targets fit the definition of *cis*-regulatory modules: an individual miRNA binding site acts *in cis* and is subject to different constraints than other regions of a given transcript. The small size of these binding sites can allow rapid change, though the the miRNA-TF pairs listed in Table 1.1 are highly conserved. Coordinate changes in sets of miRNA-regulated TFs trigger changes in the downstream targets of those TFs (Figure 1.1), and miRNA-TF target pairs can therefore be considered together as subcircuits [4]. Because of the hierarchy of control, changes

¹⁰ The set of plant-kingdom-specific TF families includes the SPL, TCP, and ARF groups—see below for acronym definitions and discussion.

in an upstream regulator can have large phenotypic consequences. A single module can function in multiple contexts, as exemplified by the network shown in Figure 1.6B, which specifies abaxial-adaxial polarity in leaves, flowers, and fruits. Modules are also deployed differently in different groups, as can be seen from the leaf-shape outputs of several miRNA-TF pairs: leaf architecture changes (heteroblasty) are manifested differently in different plant groups [88, 89]. Studying this process (often referred to as “vegetative phase change” or the “juvenile-to-adult transition”¹¹) has been a very productive approach to dissecting silencing functions; studies of abaxial-adaxial polarity have been equally productive.

1.2.1 The miR166-HD-ZIP module controls shoot identity and polarity

HD-ZIPs (homeodomain leucine-zipper TFs) have multiple functions in patterning, beginning in embryogenesis [94, 95]. Misexpression of *HD-ZIP* genes in embryos can yield homeotic transformation of the root pole into a second shoot [96]. Class III *HD-ZIP* transcripts have miR166 binding sites; mutations in these sites act dominantly, causing developmental defects, including upward-curling leaves [62, 97–99].¹² Severely affected leaves lose polarity, yielding needle-like or trumpet structures.

It appears that the main molecular function of AGO10 is to sequester miR166 away from AGO1, and thus increase *HD-ZIP* levels [100]. This process may involve AGO10-enhanced degradation of miR166 by specific nucleases [101]. At least one HD-ZIP protein directly binds the *AGO10* promoter and activates its transcription [102], creating a positive feedback loop that may explain

¹¹ Some authors discourage the use of these terms, because of potential confusion with the reproductive transition [90, 91]. In this document I simply refer to “early leaves” and “later leaves”, where possible—see Figure 1.7.

¹² *MIR165* genes also regulate *HD-ZIPs* in *A. thaliana* [62] because they derived from the *MIR166* family. The mature miRNA of these families differ at only a few nucleotide positions, so for simplicity I refer only to “miR166” throughout. The relation between the miR156/157 and miR170/171 families is similar [62].

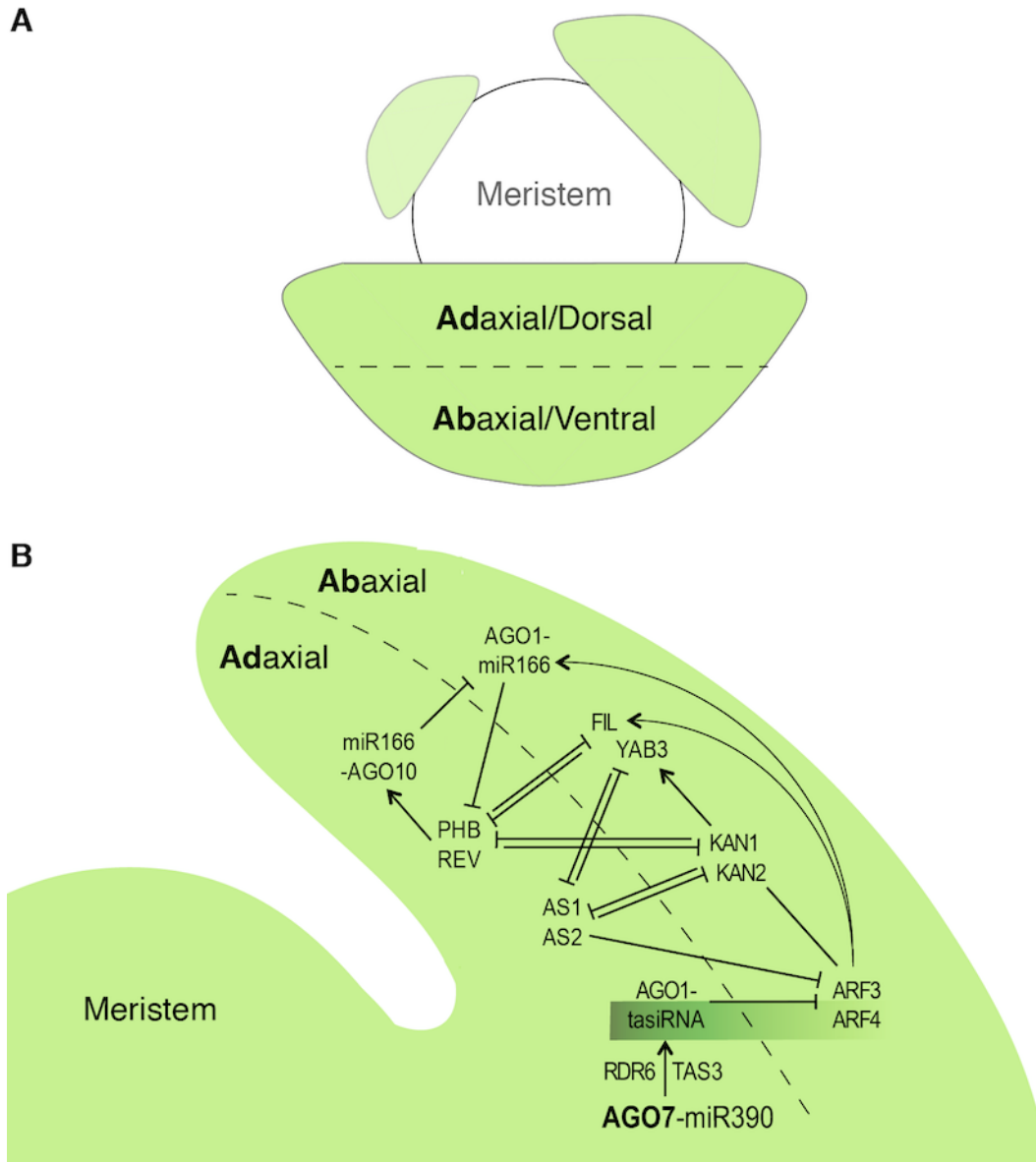


Figure 1.6: Schematics of leaf primordia illustrating abaxial-adaxial (dorsal-ventral) polarity.

A. Top-view schematic showing leaf primordia (transverse cross-sections) emerging in a spiral pattern from the flanks of the vegetative shoot apical meristem. Primordia are (pre)patterned relative the central-peripheral axis of the shoot: the side closer to the meristem center becomes the upward-facing surface of the leaf (specialized for photosynthesis), and the side opposite becomes the downward-facing surface (specialized for gas exchange).

B. Schematic of longitudinal cross-section through an expanding leaf primordium, with repressive genetic interactions controlling polarity indicated. Polarized expression of AGO10 and AGO7 is thought to modulate the activity of their highly specific miRNA binding partners, as described in the text. See references 92 and 93 for details on mutually antagonistic TF families.

the adaxial and vascular expression pattern of *AGO10* [103]. *AGO10* contributes to maintenance of stem cell identity in embryos [104], possibly by restricting the spread of miR166, which is known to move between root cells [105].¹³

Mobile signals have been suggested to similarly contribute to leaf axis establishment, based on experiments in which primordium-adjacent incisions disrupt leaf polarity, indicating that such patterning is specified relative to the center of the shoot apical meristem [92, 110, 111].¹⁴ Candidates for the relevant signaling molecules have included small RNAs (particularly some form of miR166 and/or *TAS3* tasiRNAs) and the hypothesized lipophilic ligand(s) of Class III HD-ZIPs, but there is no convincing evidence for either type of molecule [92]. The miR166-HD-ZIP circuit genetically interacts with the *TAS3* tasiRNA pathway, reflecting their shared involvement in control of polarity [112].

1.2.2 The *AGO7*-miR390-*TAS3*-ARF module controls polarity and timing

AGO7 was uncovered in screens for mutants prematurely showing adult leaf traits: *ago7* mutants have elongated downward-curved leaves with increased serration and early appearance of abaxial trichomes [113]. Additional screening uncovered several biogenesis factors [45, 114, 115] that define the molecular pathway shown in Figure 1.5. Downward-curling of leaves often indicates abaxialization, but *ago7* mutants do not have overt polarity defects [113]. However, both *ARF3* and *ARF4* specify polarity, and *ARF4* mRNA accumulates adaxially [116]. The adaxial localization

¹³ I discuss the genetic and molecular function of *AGO10* (including an apparent antiviral function) further in Appendix A. Members of the monocot-specific *AGO18* subclade also appear to compete with *AGO1* for small RNA binding: *AGO18* can promote virus resistance in rice [106, 107] and may also have specialized developmental functions [27, 108, 109].

¹⁴ An alternative hypothesis is that these incision and ablation studies disrupted primordium prepatterning. Genetic studies increasingly favor this interpretation and point to prepatterning via depletion of auxin, as discussed briefly in section 4.2.

of *AGO7* and other biogenesis factors suggests that *TAS3* tasiRNA biogenesis is spatially limited [117, 118], and that the gradient generated by movement of tasiRNAs may have a role in patterning, as depicted in Figure 1.6.¹⁵ Recent evidence indicates that *ARF3*, also known as *ETTIN*, directly responds to auxin, which modulates its physical interaction with other TFs [119, 120]. *ARF* repressors also control patterning in fruits [121–123] and lateral root emergence [124, 125]. *TAS3* tasiRNAs have been suggested to contribute to robustness of patterning in moss [126], and current evidence suggests that the *ARF* domain and the *TAS3* tasiRNA biogenesis pathway appeared coincident with the emergence of land plants [127]. *ARF3* and *ARF4* mRNA levels remain relatively constant as *A. thaliana* plants age [128], suggesting that their expression is not used directly for timekeeping.

1.2.3 The miR156-SPL module controls heteroblasty and flowering

SPB/SPL TFs are a plant-specific family of master regulators named after *SQUAMOSA* PROMOTER BINDING PROTEIN, a direct upstream regulator of the *SQUAMOSA* flower identity gene [129]. SPL levels go up as plants age [130], as depicted in Figure 1.7. This timekeeping mechanism seems to involve movement of carbohydrates to the shoot apical meristem as total photosynthetic leaf area increases: increasing sugar levels trigger gradual repression of *MIR156* transcription in the shoot apical meristem, leading to a gradual increase in SPL levels and thus leaf shape changes [131–133].

miR156 controls the transition from gametophytic to sporophytic growth in mosses [135, 136], consistent with an ancient timing role. The structure of the *TAS3* noncoding RNA differs in basal plants: both miR390 (as described above) and miR156 directly target *TAS3*, and the resulting

¹⁵ In section 4.1, I resummairize our understanding of this pathway and critically evaluate models for patterning by *TAS3* tasiRNAs based on new results presented here and elsewhere.

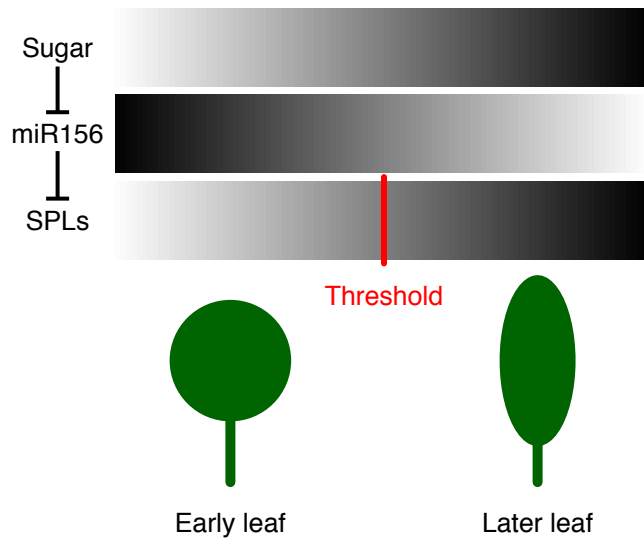


Figure 1.7: Schematic illustrating temporal steps in control of leaf shape by miR156 and SPL transcription factors. As leaf number increases, increased movement of photosynthate to the shoot apical meristem triggers repression of *MIR156* transcription, leading to an increase in *SPL* target mRNA. A hypothetical threshold for SPL activity controls transitions in leaf characteristics [128, 134]. In *A. thaliana* this transition is manifested as a shift from round early leaves to elongated downward-curved later leaves. See also Figure 4.1.

tasiRNAs target not only *ARFs* but also *APETLA2* (*AP2*) TF mRNAs, [127, 136]. In flowering plants *TAS3* tasiRNA target *ARF* mRNA exclusively and the appearance of the *MIR172* family provided another mechanism for regulating *AP2* family mRNAs [127]. SPLs activate *MIR172* genes and thus repress their *AP2* targets [137]. *AP2* TFs function primarily in flower development [138], but also affect maize leaf epidermal development [139–141]. SPL action interacts with gibberellic acid signaling and other pathways to control flowering [142–145], and multiple SPLs also inhibit lateral root production [146, 147].

Overexpression, miR156-resistant, and target mimic transgenic approaches all indicate that SPLs control leaf initiation [130, 148, 149]. Overexpression of *MIR156* in the maize *CORNGRASS1* mutant similarly has increased leaf initiation rate and prolonged juvenility [150]. In addition to leaf shape, SPLs control the two other main heteroblastic shoot traits of *A. thaliana*. They control

trichome production via transcriptional activation of MYB TFs¹⁶ and via physical interaction with other TFs, including miR171-target LOST MERISTEM proteins [153, 154]. SPLs were recently shown to control timing of leaf serration by directly sequestering the TFs described in the next section [155].

1.2.4 The miR319-TCP module controls organ margins and maturation

TCPs are a plant-specific family of basic-Helix-Loop-Helix TFs named after the maize *teosinte branched 1* locus [156], the snapdragon *CYCLOIDEA* flower symmetry gene [157], and two Proliferating Cell Factors [158, 159]. This third group (PCFs) was identified in rice based on their affinity for the promoters of proliferating cell nuclear antigen genes, reflecting the role of TCPs in arrest of cell division during patterning of leaves and flowers [160].

There are two main classes of TCPs and the second class includes many miR319 targets [161, 162]. The importance of this targeting was revealed by activation tagging of *MIR319*; downregulation of a set of functionally redundant TCPs produces dramatic wavy leaf phenotypes [163]. Activation of TCPs by abaxial factors is essential for proper control of margin growth to yield flat leaves [164, 165]. TCPs also control leaf senescence, in part by inducing jasmonic acid synthesis [166].

MIR319 and *MIR159* genes are very close in sequence, due to their common origin, and the combined family interacts with at least five other conserved miRNA-TF circuits. TCPs antagonize CUP-SHAPED COTYLEDON (CUC) proteins, both through induction of *MIR164* genes [167, 168], and through physically inhibiting their function as transcriptional activators [155]. As SPL protein levels increase, they compete for TCP interaction surfaces, titrating TCP molecules away

¹⁶ MYBs have helix-turn-helix DNA-binding domains and were discovered during studies of myeloblastoma oncogenes [151, 152]. The important miR159 target MYB33 is discussed below.

from CUC proteins, and thus allowing the CUC regulatory program that triggers appearance of leaf serrations to proceed [155]. An additional conserved miR159 target, MYB33, activates both *MIR156* and *SPL* genes [169], though this interaction plays only a minor role in leaf shape changes. MYB33 and TCPs activate miR167 to repress a set of ARF targets in flowers [170]. Finally, as part of their role in controlling cell proliferation, TCPs directly activate *MIR396* genes, and thus downregulate their Grown Response Factor targets [171]. The miR396-Grown Response Factor pair genetically interacts with *TAS3* pathway member RDR6 [172], as do the TCP target and polarity factor ASYMMETRIC LEAVES1 [168, 173–177].

A large number of TCPs contribute to function of the plant circadian clock [178], which interacts with light signaling many aspects of plant development, movement, and metabolism. This type of interaction was discovered based on a “promoter hiking” strategy [179] described below and in Chapter 2.

1.3 Gaps in our knowledge of gene regulation in plants

Plant development is worth studying for its fundamental importance and also for its practical utility. Master regulator TFs have long been considered attractive biotechnology targets, and the roles of TFs in crop domestication is well established, with notable examples from the *SPL* and *TCP* families [180–182]. Artificial miRNA and tasiRNA are useful for highly specific and predictable knock-down [183], and manipulation of a single miRNA or TF can have large beneficial effects [184, 185]. Mechanistic understanding can help us anticipate and quantify tradeoffs (between plasticity and robustness, growth and defense), bypass these tradeoffs in some cases [186], and optimize sensitivity in others, via both breeding and transgenic approaches. Plant architecture is a major determinant of yield, and therefore central to these efforts. Overexpression of *AGO7* has

been suggested as one useful method for modulating plant architecture, with promising results in rice and tomato [187, 188]. More broadly, altering core timing pathways identified in *A. thaliana* can be used to improve characteristics of diverse crops, including trees and other biofuel feedstocks [184, 189–193].

1.3.1 *AGO* genes as key control points for polarity and other processes

Thorough understanding of plant development requires understanding gene regulation at all levels. The research outlined above has uncovered suggestive examples indicating that plant AGO proteins act as regulatory hubs. Great progress has been made toward understanding AGO-small RNA complex formation and action, but further progress will likely depend on understanding transcriptional specialization of both *AGO* and small RNA genes [4]. Quantitative understanding of tissue- and stage-specific levels of *AGOs*, their guide small RNAs, their targets, and their presumed interacting proteins would enable holistic understanding of their developmental functions (Figure 1.1). We know very little overall about how different signals activate and repress AGOs; a near-total lack of information on the direct upstream regulators of *AGO* genes prevents us from reasoning about their function and evolution. For example, we can only speculate about the mechanistic basis for the tissue-specific defense roles of different AGOs, inferred from genetic analysis of Turnip mosaic virus infection [194].¹⁷ A related gap, particularly relevant to the highly specialized functions of AGO10 and AGO7, is that we do not know the order of events in polarity establishment, nor how many of the repressive interactions involved (shown in Figure 1.6) are caused by direct TF-promoter interactions. These antagonistic interactions have been difficult to dissect in mutants, due to all-or-nothing transformations, pleiotropic effects, and genetic redundancy. New strategies are being applied to assay direct action and thus sidestep issues such as redundancy.

¹⁷ See Appendix A for discussion of likely tissue-specific AGO defense roles.

1.3.2 New resources for promoter analysis

Large collections of *A. thaliana* TF clones have recently been applied to identification of direct TF-promoter interactions in two main ways. First, long regulatory sequences can be screened for TF binding directly using yeast 1-hybrid (Y1H) systems. Use of defined clone collections has dramatically increased the ease and throughput of this method beyond cDNA library screening; this improved approach was initially applied to define “gene-centered” regulatory networks in *C. elegans* [195]. One notable effort in *A. thaliana* successfully used Y1H to circumvent genetic redundancy in the robust core circadian transcription network [179]. Indeed, Y1H systems capitalize on the overlapping functions between TFs: detection of multiple hits from the same family can provide confidence, because such TF paralogs often have similar DNA-binding specificities.

A second way in which TF collections are enabling systematic investigations is by facilitating large-scale determination of *in vitro* DNA binding specificities with protein-binding microarrays [196, 197], “DNA affinity purification sequencing” [198], and a variety of other methods [199]. The resulting specificity models, now available for essentially all families of plant TFs, allow genome-wide prediction of direct binding. Y1H and specificity-model-based prediction tools are complementary to each other, and also to other methods such as chromatin immunoprecipitation. These methods are particularly useful when combined with information on function, co-expression, and/or DNA accessibility [200].

1.3.3 New tools for automated measurement of growth and development

Another bottleneck to studies of development is simple measurement of growth at the organ and whole-organism level. Many have suggested that this gap can be addressed using improvements

in optics, automation, and computation [201], paralleling the recent renaissance in microscopy methods. Most discussion of automated imaging methods have focused on their potential for accelerating crop breeding, but studies of model systems can benefit as well. Relative ease of imaging has been essential for studies of fly embryos, whole worms, and other systems. The *A. thaliana* rosette is larger than these small model animals, but has other features that make it well-suited to imaging and interesting to computer vision experts [202]. New automated imaging methods remain inaccessible to many laboratories because of their high cost, an issue I discuss at length below.

1.3.4 Overview of contributions

This dissertation has two main messages, which are evident in the framework outlined above (particularly references 50, 113, 117, and 202) and were reinforced by the results I obtained.

- *AGO7* is a key regulator of plant growth that integrates temporal and spatial signals.
- Low-cost timelapse imaging methodology is rapidly improving and likely to become a standard tool, which will be particularly beneficial for *A. thaliana* geneticists.

Our main goal, described in the next chapter, was to identify direct upstream regulators of the three *AGO* genes central to post-transcriptional control of development (*AGO7*, *AGO10*, and *AGO1*), and then characterize related functional linkages. Based on initial results (subsection 2.3.1), efforts focused on *AGO7*. Surprisingly, screening did not identify factors known to control *polarity*, but rather suggested a role for miRNA-targeted TFs involved in different aspects of *timing*, i.e. leaf heteroblasty. Functional analysis defined requirements for two short proximal promoter regions (Figure 2.17). In the course of this functional characterization (described in section 2.3.3), I

refined a simple low-cost method for timelapse photography of rosette growth. A technical report describing this methodology (chapter 3) should be useful to many labs, given the widespread interest in this subject. A series of appendices (B to D) further substantiate the argument that imaging of vegetative growth is becoming a powerful tool. Collectively, this work advances our knowledge of leaf development in an important model system.

Chapter 2

Direct regulation of *ARGONAUTE7* by miRNA-targeted transcription factors

This chapter will be revised and submitted for peer review, with contributors tentatively listed as follows:

J. Steen Hoyer^{1,2}, Jose L. Pruneda-Paz³, Ghislain Breton^{3,4}, Mariah A. Hassert¹, Emily E. Holcomb¹, Halley Fowler¹, Kaylyn Bauer¹, Jacob Mreen¹, Steve A. Kay^{3,5}, James C. Carrington¹.

1. Donald Danforth Plant Science Center, St. Louis MO, USA.
2. Computational and systems biology program, Washington University in St. Louis, MO, USA.
3. Division of Biological Sciences and Center for Chronobiology, University of California San Diego, CA, USA.
4. Department of Integrative Biology and Pharmacology, McGovern Medical School, Houston, TX, USA.
5. Department of Neurology, University of Southern California, Los Angeles, CA, USA.

2.1 Abstract

ARGONAUTES are the central effector proteins of RNA silencing which bind target transcripts in a small RNA-guided manner. *Arabidopsis thaliana* has ten *ARGONAUTE* (AGO) genes, with specialized roles in RNA-directed DNA methylation, post-transcriptional gene silencing, and antiviral defense. To better understand specialization among *AGO* genes at the level of transcriptional regulation we tested a library of 1541 transcription factors for binding to the promoters of *AGO1*, *AGO10*, and *AGO7* using yeast 1-hybrid assays. A ranked list of candidate DNA-binding TFs revealed binding of the *AGO7* promoter by a number of proteins in two families: the miR156-regulated SPL family and the miR319-regulated TCP family, both of which have roles in developmental timing and leaf morphology. Possible functions for SPL and TCP binding are unclear: we showed that these binding sites are not required for the polar expression pattern of *AGO7*, nor for the function of *AGO7* in leaf shape. Normal *AGO7* transcription levels and function appear to depend instead on an adjacent 124-bp region. Progress in understanding the structure of this promoter may aid efforts to understand how the conserved AGO7-triggered *TAS3* pathway functions in timing and polarity.

2.2 Introduction

Small RNAs regulate developmental timing and morphogenesis in a wide range of eukaryotes. Heterochronic (abnormal timing) mutants of the model nematode *Caenorhabditis elegans* led to the discovery of the first microRNA (miRNA)-target pair [9, 10]. Similar screens for *A. thaliana* heterochronic mutants led to elucidation of a specialized pathway in which *trans*-acting small interfering (tasi)RNA are produced from noncoding *TAS3* transcripts [45, 113, 114, 203]. Genetic analysis of leaf morphology has also led to the discovery of several other aspects of RNA silencing,

including the cloning of the first *ARGONAUTE* (*AGO*) gene [38]. AGO proteins bind small RNAs and effect small-RNA-guided regulatory changes. Several families of *MIRNA* genes are conserved in all land plants [54], and miRNA from the majority of these families repress TFs controlling developmental programs, suggesting that AGO-miRNA-TF circuits became embedded in the core regulatory networks for the plant body plant early in land plant evolution [5].

The *A. thaliana* genome contains ten *AGO* genes, which function in development, stress resistance, and defense against viruses and transposons [23]. AGO7 and AGO10 are highly specialized: each has limited adaxial and vascular expression [103, 117] and a single main binding partner: miR390 and miR166, respectively [50, 100]. AGO7 triggers production of phased siRNAs from *TAS3* noncoding transcripts [48–50, 204]. Effects on ARF3, ARF4, and possibly ARF2 are the main downstream output of the AGO7/*TAS3*/SGS3/RDR6/DCL4 pathway [128, 175, 205, 206]. AGO7 action is thought to limit production of *TAS3* tasiRNAs such that tasiRNA movement creates a graded accumulation pattern in developing leaf primordia [117, 118; see Figure 1.6]. This gradient contributes to patterning of *ARF* target mRNA, establishing either an opposing gradient or a sharp boundary, which may contribute to robust maintenance of polarity [207]. The *TAS3* pathway has important roles in leaf development in all plants examined thus far, including moss [126], maize [112, 208, 209], tomato [36], lotus [210] and alfalfa [211].

Understanding the functions of miRNA such as miR390 and miR166 will require information on the signals controlling tissue-specificity of their AGO partners. Our objective in this work was to identify upstream regulators of *AGO* genes and link them to existing genetic knowledge. We capitalized on new yeast-based tools that provide a fast way to identify upstream regulators. We identified unexpected connections to two other conserved miRNA-TF circuits that control leaf morphogenesis and defined two other functional regions of the *AGO7* promoter.

2.3 Results

2.3.1 Multiple SPLs and TCPs bind the *AGO7* promoter

We sought to identify TFs controlling the expression of the three main *AGO* genes involved in post-transcriptional control of development (*AGO1*, *AGO10*, and *AGO7*) using high-throughput yeast 1-hybrid assays. Our automated strategy, described previously [178, 179], uses a large collection of arrayed *A. thaliana* TFs (details below) and also short promoter bait sequences, for high resolution and sensitivity. We considered four fragments for each promoter, with ~50 bp of overlap between fragments, to ensure that fragment-edge binding sites were assayed. For *AGO7* these fragments spanned a 1934 bp region (Figure 2.1A). Transgenes driven by the collective sequences represented by these fragments are sufficient to complement corresponding *ago* mutants [50, 104, 212], suggesting that they contain the most important upstream regulatory elements. Promoter fragments were screened against a TF-activation domain fusion library in 384-well format with one prey TF per well [178], using β -galactosidase reporter activity from fusion to promoterless *uidA* coding sequence as a quantitative readout (Figure 2.1).

A total of 1497 TFs were tested for *AGO* promoter binding (Table 2.1). This collection consists mainly of sequence-specific TFs, but also includes transcriptional co-factors and empty vector control wells [178]. Each TF was tested against each promoter fragment a single time. We ranked TF candidates based on normalizing promoter-fragment-driven β -gal activity by the median value for each plate (as illustrated in Figure 2.1B), to account for systematic differences between plates. We separately plotted signal distributions across all twelve screens (Figures 2.2, 2.3, and 2.4) to assess which TFs “hits” act as nonspecific activators in this system, as described below.

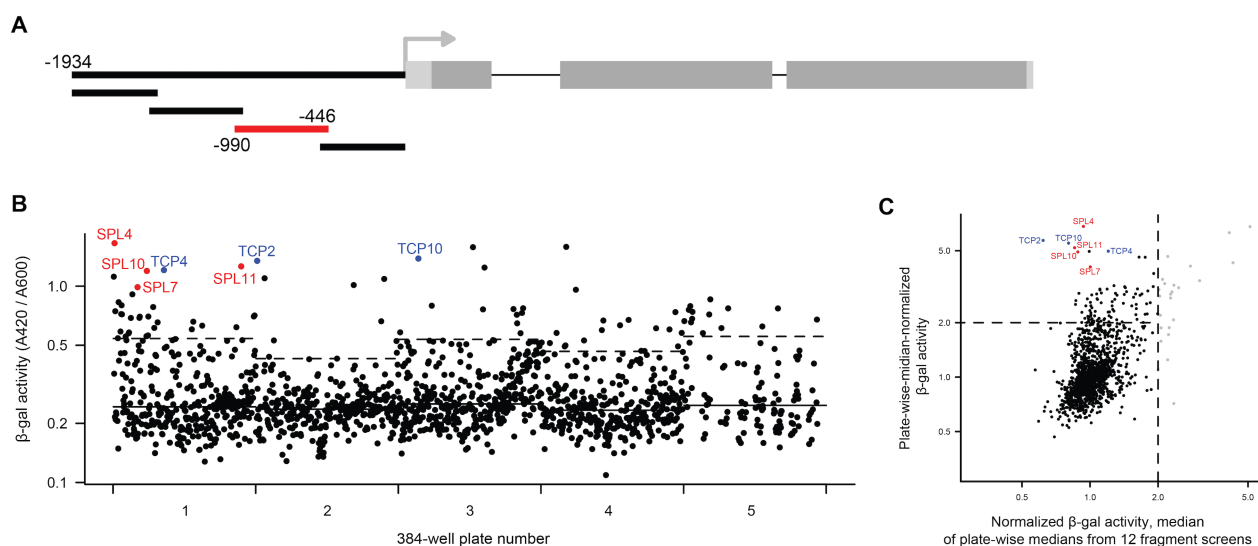


Figure 2.1: SPL and TCP TFs bind the *AGO7* promoter in yeast.

A. Schematic of *AGO7* promoter illustrating four fragments screened with Y1H assays. Subsequent panels show results for the fragment indicated in red, which spans the region from 990 bp to 446 bp upstream of the transcription start site.

B. Scatterplot of β -gal activities for each prey TF constructs screened. Wells are shown in row-first order for each of the five plates. Median activity for each plate is indicated with solid lines. Dashed lines indicate a cutoff of 6 median absolute deviations above the median for each plate. Hits from SPL and TCP families are highlighted.

C. Diagnostic plot incorporating data from 12 screens. Y-dimension reflects the same values as panel B, normalized by plate median. X-dimension results from taking the median of plate-wise-median activities from all twelve *AGO* promoter fragment screens. Vertical dashed line demarcates TFs for which median reporter activity is two-fold higher than the median for their plate (nonspecific activators, light gray).

Table 2.1: Transcription factors included in Y1H prey library for *AGO* promoter screens: counts by family. Thirty-five families are represented by ten or more TFs. thirteen families are represented by two TFs, and ninety-eight TFs are single representatives or unassigned. Cf. reference 178.

Family	Count	Family	Count
MYB	118	NF-YB	9
bHLH	98	SRS	9
MADS	87	C2C2-CO-like	8
C2H2	85	ARID	7
HB	71	AS2/LOB	7
AP2-EREBP	64	BES1/BZR	7
ERF	58	COL (C2C2)	7
NAC	58	ZF-HD	7
bZIP	55	CPP	6
MYB-related	55	FHA	6
WRKY	53	Alfin	5
C3H	41	ARR-B	5
GRAS	31	EIL	5
AUX-IAA	27	JUMONJI	5
LOB	21	LIM	5
TCP	20	NF-YC	5
G2-like	19	RWP-RK	5
ABI3VP1	16	SNF2	5
C2C2-Dof	16	TLP	5
HSF	16	E2F-DP	4
RING-REGIA	16	Histone 2A	4
CCAAT	15	NF-YA	4
Trihelix	15	PLATZ	4
SET	14	TUB	4
C2C2-GATA	13	YABBY (C2C2)	4
DOF (C2C2)	12	AP2 (Single domain)	3
GARP-G2-like	12	ARR-A	3
GATA (C2C2)	12	BBX	3
GeBP	12	BT	3
PHD	12	DBP	3
SPL	12	GIF	3
TIFY	12	MBF1	3
ARF	10	PBF-2-like	3
HMG	10	Pseudo ARR-B	3
REM (B3)	10	Sigma70-like	3
		TAF	3
		Other (one or two TFs)	122
		Total	1497

Of the TFs families assayed, only two were represented by multiple hits 6 absolute deviations or more above the median for their plate (Figure 2.1B). The first group, Teosinte Branched/Cycloidea/PCF family factors (TCPs),¹⁸ had previously been suggested to directly regulate *AGO7* [168]. The three TCP hits identified are miR319 targets [163] and redundantly control leaf margin development and senescence [166]. The second group, SQUAMOSA-PROMOTER-BINDING PROTEIN-LIKE (SPL) factors,¹⁹ are master regulators of heteroblasty in *A. thaliana* and other plants [88], the same context in which *AGO7* was discovered [113].

We examined the distribution of reporter activity for other promoter fragments screened, confirming that these SPL and TCPs specifically hit the second proximal region of the *AGO7* promoter. Plate-wise median β -gal activities for the SPL and TCP hits were close to the median (across all twelve screens) for their plate (Figure 2.1C), indicating that they do not fall in the group of TFs that are nonspecific reporter gene activators.

We further tested a group of SPL and TCP factors with a second Y1H system, based on a secreted luciferase reporter with an improved dynamic range [213]; repeated testing reduces statistical false positives and use of alternative reporters can reveal reporter-gene-specific technical false positives [214]. This secondary screening confirmed that multiple SPL and TCP TFs bind the second proximal *AGO7* promoter fragment tested, despite considerable experimental noise (Figure 2.5). Some TFs yielded a small degree of activation relative to two different empty vector controls; it is not clear whether these small differences reflect lack of binding (i.e. nonspecific binding only) or indicate binding that is weak but specific.

¹⁸ Recall that the TCP family was described in subsection 1.2.4.

¹⁹ Recall that the SPL family was described in subsection 1.2.3.

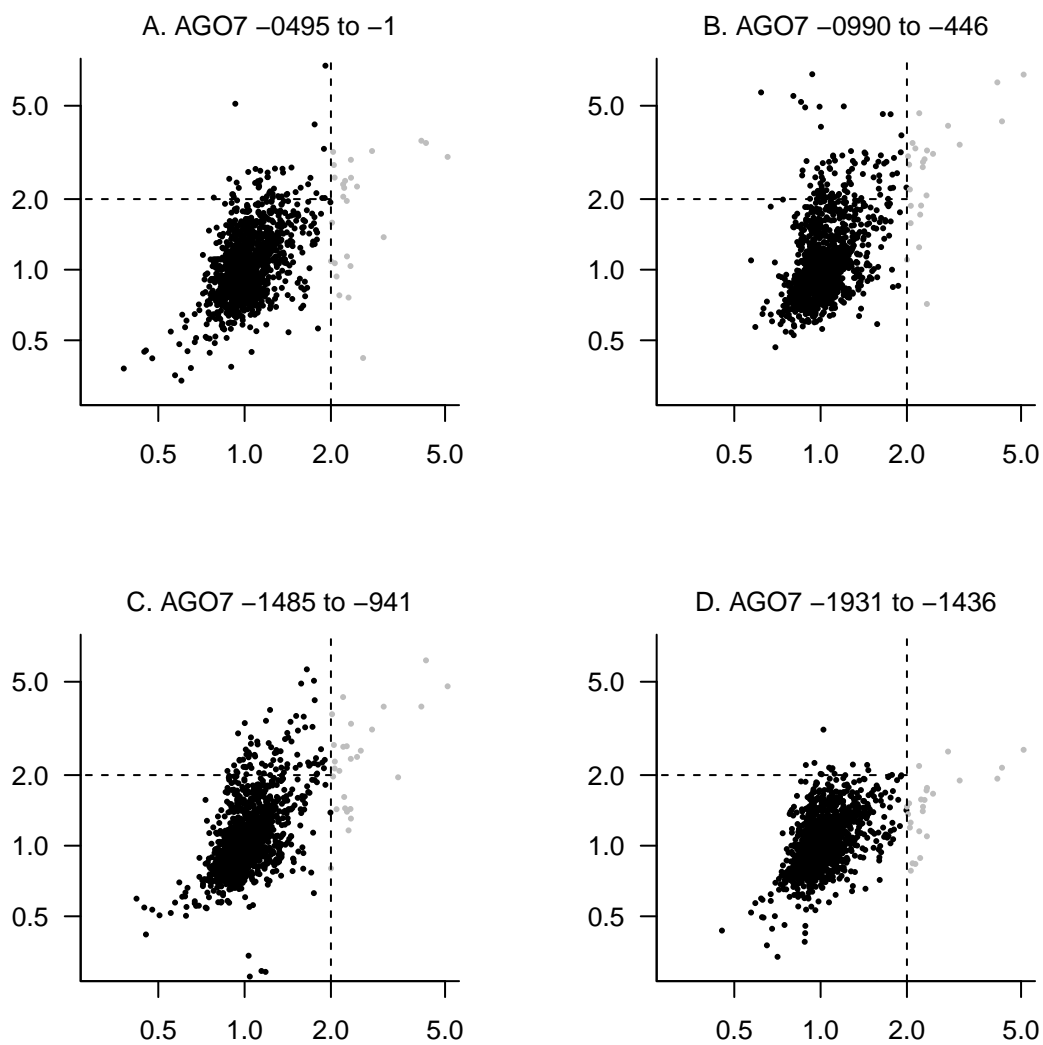


Figure 2.2: Scatterplots of β -gal activities with likely nonspecific activations indicated for *AGO7* promoter fragment screens. Panel B is equivalent to Figure 2.1C.

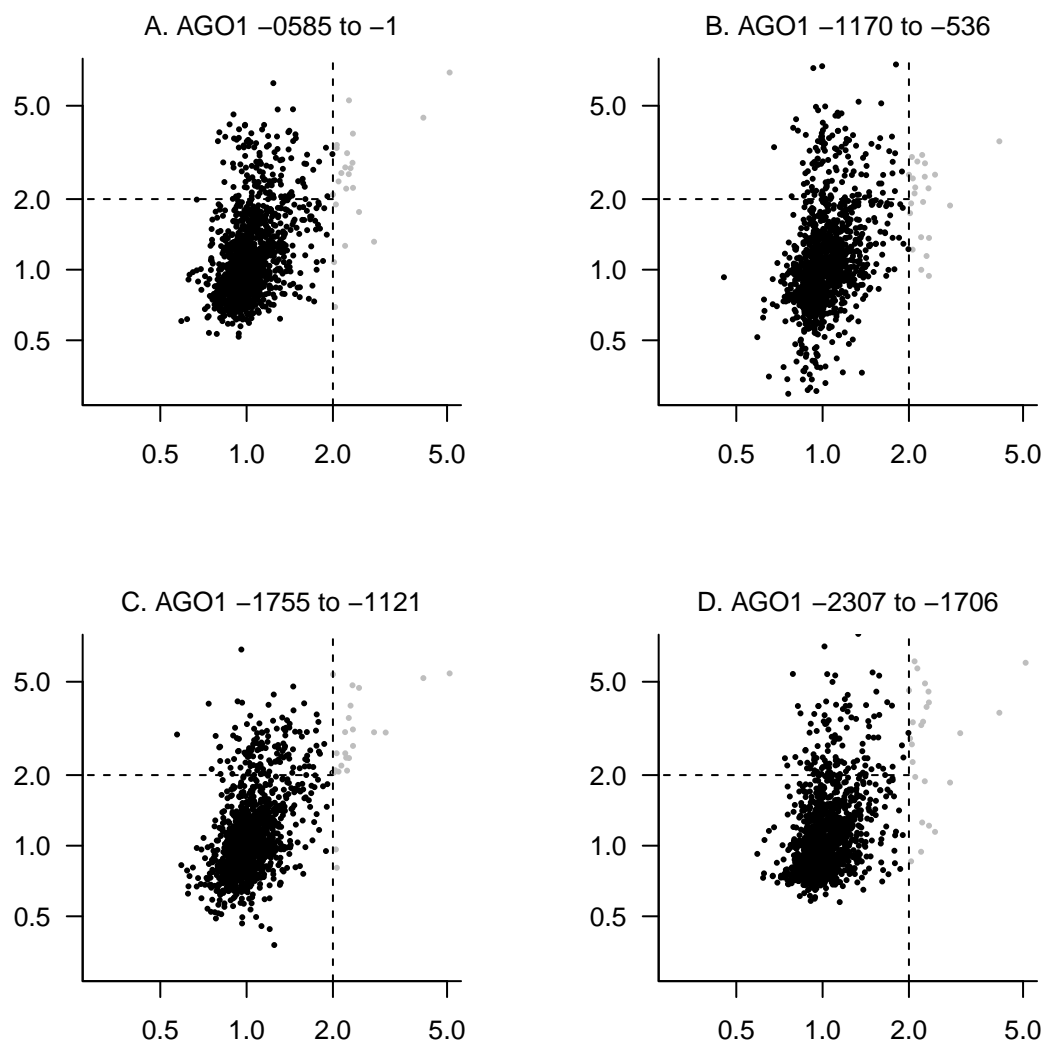


Figure 2.3: Scatterplots of β -gal activities with likely nonspecific activations indicated as in Figure 2.1C for *AGO1* promoter fragment screens.

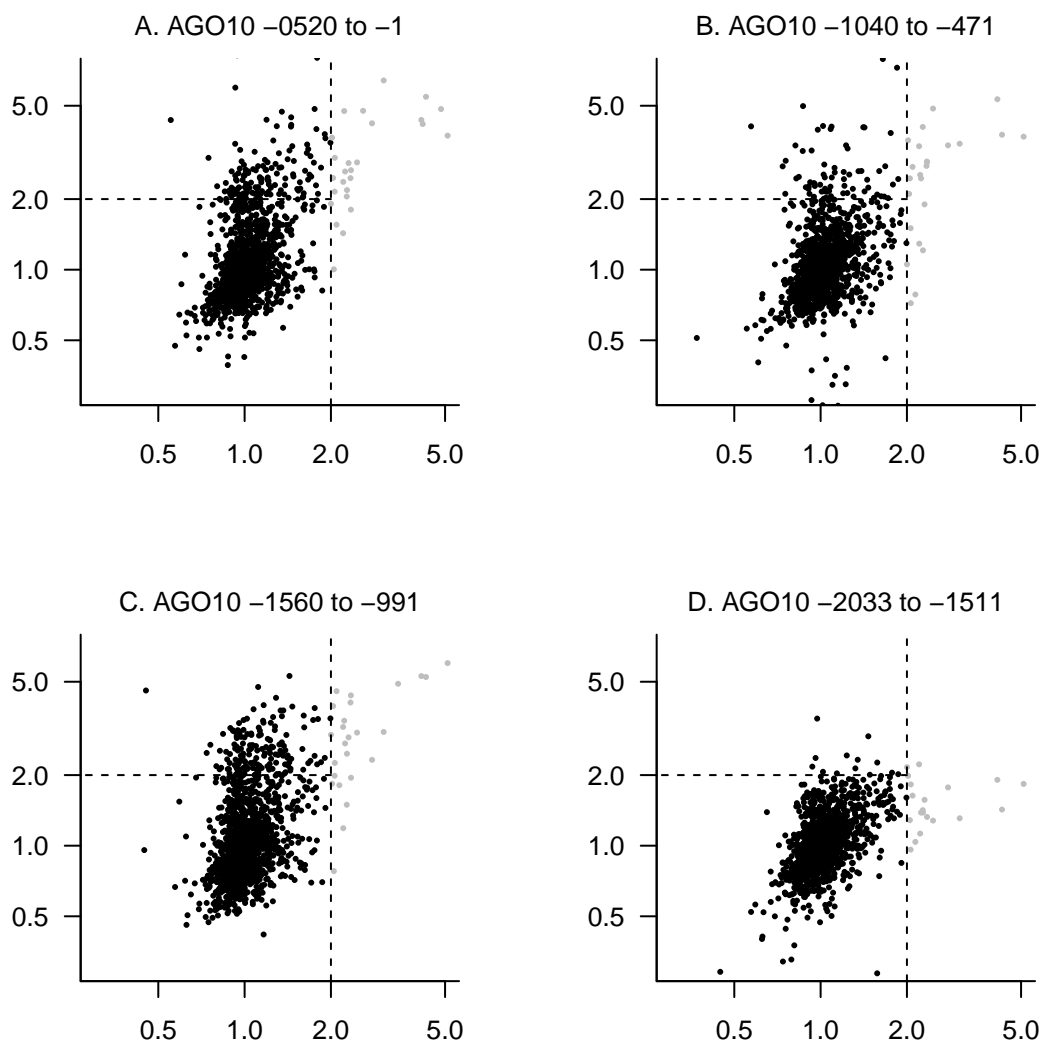


Figure 2.4: Scatterplots of β -gal activities with likely nonspecific activations indicated as in Figure 2.1C for *AGO1* promoter fragment screens.

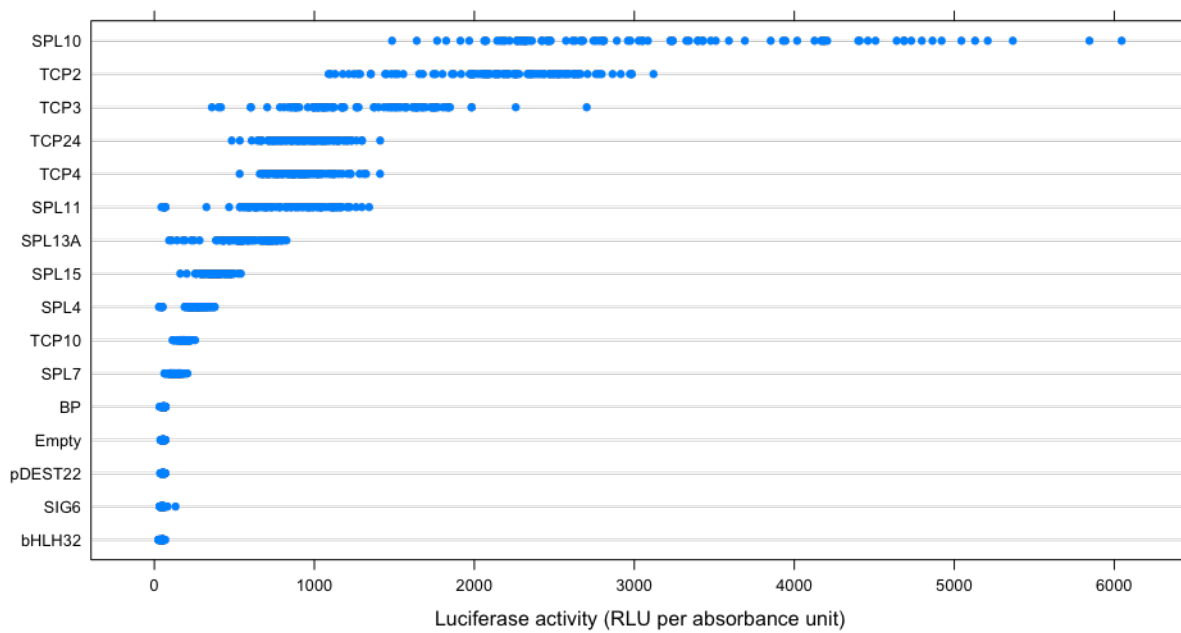


Figure 2.5: Targeted Y1H assay using *Gaussia* luciferase reporter, quantified in terms of relative luminescence units per absorbance unit at 600 nm. TFs were tested against the *AGO7* -990/-446 region and are displayed in order by mean reporter activity.

We assessed possible SPLs and TCPs binding sites using DNA-binding specificity models determined based on *in vitro* sequence affinity with protein-binding microarrays [196]. These position-weight matrices (PWM, downloaded from the CisBP database) match consensus binding sequences previously determined with *in vitro* selection for SPLs [215] and TCP4 [166]. An example sequence logo for one of these models, for SPL11, is shown in Figure 2.6A. Because the Y1H bait of interest extends to position -990 (Figure 2.1A), we considered the 1 kb region adjacent to the annotated *AGO7* transcription start site. For SPL11, the highest-scoring positions (on both strands) were centered on the only two ‘GTAC’ motifs (SPL core binding sites) in that region, at -500/-497 and -486/-483 (Figure 2.6, panels B and C).

We tested the significance of these core ‘GTAC’ sequences using the luciferase reporter gene in yeast. Truncated Y1H bait sequences (-531/-446 and -750/-476) containing core binding sequences yielded activation of the reporter when tested against SPL11, but not with the corresponding empty prey vector (Figure 2.6D). By contrast, activation was not observed for a 3’-truncated bait lacking ‘GTAC’ sites (-750/-501), nor for modified -531/-446 bait sequences with one or both 4-mers deleted or scrambled (Figure 2.6D). Deletion of an unrelated 6-bp region reduced reporter activation (compared to empty vector) but not to the same extent. These results are consistent with direct SPL binding, possibly with some degree of cooperativity, at one or both ‘GTAC’ sites in the yeast system.

We similarly scanned the promoter sequence with empirically determined PWM for five of eight *CINCINNATA*-like TCPs, a set that includes four of the five miR319 targets in *A. thaliana* [160, 163]. The highest scoring positions for four TCPs were centered on a ‘TGGTCC’ motif at -459/-454 (Figure 2.7, panels E to I). This 6-mer was the most highly enriched sequence in the promoters of a set of experimentally defined TCP targets [166], and is present in the “most preferred” sequences for TCP3, TCP4, and TCP5 PWMs. A second ‘TGGTCC’ site at -428/-423 was among the four

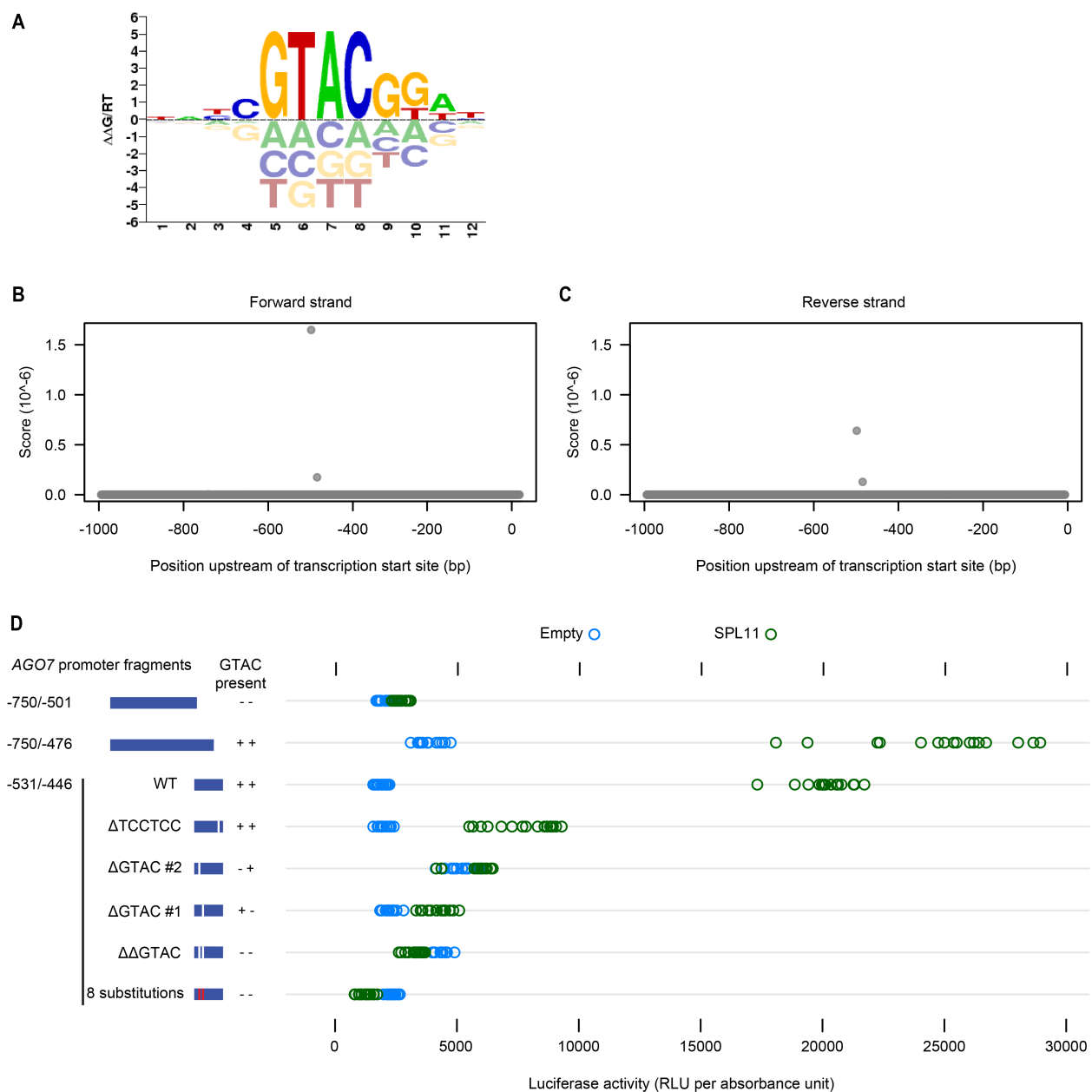


Figure 2.6: Identification of SPL11 binding sites.

A. Sequence logo for SPL11 PWM, as downloaded from CisBP. Individual position weights can be interpreted as binding specificity contributions (changes in free energy, arbitrary units).

B and C. Scores for SPL11 PWM at each position of the 1 kb region upstream of the annotated *AGO7* transcription start site.

D. Reporter activity (relative luminescence units normalized by A600) for SPL11 and pDEST22 (empty vector) tested in yeast against *AGO7* promoter baits including several derivatives of the -531/-446 region. Modifications included one or two 4-bp deletions, 8 substitutions (TCCG/AAGG), and an unrelated 6-bp deletion; see Table 2.4, below.

highest-scoring sequences for all five TCPs considered (Figure 2.7), but was absent the -990/-446 region that yielded TCP hits in the initial Y1H screen. High-scoring positions for the TCP2 PWM included a related ‘GGGACC’ sequence at -764/-770 followed by the -459/-454 ‘TGGTCC’ motif (Figure 2.7, panels A and F). The second highest scoring position for TCP24 was centered on a nearby ‘GTTCCC’ sequence (Figure 2.7J).

We tested requirements for candidate TCP binding sites with the luciferase Y1H system. Truncated bait sequences (-750/-501 and -750/-476) lacking all four sites described above did not drive reporter activation (relative to the empty prey vector control) when tested with TCP2 (Figure 2.8). The -990/-446 region used in the initial screen yielded reporter induction, as did a 5'-truncated 86 bp bait region (-531/-446) containing the higher-scoring ‘TGGTCC’ motif (Figure 2.8). The same truncated bait sequence with the ‘TGGTCC’ 6-mer deleted did not yield reporter activation (Figure 2.8). We conclude that the -459/-454 ‘TGGTCC’ is a high-affinity TCP binding site that functions in the yeast system and possibly *in planta*.

2.3.2 SPL binding sites are not required for polar *AGO7* transcription

To test the possibility that SPL and/or TCP binding sites contribute to polar *AGO7* transcription, we fused a series of truncated versions of the *AGO7* promoter to *GUS* for comparison to previously described transcriptional reporter lines [50, 117]. Consistent with previous results [117], the 1934 bp region upstream of the annotated *AGO7* transcription start site yielded clear adaxial signal in transverse sections of leaf primordia (Figure 2.9A). A 482 bp version of the promoter yielded the same pattern in almost all plants tested (Figure 2.9B), indicating that SPL core binding sites (-500/-496 and -486/-483) are not required for this pattern. Assaying a single transgenic family for a 422 bp promoter construct provided tentative evidence that core TCP binding sites are not

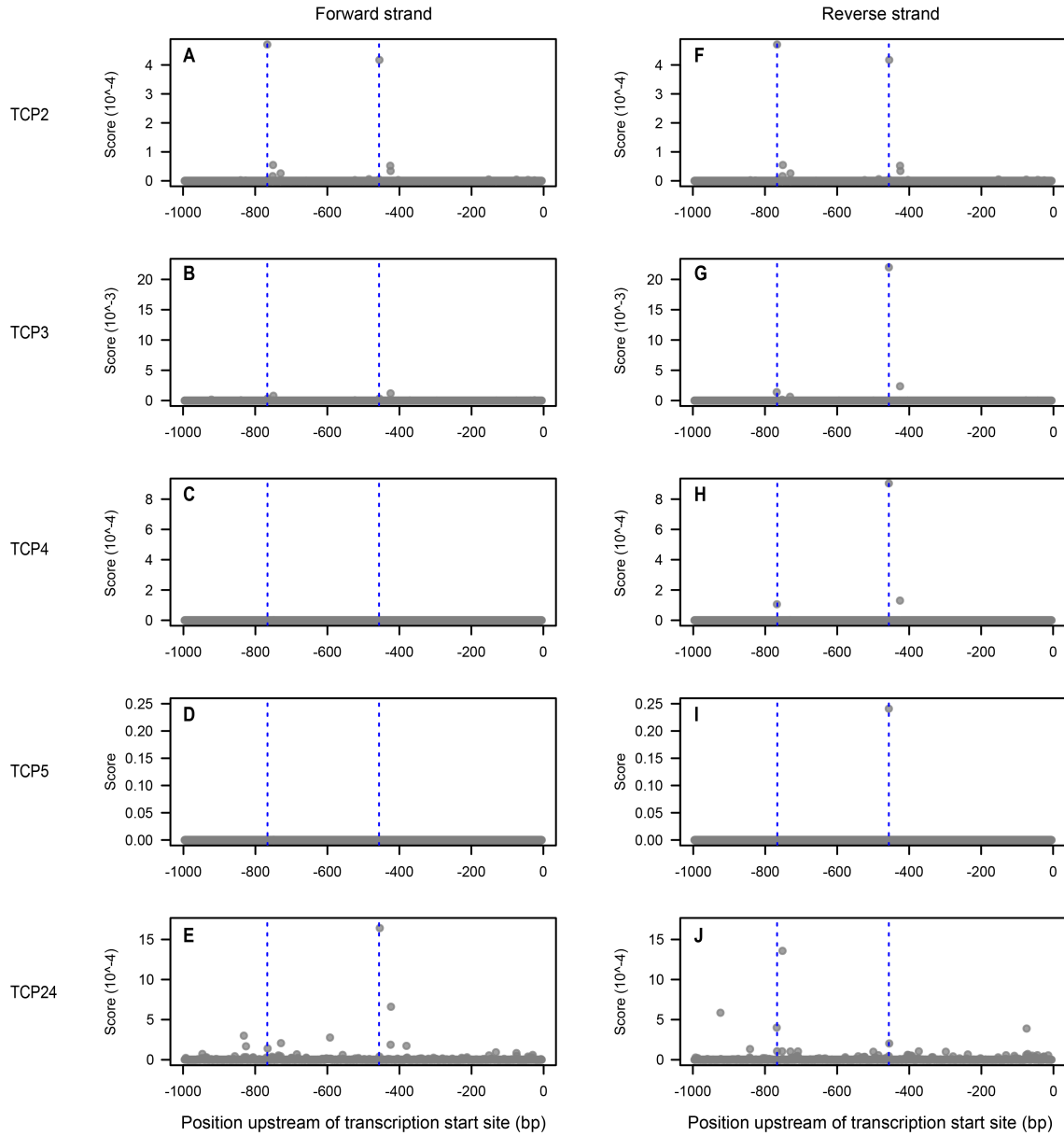


Figure 2.7: Identification of TCP binding sites.

A to J. PWM scores at each position of the 1 kb region upstream of the annotated *AGO7* transcription start site for the TCPs indicated. Dashed blue lines indicate the two highest scoring positions for TCP2: a 'GGGACC' sequence at -764/-770 and a 'TGGTCC' sequence at -459/-454. Panels A and F are identical, because the CisBP model for TCP2 is perfectly symmetrical.

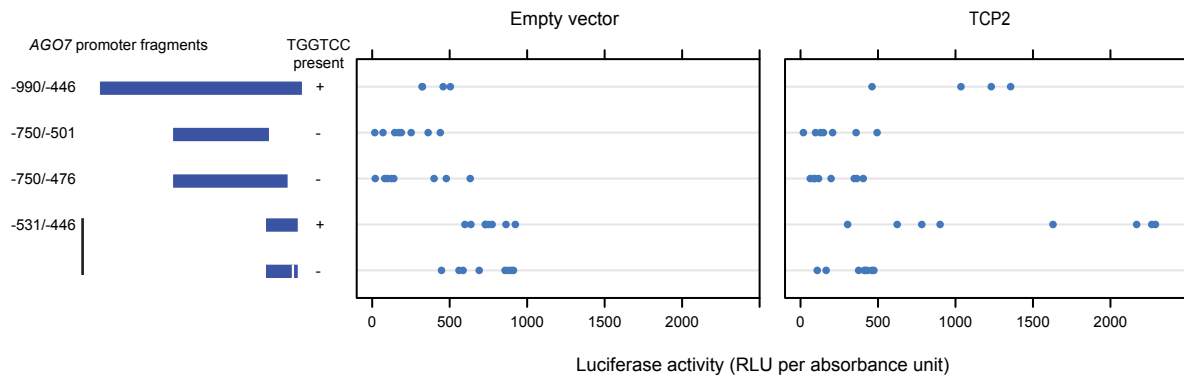


Figure 2.8: Testing of TCP binding sites in yeast. Reporter activity for TCP2 and pDEST22 (empty vector) tested against Y1H bait from initial screens (-990/-446), two truncated versions lacking candidate TCP binding sites, and the -531/-446 region, with candidate TCP binding site ('TGGTCC') deleted or intact.

required for this pattern either (Figure 2.9A). By contrast, the TSS-proximal 298 bp region rarely yielded visible blue reporter signal (Figure 2.9C). Weak adaxial signal was visible for a small proportion of plants (Figure 2.10B), including 2 of 7 plants for one of two transgenic families for the experiment illustrated. Surprisingly, one of two 150 bp construct transgenic families yielded stronger polar signal (Figure 2.10C). It is possible this pattern may have been enhanced by elements adjacent to T-DNA insertion (position effects) or some other technical factor caused higher staining intensity. Further experiments with additional independent transformants would be required to distinguish these possibilities. A promoterless 5' UTR construct appeared to yield faint blue signal (Figure 2.10D), but promoterless-GUS transformants did not yield visible blue signal (Figure 2.9D) in any of our experiments.

Staining of whole seedlings yielded qualitatively similar results (Figure 2.11), notably in the clear difference in staining intensity for 495 bp vs. 298 bp promoter:GUS transgenes. An additional high-signal control (*35S::GUS*) was included in the experiment shown. Signal from this transgene was surprisingly uneven and difficult to detect in all histological experiments done with moderate

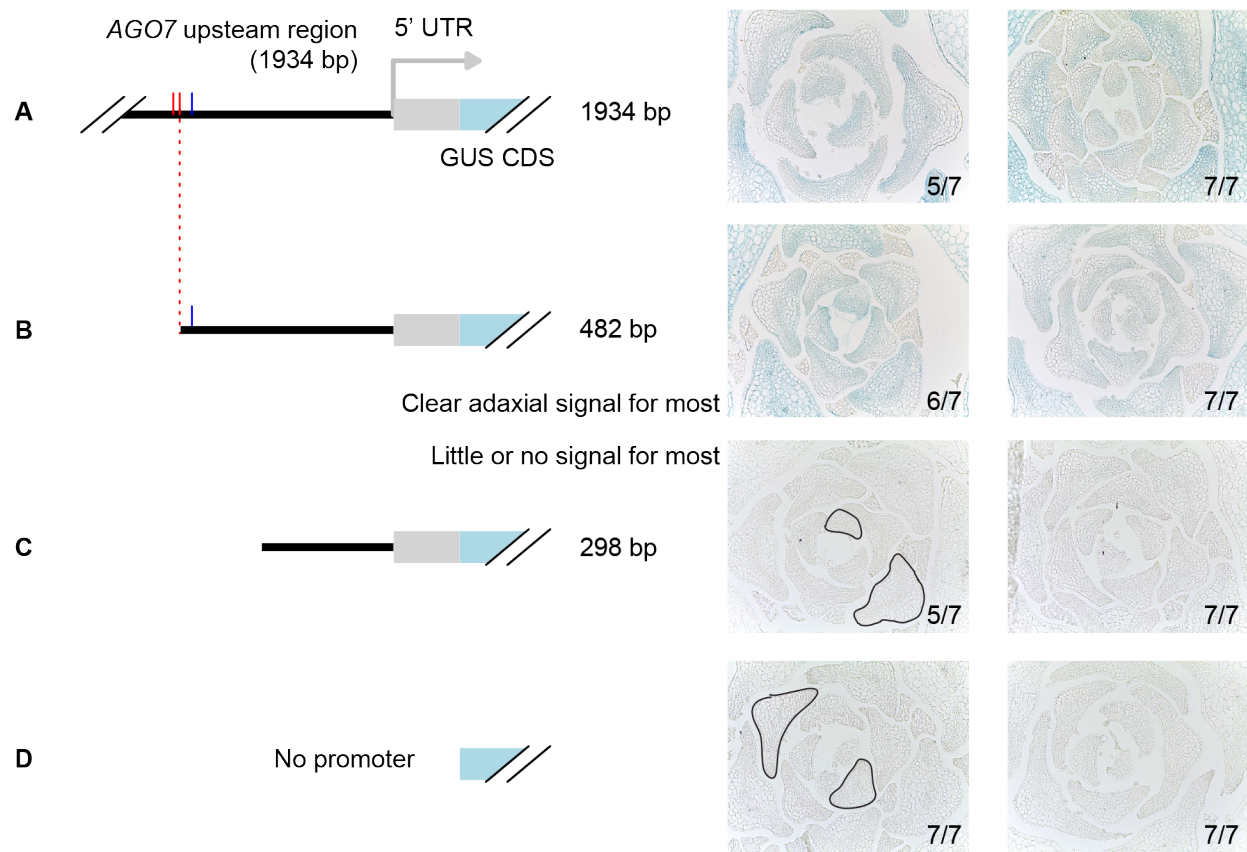


Figure 2.9: Histological analysis of GUS reporter gene activity driven by truncated *AGO7* promoter constructs. Core SPL binding sites are indicated in red; the 482 bp promoter construct illustrated in panel B ends immediately adjacent to the second site. Blue tick mark indicates TCP binding motif at -459/-454. For each construct, results are shown for two independent transgenic families (groups of T₃ siblings, each descended from a different transformant; each group was stained in a separate scintillation vial). The predominant class for each family is illustrated with a representative transverse section through young leaf primordia, and the number of plants in the predominant class is indicated as a fraction. Two primordia are outlined for both of the transgenes for which signal is not visible (panels C and D). Two and one plants yielded a weaker and/or less strongly adaxial pattern than shown here for 1934 bp and 482 bp promoter constructs, respectively. The less-frequent pattern for 298 bp promoter constructs (panel C) is shown in Figure 2.10B. Between three and nineteen independent lines were tested for all of the constructs shown here, with broadly similar results across multiple experiments.

The schematic shown in Figure 1.6A may help with interpretation of the transverse perspective shown.

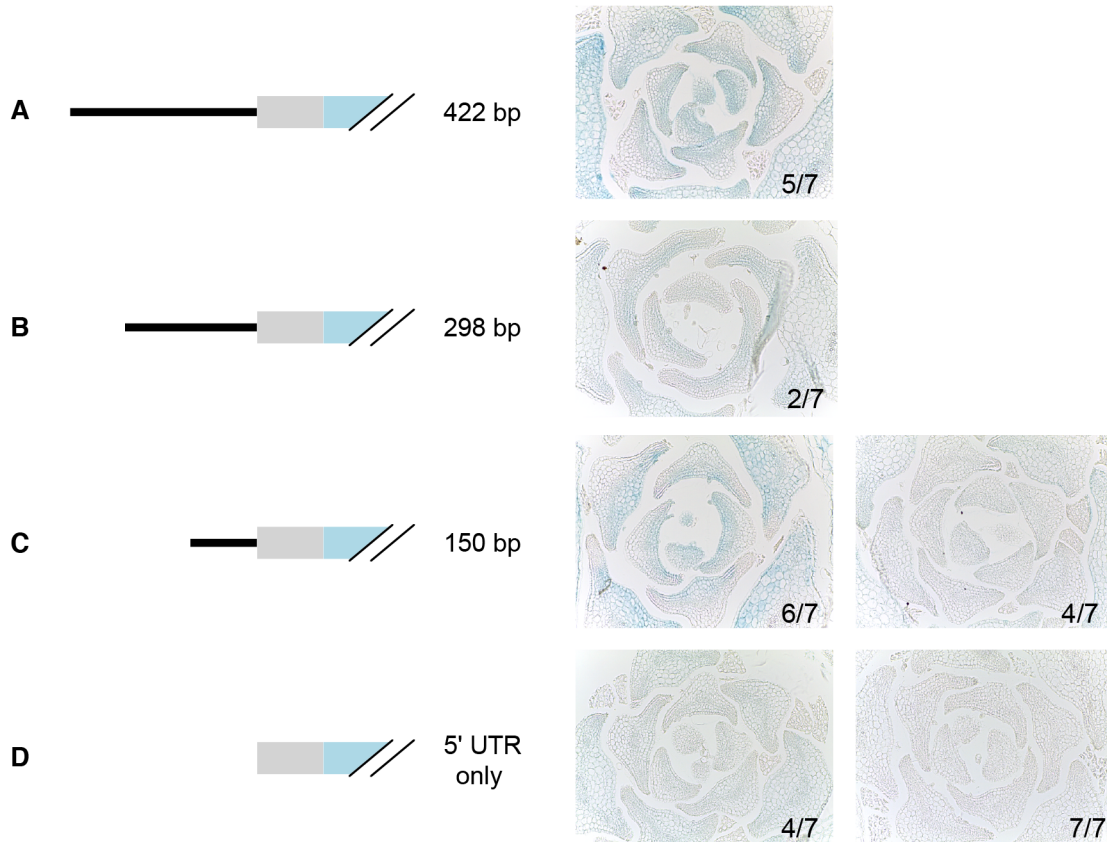


Figure 2.10: Transverse sections through leaf primordia, showing GUS reporter gene activity driven by shorter truncated *AGO7* promoter constructs. Sections are from the same experiment as Figure 2.9. Core SPL and TCP binding sites are absent from all promoter fragments illustrated. A. Predominant staining pattern (5/7 plants) observed for a single transgenic family with 422 bp truncated promoter. B. Lower-frequency pattern (faint adaxial signal) observed for 2/7 plants for one of the transgenic families shown in Figure 2.9C (left side). C and D. Predominant staining patterns for two transgenic families each for transgenes including 150 bp and 0 bp of sequence upstream of the *AGO7* 5' UTR.

or high stringency staining (not shown). This weak staining in histological sections from multiple lines is difficult to explain and may have involved transgene silencing.

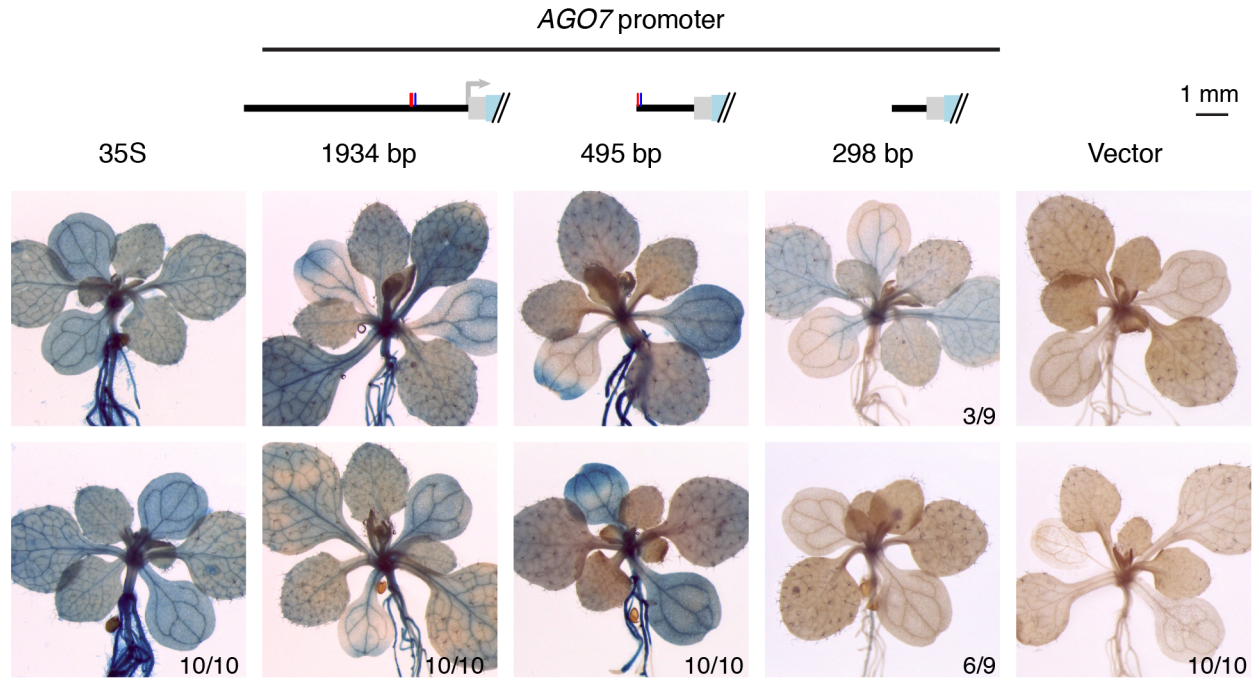


Figure 2.11: GUS reporter gene activity driven by truncated *AGO7* promoter in whole seedlings 12 days post-stratification. Two representative plants are shown for each transgene. Patterns were consistent among siblings, except for the 298 bp promoter construct, for which weak blue signal was visible for 3 of 9 plants but not the others.

Overall these results raise the possibility that *cis* elements in a short proximal promoter region or 5' UTR can confer adaxial polarity to *AGO7* transcription. The -482/-299 region, however, is a larger determinant of *AGO7* transcription level, as discussed further below.

2.3.3 SPL and TCP binding sites are not strictly required for *AGO7* function

We similarly tested *cis* requirements for transgenic complementation of *ago7* mutants. We inserted a series of truncated versions of the *AGO7* promoter upstream of the *AGO7* coding sequence

(including an N-terminal 3x-hemagglutinin (HA) tag). Previous results [50] indicated that the 1934 bp promoter version of this transgene is functional for complementation of transformed *ago7* mutants. For the experiment illustrated in Figure 2.13, blinded classification of downward leaf curling assigned 100% of empty-vector-transformed reference genotype plants (*ago7* mutant and wild-type Col-0, $n = 21$ and 20 plants, respectively) to the expected phenotype class. Groups of mutant plants transformed with 3xHA-AGO7 constructs were predominantly assigned to one or the other class: primary transformants for 422 bp to 1934 bp promoter constructs were mostly scored as complemented, whereas most transformants for 298 bp and 0 bp promoter construct displayed the downward-curved-leaf mutant defect (Figure 2.12).

We extended this result by quantifying leaf shape for a smaller number of transformants, by dissecting, scanning, and measuring leaves in order [216]. For the reference genotypes, leaf blade length-to-width ratios were higher for wild-type relative to mutant plants, due to increased curling and/or elongation (Figure 2.13, panels A and H). Promoterless and 298 bp promoter construct transformants were not distinguishable from empty vector mutant controls (Figure 2.13, panels F and G). Longer promoter constructs shifted blade length-to-width ratios down towards wild-type levels (Figure 2.13, panels B to E), which we interpret as partial complementation, consistent with the rosette-level results in Figure 2.12. Independently measuring these leaf dimensions at one position (true leaf 6)²⁰ with calipers yielded similar results (Figure 2.14).

Results were similar for a related metric that quantifies leaf elongation, the ratio of leaf blade length to petiole length (Figures 2.15 and 2.14). The difference between wild-type and mutant background control plants was smaller for this metric (Figure 2.15, panels A and H), as was the difference, if any, between means for the 1934 bp promoter construct lines and wild-type empty vector control lines

²⁰ True leaf 6 was selected for measurement for consistency with other studies, including references 206 and 49.

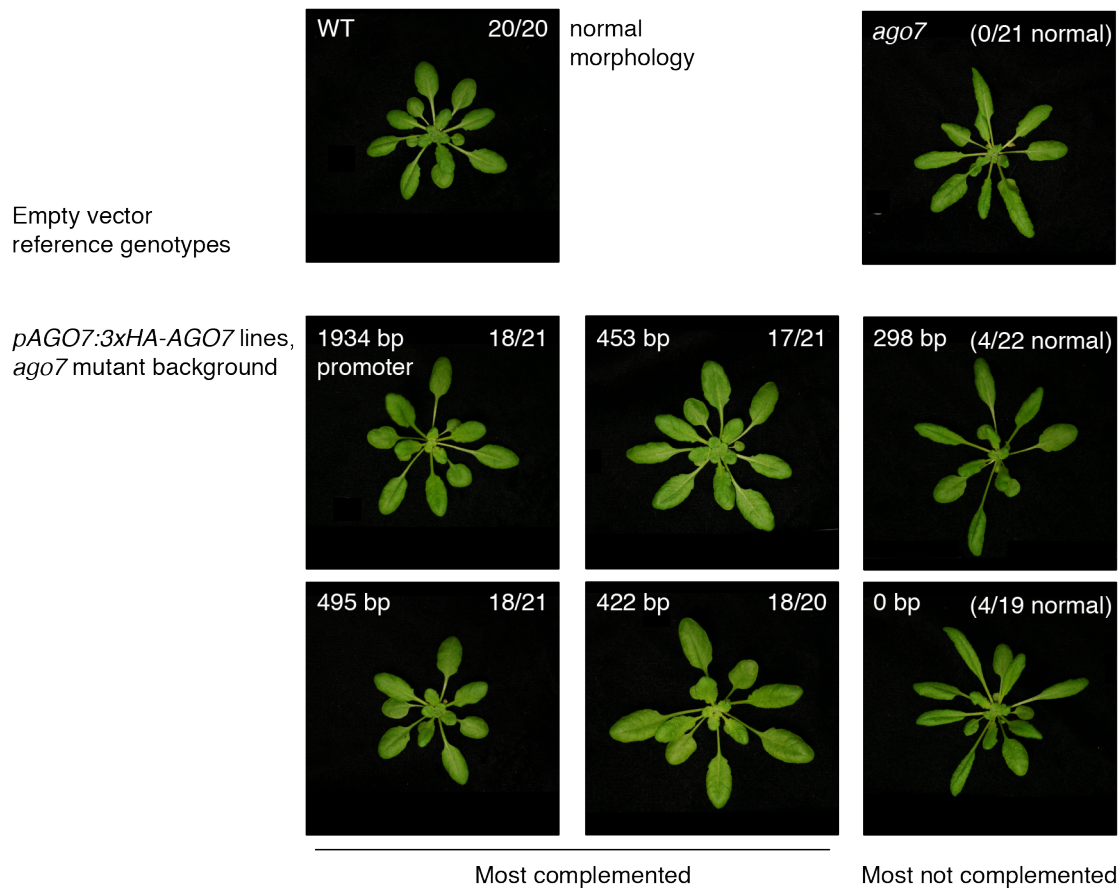


Figure 2.12: Complementation of *ago7* mutant leaf shape phenotype (top right) with 3xHA-AGO7 transgenes driven by truncated versions of the *AGO7* promoter. One representative (major class) primary transformant is shown for each genotype. Upper-left corner labels for middle and bottom rows indicate the length of upstream *AGO7* regulatory sequence used to drive the 3xHA-AGO7 coding sequence in each construct. Upper-right corner numbers indicate the fraction of plants blindly assigned to the normal morphology category.

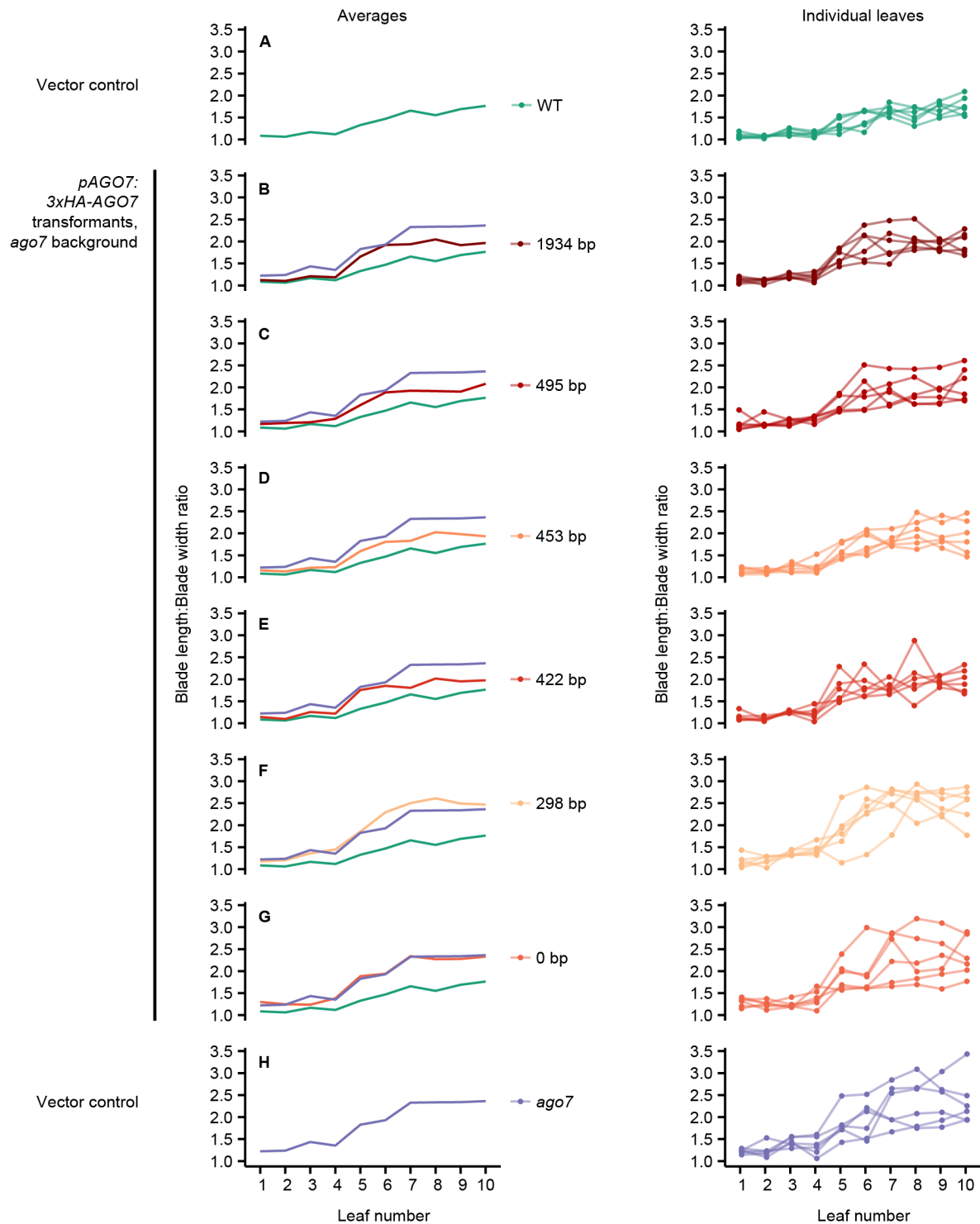


Figure 2.13: Transgenic complementation of *ago7* leaf shape defects, quantified based on leaf blade length to width ratio for true leaves 1 to 10. (Caption continues on next page.)

Figure 2.13: (continued) Values for each individual plant are connected with lines on the right-hand graphs, and the average of these values is plotted on the left. Averages for empty vector control genotypes (panels A and H) are repeated in each left-hand panel to facilitate comparison.

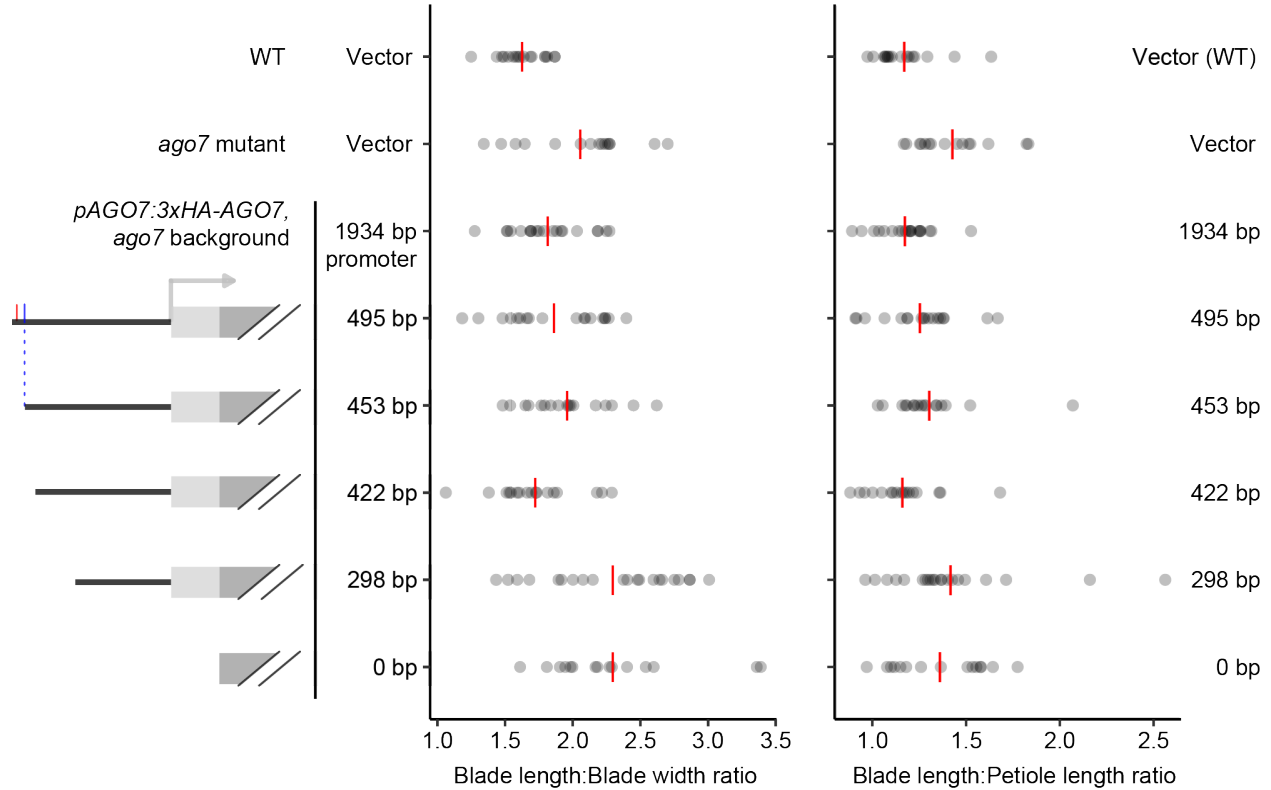


Figure 2.14: Complementation of *ago7* leaf shape defects, quantified based on leaf blade length-to-width ratio (left) and leaf-blade-length to petiole-length ratio (right) measured for true leaf 6 with calipers on days 28 to 30 days post-stratification. Each datapoint shows the ratio for a distinct primary transformant. Red lines indicate the mean for each genotype.

(Figure 2.15B). Means were longer at most leaf positions (i.e. closer to wild-type) for intermediate-length promoter constructs (Figure 2.15, panels C, D, and E) than for short promoter constructs (Figure 2.15, panels F and G). Exceptions at one position (true leaf 10) were caused by recently emerged leaf “outliers”, the petioles of which were very short and thus disproportionately affected by technical variation (Figure 2.15C).

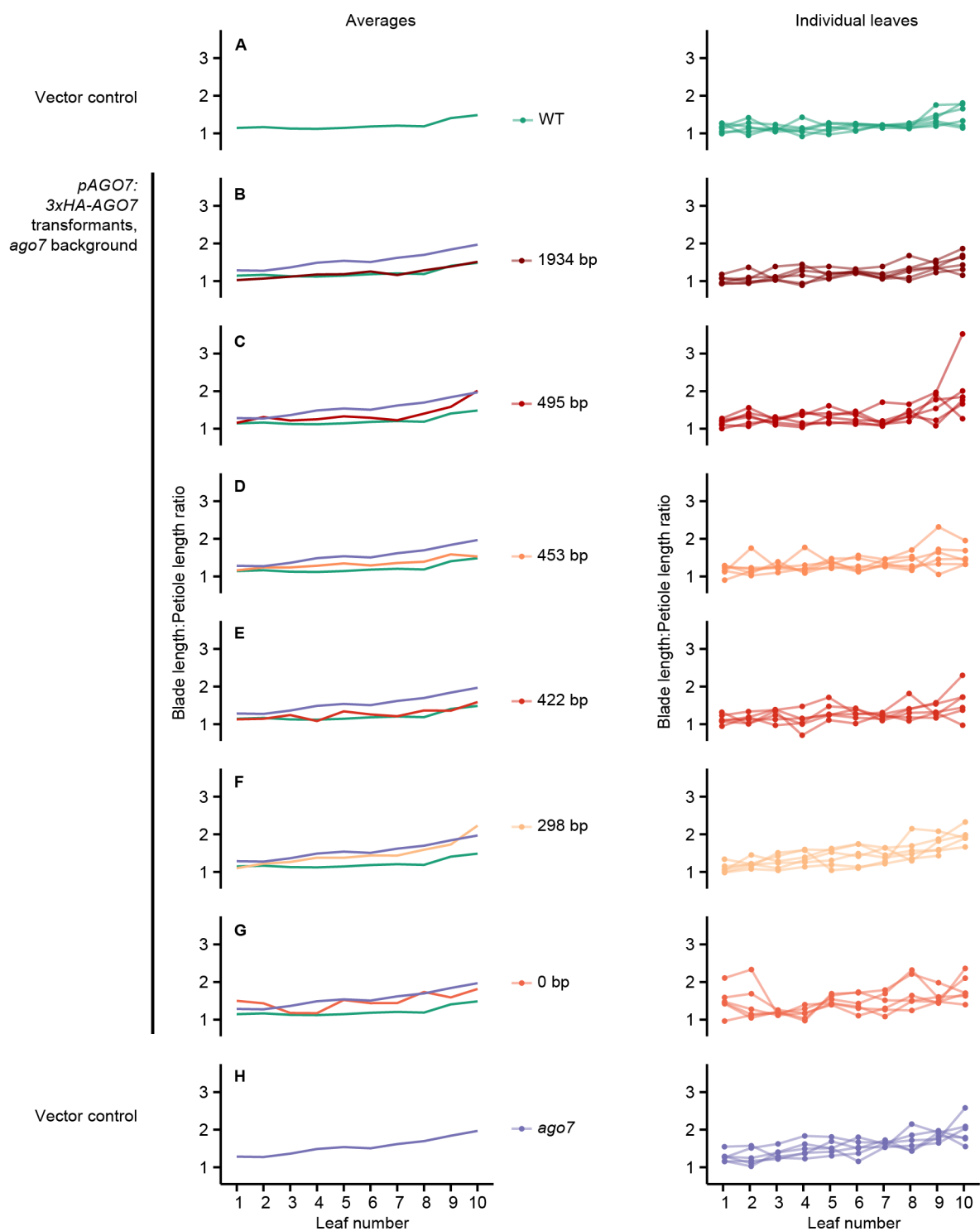


Figure 2.15: Complementation of *ago7* leaf shape defects, quantified based on leaf blade length to petiole length ratio. Panel layout is as in Figure 2.13.

The promoter lengths tested end immediately adjacent to core SPL and TCP binding sites (two ‘TGGTCC’ sites and one of two ‘GTAC’ motifs discussed above; see Figure 2.16). The 422 bp promoter transgene lacks all of these sites, but is sufficient for partial complementation (Figure 2.12, Figure 2.13E, Figure 2.15E). We therefore tentatively conclude that SPL and TCP binding is not required for *AGO7* transcription at levels that are sufficient for normal leaf morphology. The morphological data described allow us to estimate possible small differences between leaf shape in the complemented lines, but further experimentation would be necessary to relate such differences to cellular parameters or promoter structure.²¹

Finally, we scored appearance on trichomes on abaxial leaf surfaces to assess complementation of the forward shift in *ago7* mutants [113]. Consistent with results from previous transgenic experiments [50, 217], abaxial trichomes were visible on an earlier leaf for empty-vector-transformed mutant plants relative to corresponding wild-type plants (Figure 2.16); abaxial trichomes appeared 1.7 leaf positions earlier on average (95% confidence interval 0.5 to 2.9, $p = 4 \times 10^{-4}$, Tukey’s honest significant difference method). However, there was considerable variability, possibly due to effects from hygromycin selection. No 3xHA-AGO7 transgenic line showed a detectable increase in earliest abaxial trichome position (relative to empty-vector-transformed mutant plants; $p > 0.3$), indicating that none of the promoter lengths tested were able to drive full complementation of this defect. Alternative strategies may be required to assess ARF-mediated effects of AGO7 levels on trichome production.

²¹ As noted briefly in the introduction, chapter 3 describes additional time-resolved morphological analyses of these transgenic lines. None of the measurements presented in this chapter used Raspberry Pi camera photos (as described in the next chapter), except indirectly, as an aid for identification of leaf phyllotactic order. To reiterate, figures 2.13 and 2.15 are based on a different type of image processing, using a flatbed scanner and LeafJ [216].

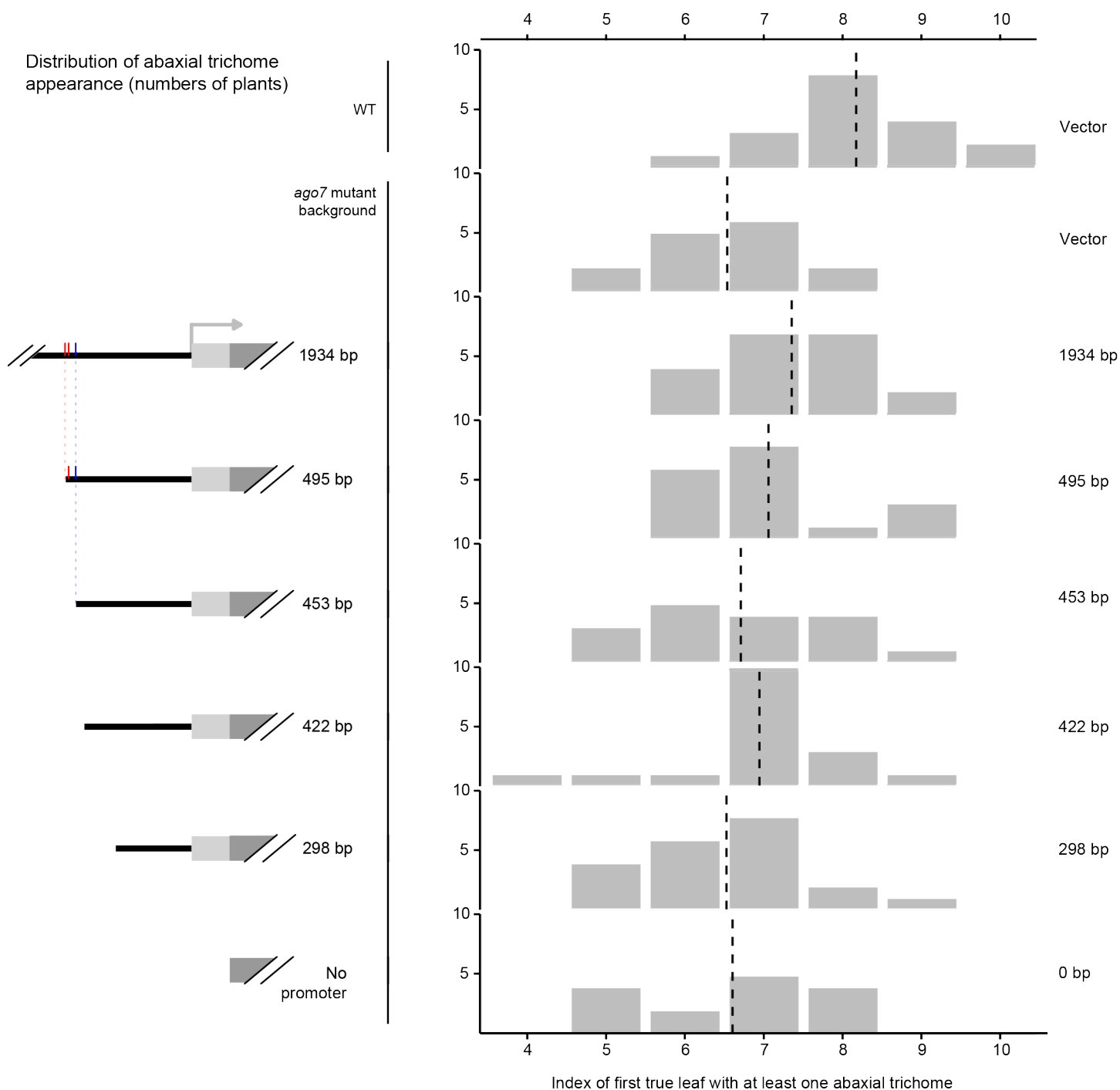


Figure 2.16: Assay for complementation of *ago7* early abaxial trichome appearance phenotype with 3xHA-AGO7 transgenes driven by truncated versions of the *AGO7* promoter. Dashed lines indicate the mean for each genotype. Core SPL (red) and TCP (blue) binding sites are indicated as in Figure 2.9.

2.4 Discussion

We characterized the structure of the *AGO7* promoter with transgenic analyses and a large-scale screen for upstream regulators. The most notable result from our Y1H analysis was a direct connection to multiple miR156-targeted SPL and miR319-targeted TCP factors. This result appears to reinforce the idea that gradual repression of *MIR156* transcription is the key regulatory step controlling heteroblasty in plants [88], and provides an additional example of functional linkage between SPL and TCP TFs [155, 218]. However, we were not able to assign a clear function to the candidate SPL and TCP binding sites in the *AGO7* promoter, particularly because a 422 bp proximal promoter region lacking all these sites is sufficient for substantial transgenic complementation of leaf morphology defects in *ago7* mutants (Figures 2.12 to 2.15).

Our truncation analysis provided preliminary evidence for two other functional regions of the *AGO7* promoter (Figure 2.17). We obtained different outcomes for mutant plants tested with 422 bp promoter constructs (largely complemented) versus 298 bp promoter constructs (not complemented). This difference suggests that one or more functionally important binding sites is present in the -422/-299 region. In general agreement with this idea, signal was qualitatively weaker for a 298 bp promoter:GUS reporter than for the next-longest promoter fragment tested (Figure 2.9). Multiple experiments suggest that the minimal core promoter and possibly one or more polarizing *cis* elements are intact in the 298 bp proximal region, but dissecting this further has been technically challenging because of the faintness of the signal. Despite progress, we did not succeed in our goal of identifying TF binding events necessary and/or sufficient for polar expression of *AGO7* and *AGO10*. It will be useful to integrate our results with other datasets, including for the *a priori* candidates shown in Figure 1.6. The YABBY1 gene is one promising candidate (Table 2.2), but did not emerge as a hit from the Y1H screens. Surprisingly, we also did not recover the polarity factor REVOLUTA for the *AGO10* promoter [102]; this likely represents a biological false negative.

AGO1 is ubiquitously expressed [103], and therefore expected to be under very robust transcriptional control which may be difficult to dissect. We believe the Y1H results presented here will be a useful resource, especially when combined with computational predictions such as the example shown in Table 2.2.

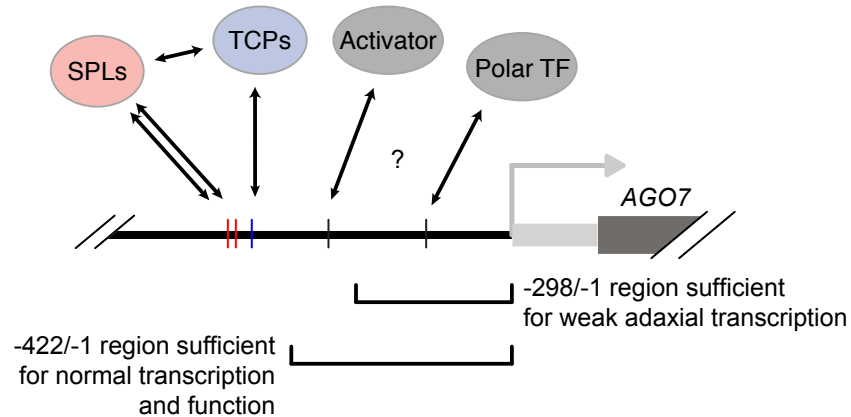


Figure 2.17: Schematic of the *AGO7* proximal promoter region with hypothesized TF binding sites and summary results from transgenic analyses indicated.

The truncation strategy used for our transgenic assays preserves the distance between *cis* elements, but also has inherent limitations. We did not test the possibility SPL and TCP core binding sites are *sufficient* for specific genetic functions. The apparent enhancer(s) in the -422/-299 region may be functionally redundant with these binding sites, and therefore largely masked any contributions to morphology through *AGO7*. Redundant clusters of activator binding sites appear common, as discussed below (section 4.3). Effects may be larger in other tissues, given the important functions of ARF repressors in fruits and roots [121, 123, 124]. Alternatively, the sites may simply be nonfunctional, at least in *A. thaliana*. Testing SPL and TCP binding in different tissues would help in assessing these possibilities, as discussed in the next section.

Table 2.2: Highest-scoring PWM scan match positions for *AGO7*, within first 425 bp for motif collection described in reference 197. PWM models were scanned across the 425 bp upstream region proximal to the annotated *AGO7* transcription start site using the ‘Find Individual Occurrences of Motifs’ tool (FIMO) [219] via the online MEME Suite [220] version 4.12.0, with default settings ($p < 10^{-4}$ cutoff). Complete scan results for the promoters of *AGO1*, *AGO7*, and *AGO10* with the PWM collections described in references 196, 197, and 198 will be provided in a supplemental Excel file.

TF	start	stop	strand	score	<i>p</i> -value	<i>q</i> -value
WRKY45	35	42	-	11.32	4.14e-05	0.13
WRKY12	35	42	-	10.99	9.01e-05	0.28
AHL20_2	61	68	-	10.52	9.69e-05	0.11
YAB1	134	141	+	10.93	5.88e-05	0.12
GLK1_2	299	306	+	11.71	3.96e-05	0.41

2.5 Follow-up experiments

Four types of experiments would be particularly useful to extend this work in the near-term future, for more direct inferences. Several of these are related to improved quantification or to well-understood limitations of Y1H validation approaches [214].

Test binding *in planta* with ChIP: We inferred direct DNA binding by SPLs and TCPs based on the combination of Y1H assays and PWM scans. DNase sensitivity data [200] suggests that these binding sites are accessible under a variety of conditions in at least some tissues of whole seedlings (not shown). These combined results, however, are no guarantee that SPL-promoter and TCP-promoter interactions occur in plant cells, let alone cause transcriptional changes. Two other factors that could reduce binding are competition with other TFs for sites and the possibility that protein-protein interactions interfere with SPL and/or TCP binding. Given the unclear function of the candidate sites described, it would be useful to biochemically test binding with chromatin immunoprecipitation. Timecourse analysis of differential binding would be valuable, as discussed further below in section 4.2.

Genetic sufficiency tests: As noted directly above, we demonstrated that core SPL and TCP binding sites are not required for adaxial transcription or transgenic complementation. We did not, however, test whether SPL and TCP binding sites are sufficient for normal expression and function, so it remains possible that the function of these sites overlaps with other TF binding sites. If this is the case, we would predict that fusing the core-binding-site-containing region to a heterologous minimal promoter would yield high-level GUS reporter signal and complementation of *ago7* mutants. The minimal 35S promoter [221] is the most common core sequence used for this purpose.

Quantitative GUS assays: We qualitatively assessed requirements for normal *AGO7* transcription using histochemical staining in transgenic reporter lines, but a more mechanistic understanding of developmental thresholds for *AGO7* mRNA levels will require quantification. Given the truncated-promoter lines in hand, quantitative assays based on 4-MU fluorescence are an obvious priority. It would be particularly interesting to test a possible role for the sites of interest in the (small) increase in *AGO7* levels as plants age [113]. Testing a larger number of independent transformants would also strengthen the results shown in Figures 2.11 and 2.10.

Molecular phenotypes: It seems likely that complemented transgenic lines have near-normal levels of *AGO7* protein and *TAS3* tasiRNAs, and thus low *ARF* mRNA levels. Unfortunately, we did not formally test these predictions, via immunoblotting, small RNA blotting, or qRT-PCR (per reference 217). Samples were collected with this end in mind but were never processed.

These four lines of work represent short-term experimental priorities, but many of other routes are available for approaching a holistic understanding of *AGO7* function. In chapter 4, I revisit the

key uncertainties about the action of *TAS3* tasiRNAs and related pathways throughout the plant life cycle.

2.6 Methods

2.6.1 Plasmid construction

Promoter fragments were PCR-amplified from previously described plasmids [50], with the primers listed in Table 2.3. Gel-purified PCR products were cloned with the pENTR D-TOPO kit (Invitrogen), and LR-recombined into several destination vectors: pGLacZi for Y1H screens [222], pMDC162 for GUS transcriptional reporters, and pMDC99 for transgenic complementation assays [223]. The destination vector pY1-gLUC59(GW) used for the secreted *Gaussia* luciferase Y1H reporter system has been described [213].

2.6.2 Y1H screens

Automated *lacZ* screens were done as previously described [178, 179] using a collection of 1541 TFs and an Agilent BioCel 1200 robotic platform. The TF-activation domain fusion yeast strain collection (arrayed in 384-well plates) was mated to bait strains. Diploid cells were selected in media lacking uracil and tryptophan, lysed by freeze-thaw, and assayed for β -galactosidase activity. Targeted Y1H assays were done similarly, with the lysis and assay steps replaced, essentially as described [213]. Briefly, diploid cells were resuspended in phosphate-buffer saline, 50 μ L of cells were transferred to a clear-bottom plate, and a Synergy H1 plate reader (Biotek) was used to inject

Table 2.3: Oligonucleotide sequences used for *AGO* promoter TOPO cloning. Primer names indicate position of 5'-most genomic base relative to the annotated transcription start site. Names also list the nucleotides added to create 'CACC' sequences for directional TOPO cloning.

Oligo name	Sequence
AGO1_-2308_FWD_cac	CACCCGCTTGTTAAACTCATAATC
AGO1_-1706_REV	TTAGGTGAAAGAATATCTAGAC
AGO1_-1755_FWD_cacc	CACCATCTAGACAATCTTTTGTTAG
AGO1_-1121_REV	GTTGCTCGTGCGTGAAGA
AGO1_-1170_FWD_cacc	CACCTACTCGTGACATATTCTCTA
AGO1_-536_REV	TATAAAGGATGTTATACAGTTAAG
AGO1_-585_FWD_cacc	CACCACAAGTACCAATTTTAAACTG
AGO1_-1_REV	TGCTACACTTTAAATTCAAGG
AGO7_-1934_FWD_c	CACCTGTCTCTTCTTCTGTACATGC
AGO7_-1436_REV	TAAGTATATTAAAAAATATCAGATGAC
AGO7_-1485_FWD_cacc	CACCTTATAGGTAAATGGATATGACT
AGO7_-941_REV	TGCTAAAACAAAAGATGCTCAA
AGO7_-991_FWD_cac	CACCCAAAGACATACATCTATAATATA
AGO7_-446_REV	AATTATGGGGACCATTCTGT
AGO7_-495_FWD_cacc	CACCAAGAAAATAGTACAAAGAATAAAT
AGO7_-1_REV	AGAAAGGGATTGTCTGAGTTT
AGO10_-2033_FWD_cacc	CACCGATTTCTATAAAAAATACATTCC
AGO10_-1511_REV	AGACCCCATTTTCGTGACT
AGO10_-1560_FWD_cacc	CACCGGAAGAAAACAAAATTAATGAG
AGO10_-991_REV	TAGTCTAGGTTAGTTTCCG
AGO10_-1040_FWD_cacc	CACCTATCACAACTAGACAATCC
AGO10_-471_REV	ACATCATTTGTTACAAGATGG
AGO10_-520_FWD_cacc	CACCTTTTTATAATAAGATTAGAGAATTAT
AGO10_-1_REV	ATAGCTTTCCTCTCAATGTG

Table 2.4: Oligonucleotide sequences directly cloned for *AGO7* promoter mutation analysis in yeast. Forward sequences (5' to 3' in the direction of *AGO7* transcription) are followed by corresponding reverse sequences. Dashes indicate bases “deleted” relative to the genomic reference sequence. The last sequence for each set is a degenerate oligo; a single clone resulting from these oligos was used, as indicated in the caption for Figure 2.6.

CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAGGTACGAAGAAAATAGTACAAAGAATAAATAATTAAACAGAATGGTCCCCATAATT
CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAGGTACGAAGAAAATAGTACAAAGAATAAATAATTAAACAGAA-----CCATAATT
CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAGGTACGAAGAAAATA----AAAGAATAAATAATTAAACAGAATGGTCCCCATAATT
CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAG----GAAGAAAATAGTACAAAGAATAAATAATTAAACAGAATGGTCCCCATAATT
CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAG----GAAGAAAATA----AAAGAATAAATAATTAAACAGAATGGTCCCCATAATT
CACCGAAAGAGTCCAAAGTGTGTATTATTAATGAGHVBDAAGAAAATAHVBDAAGAAATAAATAATTAAACAGAATGGTCCCCATAATT

AATTATGGGGACCATTCTGTTTAATTATTTATTCTTTGTACTATTTTCTTCGTACCTCATTAATAATACACACTTTGGACTCTTTCGGTG
AATTATGGTTCTGTTT-----AATTATTTATTCTTTGTACTATTTTCTTCGTACCTCATTAATAATACACACTTTGGACTCTTTCGGTG
AATTATGGGGACCATTCTGTTTAATTATTTATTCTTT----TATTTTCTTCGTACCTCATTAATAATACACACTTTGGACTCTTTCGGTG
AATTATGGGGACCATTCTGTTTAATTATTTATTCTTTGTACTATTTTCTTC----CTCATTAATAATACACACTTTGGACTCTTTCGGTG
AATTATGGGGACCATTCTGTTTAATTATTTATTCTTT----TATTTTCTTC----CTCATTAATAATACACACTTTGGACTCTTTCGGTG
AATTATGGGGACCATTCTGTTTAATTATTTATTCTTTTHVBDTATTTTCTTCHVBDCTCATTAATAATACACACTTTGGACTCTTTCGGTG

Table 2.5: Forward primer sequences used for TOPO cloning of truncated versions of the *AGO7* promoter. Names follow Table 2.3. Bases -298/-295 are a natural ‘CACC’ sequence suitable for directional TOPO cloning.

Oligo name	Sequence
AGO7_-482_FWD_cacc	CACCAAAGAATAAATAATTAAACAGAATGGTCC
AGO7_-453_FWD_cacc	CACCCCATAATTCGATTTAATGAGTGTATTG
AGO7_-422_FWD_cacc	CACCATTTTATAAAACATGTGTAACAACAACAA
AGO7_-298_FWD	CACCAAACATTATCGGTAATCACTA
AGO7_-150_FWD_cacc	CACCTATTTTCTTTTATTATTGCCAACAATT
AGO7_+1_FWD_cacc	CACCGCCTCTTTTATCTCTCTCTCATAAA

10 μ L of 20 μ M coelenterazine substrate solution into each well and read luminescence immediately afterward (0.1 s integration time).

2.6.3 Plant materials and growth conditions

All *A. thaliana* plants descended from the reference Col-0 accession. The *zippy-1* mutant allele was isolated by Hunter et al. [113], and is referred to throughout as “*ago7*”. Plants were transformed by floral dip using *Agrobacterium* strain GV3101 [224, 225].

Plants were grown under short day conditions (8 hours light, 16 hours dark) in a Conviron MTR25 reach-in chamber with PolyLux fluorescent bulbs (200 μ mol photons per second per square meter) at 22 °C with 50% humidity.

2.6.4 *ago7* mutant complementation tests

Measurement of leaf phenotypes followed previous work [205]: we scored the index of the earliest leaf with at least one abaxial trichome using a stereomicroscope at 28 to 30 days post-stratification, and concurrently measured the blade length, blade width, and petiole length for the sixth true leaf with digital calipers (Mitutoyo, Japan). At a later timepoint (33 and 35 days post-stratification), we dissected and scanned the first ten true leaves from each plant with a Canon Pixma MP190 flatbed scanner. Leaf shape parameters were measured with the LeafJ plug-in for ImageJ [216].

2.6.5 GUS assays

Histological GUS assays were essentially as described [117, 226, 227]. Seedlings were collected into ice-cold 90% acetone, incubated at -20 °C for 20 minutes and then room temperature for another 20 minutes. Seedlings were washed twice (5 minutes each) with staining buffer (100 mM sodium phosphate [pH 7], 20% methanol, 0.1% Triton X-100, 1.5 mM ferri- and ferrocyanide).

Staining buffer with 0.5 mg/mL 5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid (X-Gluc) was vacuum-infiltrated into seedlings on ice for two rounds of 15 minutes each. Samples were then incubated at 37 °C for 20 hours, taken through an ethanol/histoclear series, and infiltrated with Paraplast Plus at 60 °C, before embedding [226]. Tissue sections (10 μ m thickness) were mounted on Probe-On Plus slides (Thermo Fisher), deparaffinized with histoclear, coverslipped, and photographed.

2.7 Acknowledgments

We thank Danforth Center Plant Growth Facility staff for excellent plant care, G. Nguyen and R. Allscheid for logistical support, and members of the Carrington lab for helpful discussions. D.H. Chitwood and J.G. Hodge provided useful advice on histological analysis. We thank T.C. Mockler and members of his lab (J. Gierer, D. O'Brien, M. Wiechert) for assistance with their plate readers and liquid-handling equipment. This work was supported by US National Science Foundation award 1330562 to JCC. JSH was supported by an NSF graduate research fellowship (award 1143954).

Chapter 3

Improved methods for time-lapse imaging of vegetative growth

Time-course analysis is essential for measurement of dynamic traits and for developmental context. The *AGO7* gene plays a major role in developmental timing, but the analysis of *cis* requirements for its function described in the last chapter focused on single time-point measurements (section 2.3.3). Here I extend this analysis and document new methodology for time-lapse imaging of rosette development.

3.1 Abstract

Top-down imaging is a simple method for measuring the dynamics of *A. thaliana* rosette growth; several approaches have been described. We describe an effective fixed-position top-down Raspberry Pi camera setup for imaging sets of 180 plants in a two-level controlled-environment growth chamber, with photo captures every five minutes. Segmenting plant rosette foreground from soil background is straightforward; no special modifications of growth conditions are required. We

applied this method to transgenic complementation analysis of a mutant with downward-curved leaves.

3.2 Introduction

Improvements in automated image-based measurement have the potential to dramatically accelerate progress in many areas of plant biology [201]. Improved methodology has become increasingly important as genotyping and molecular profiling costs have dropped precipitously. A variety of tools will need to be optimized for different systems and scales of investigation [228].

Arabidopsis thaliana is the premier model plant species for molecular genetics and is thus an important target system for image-based phenotyping. Experimenters using *A. thaliana* benefit from its small size and ease of transformation, and can also draw on the exceptional genetic and genomic resources available, such as large collections of indexed mutants and fully-sequenced naturally inbred strains. These advantages have enabled many seminal discoveries [229].

A. thaliana leaves initially emerge in a relatively flat spiral rosette, so top-down imaging is sufficient to capture a great deal of information about growth. Accordingly, a body of work has emerged describing time-lapse imaging of rosettes and analysis of the resulting images. Several groups have recently begun time-lapse photography experiments using the microcomputers and cameras developed by the Raspberry Pi Foundation. The low cost of these rigs allows one to image many plants in parallel; this approach complements lower-throughput methods, which achieve high resolution depth information but rely on sophisticated instrumentation and typically incur high costs [230].

3.2.1 Biological context and approach

Analysis of leaf shape changes (heteroblasty) has been a powerful tool for understanding fundamental processes [231], including microRNA and trans-acting short interfering RNA biogenesis [232]. We have used transgenic mutant complementation analysis to identify promoter element requirements for the function of *A. thaliana* *ARGONAUTE7*. We sought to extend single-timepoint analysis to quantify the dynamics of rosette changes.

Our primary goal in developing our photography approach was to enable visual identification of leaves in their spiral phyllotactic order; accurate determination of order is essential for comparison of shape changes across genotypes. A secondary goal, anticipating that analysis algorithms will continue to improve, was to generate time series information for quantitative analysis. Our approach has been to:

- Use standard growth conditions that promote plant health;
- Begin imaging plants at the earliest possible point in their development, i.e. directly after recovery from transplanting; and
- Err on the side of photographing at higher frequency than needed, until the effect of image capture frequency on leaf tracking analysis can be assessed.

3.3 Methods

3.3.1 Growth conditions

As described in subsection 2.6.3, plants were grown in a Conviron MTR25 reach-in chamber with PolyLux fluorescent bulbs (200 μmol photons per second per square meter) at 22 °C with 50% humidity. Short day conditions (8 hours light, 16 hours dark) were used to delay shoot elongation and flowering. Fifteen 3-inch pots containing ProMix FPX growth mix were used per flat. Genotype positions were randomized within each flat, and plants were given serially increasing numbers, indicated with printed white labels (Midsci DTCR-4000, 3/8 inch diameter) affixed to the corner of each pot (Figure 3.1). Six flats were arranged on each of the two growth shelves (Figure 3.2), allowing imaging of 180 plants simultaneously. Plants were generally watered and fertilized twice per week; watering times were recorded.

3.3.2 Hardware and software configuration

Five of the components of each imaging rig were purchased from Newark element14 (Table 3.1). Three of these components have since been superseded, as discussed below.

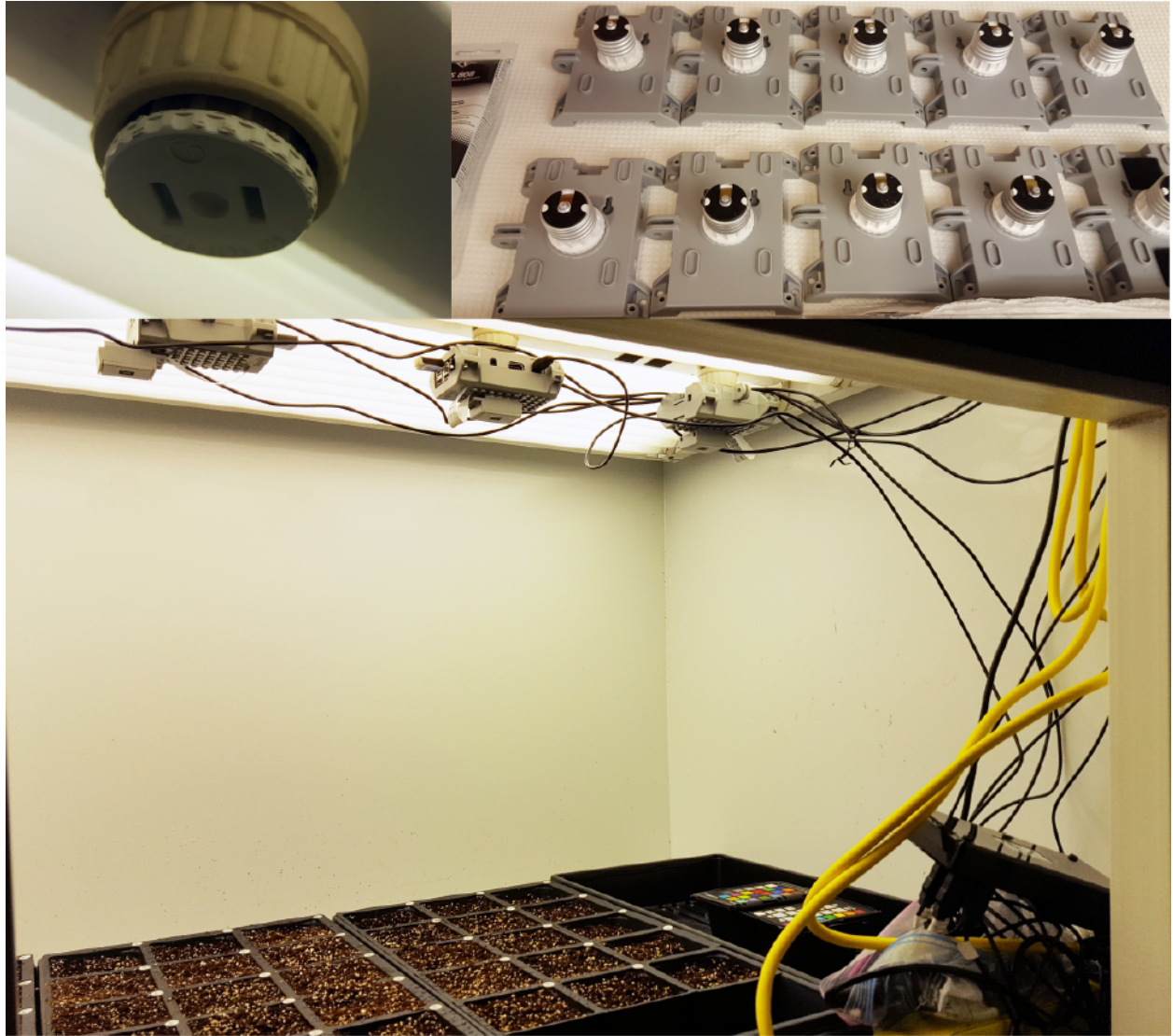


Figure 3.1: Assembly and mounting of Pi/Camera imaging rigs. Incandescent bulb socket adapters were attached to cases with silicone adhesive and then screwed into sockets. Third photograph shows three Pi/Camera rigs above plant flats and color standard cards in upper level of the controlled environment growth chamber. USB power supplies (black) and surge protector (yellow) are in view. See Figure 1 in Appendix B for additional photos.

Table 3.1: Components for a single Pi/Camera rig.

Component	Newark part
Raspberry Pi 2 Model B board	38Y6467
Raspberry Pi camera module v1 (5 megapixel)	69W0689
MicroSD card (8 GB, Sandisk)	38Y6470
Wifi module (“WiPi dongle”)	07W8938
USB to microUSB power cable	06W1049
SmartiPi case kit #2 (Smarticase LLC)	

Camera module lenses were manually focused with pliers for a lens-plant distance of 55.2 cm. To suspend Pi/Camera rigs, we first attached AC power socket adapters to the the back of SmartiPi cases with silicone adhesive (Figure 3.1). We then encased Pi boards and cameras and screwed the attached socket adapters into incandescent bulb sockets above each level of the growth chamber. We set up a total of twelve Pi/Camera rigs, one directly above each flat. Pi/Camera rigs were powered through two USB power supplies drawing power (via extension cord and surge protector) from an auxiliary power outlet built into the growth chamber.

Image acquisition Time-lapse imaging essentially followed the tutorial at RaspberryPi.org: image captures (`raspistill` command) were scheduled at five-minute intervals using a `cron` table. Images were pulled from each Pi to a remote server twice per hour (using `rsync`) by a server-side `cron` process. The concrete details are online²² and are documented further in Appendix C. Additional notes and instructions are in Appendix B.

²² See <https://github.com/jshoyer/raspi-topdown-plant-imaging-12x>

A

Flat 1	Flat 2	Flat 3	Flat 4	Flat 5	Flat 6
501 506 511	516 521 526	531 536 541	546 551 556	561 566 571	576 581 586
502 507 512	517 522 527	532 537 542	547 552 557	562 567 572	577 582 587
503 508 513	518 523 528	533 538 543	548 553 558	563 568 573	578 583 588
504 509 514	519 524 529	534 539 544	549 554 559	564 569 574	579 584 589
505 510 515	520 525 530	535 540 545	550 555 560	565 570 575	580 585 590

B Field of view 1

516	517	518	519	520
511	512	513	514	515
506	507	508	509	510
501	502	503	504	505

C Field of view 4

556	557	558	559	560
551	552	553	554	555
546	547	548	549	550
541	542	543	544	545

Figure 3.2: Schematic of plant arrangement and camera fields of view for each chamber level.

A. Plants were given unique numbers, in column-first order as one looks into the chamber. For the experiment emphasized in this chapter, plants were numbered 501 to 590 in the upper level of the chamber (as shown here) and 601 to 690 in the lower level. Black and gray lines at the bottom of the panel indicate the extent of each camera field of view; numbers of plants included in fields of view 1 and 4 are in black.

B. Camera fields of view are oriented 90° relative to the perspective in panel A, and plant numbering within photographs is therefore in row-first order, starting from the bottom (south side) of the image. Because each field of view covers more than one flat, each photo contains a “secondary” row of pots; for fields of view 1 to 3 this secondary row is in the top (north) side of the image.

C. Fields of view 4, 5, and 6 (as well as 10, 11, and 12) contain the secondary row of pots in the bottom (south) side of the image.

3.4 Results and Discussion

3.4.1 Reliable acquisition and transfer of images

We began to use Raspberry Pi cameras because their low cost makes it practical to leave multiple stationary cameras above a large number of plants. We tested 1) whether the failure rates of such low-cost hardware would affect image capture, and 2) how efficiently we could transfer images to a remote server.

Wifi signal was a concern because of the thickness of growth chamber walls, but our local wireless network sufficed for efficient “headless” operation (via Secure Shell [SSH] connection to each RasPi), and also for transfer of large number of images in preliminary tests. We could therefore use `cron` tables to schedule image capture at a high rate (5, 10, 15, . . . minutes past the hour). Higher frequency imaging is generally undesirable, because it would increase the difficulty of avoiding capture times when opening the chamber doors for watering or other manipulations. We scheduled image transfer from each RasPi twice per hour using `rsync`, a standard Linux utility designed for high-latency low-bandwidth connections [233]. Transfers were staggered, to reduce wifi interference. Because of the predictable file naming scheme employed, it was straightforward to programmatically confirm that all photo timepoints were captured and transferred as scheduled.

In several months of operation, image capture has only been interrupted by a single crash of one computer. This crash was detected within half an hour (via `cron` error emails and Ganglia Monitoring System dashboard) and corrected by restarting the affected machine. Several image transfers per day initially failed, do to transient wifi interference from passerbys carrying phones and the like, but the affected files were always transferred in subsequent rounds, after wireless signal improved.

More recent Raspberry Pi boards, including the lower-cost Zero W model, include a built-in wifi module. Transfer of images over a wireless network will therefore continue to be a useful option for small-to-moderate-sized imaging installations.²³

3.4.2 Image processing

We have begun developing tools for extraction of quantitative plant trait information from the photographs described using the PlantCV and OpenCV software packages. The rationale for using and enhancing PlantCV is outlined in reference 234 and Appendix D. Simple thresholding based on the green-magenta a^* component of $L^*a^*b^*$ color space is largely sufficient for distinguishing plant foreground pixels from background (Figure 3.3). Use of this transformation is reasonably common—see for example reference 235. Plants can be identified by finding contours (`cv2.findContours` function) within binary thresholded images. Contours for many plants are not connected, often because bits of soil occlude thin leaf petioles (Figure 3.4). Single plant objects are defined by joining all contours that partially overlap with a circular region of interest around the center of each plant (Figure 3.3B). We manually identified plant center coordinates with the ImageJ point tool, which is a quick process for an experiment with this number of plants. At least one other group has found it expedient to use essentially the same method [236] and the plant center landmark position data will be useful for future work.

Once individual plants have been adequately segmented, one can use the shape analysis functions in PlantCV to measure rosette shape descriptors (Figure 3.3C, Figure 3.4). Multiple metrics depend on calculating shapes that surround each plant, particularly the minimum bounding ellipse and minimum convex polygon (convex hull) required to surround the plant. The current `analyze_object`

²³ One wifi-networked Raspberry Pi imaging system ten-fold larger than the one described here is documented at <https://github.com/calizarr/PhenoPiSight>

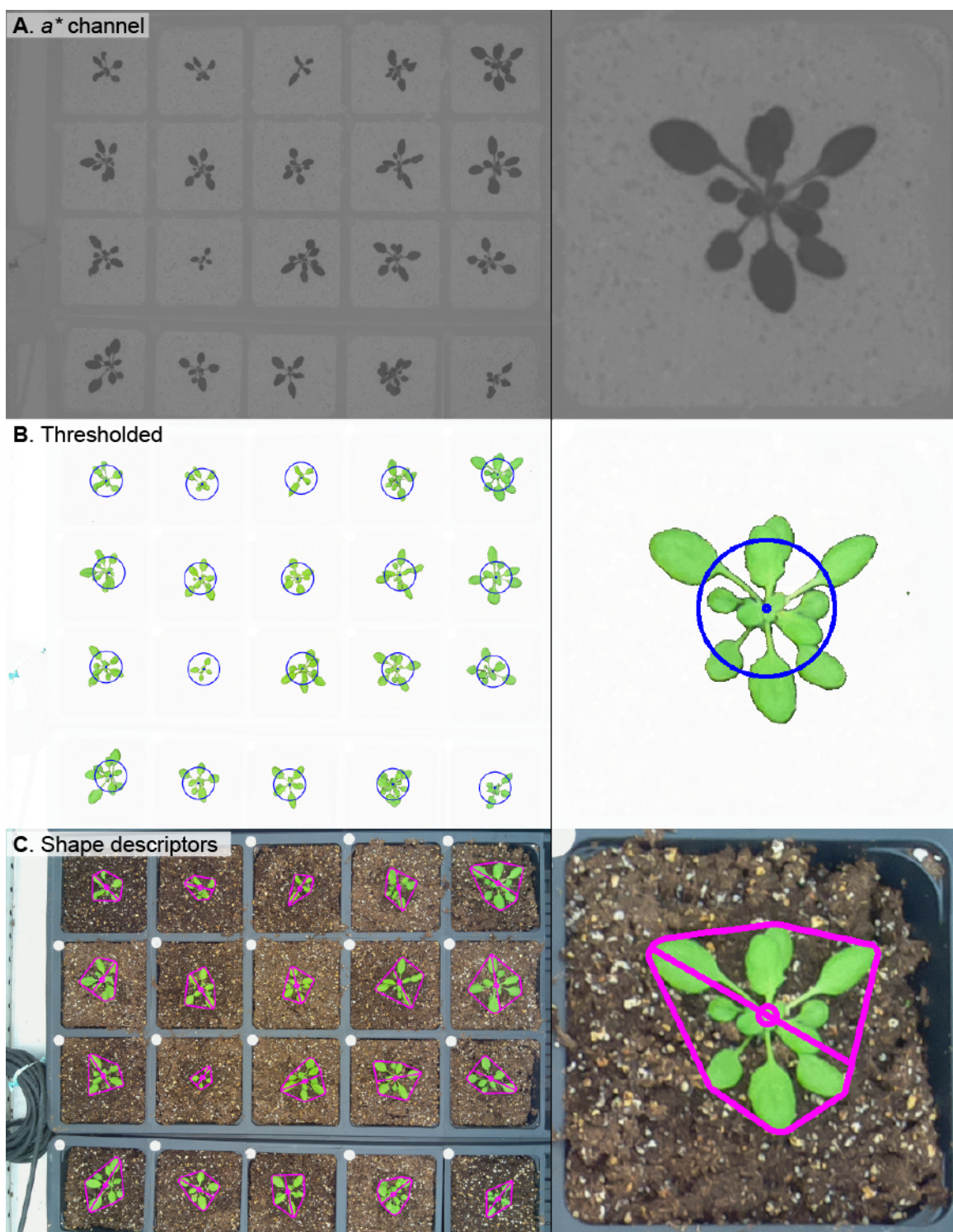


Figure 3.3: Flow diagram illustrating steps in image processing. A. The a^* channel from each image is extracted and used for binary thresholding. The top right plant is shown in inset. (Caption continues on next page.)

Figure 3.3: (Continued) B. Thresholding yields a “masked” image (background masked here with 97.5% transparency), and circular regions of interest are used to group disconnected contours into rosette objects.

C. The minimum bounding ellipse (not shown) and convex hull for each plant object are identified, allowing calculation of size and shape parameters. The longest dimension and convex hull outline are represented on the original image with magenta lines.

function returns sixteen measurements, of which nine are relevant for top-down views of a plant. The “simple” metrics are either integer counts (number of convex hull vertices) or are provided as pixel counts reflecting areas (projected plant area, convex hull area) or lengths (perimeter length, longest axis length, and minimum bounding ellipse major and minor axis lengths). Two composite metrics are ratios, namely solidity and object bounding ellipse eccentricity (0 to 1 scale). For the preliminary analysis presented here we use pixel counts rather than converting to physical distances. Because of the overlap between fields of view (Figure 3.2), one third of the plants were photographed continuously with two different cameras. For the comparison below we consider only the “primary” view, for each plant (field of view 1 for plants in flat #1, field of view 2 for plants in flat #2, etc.) but correcting for optical distortion would likely improve spatial accuracy, as discussed further in subsection 3.5.2. Shape descriptor measurements for primary and secondary views (for the sixty plants photographed twice) are well correlated (not shown).

The method for plant segmentation described above suffers from two main deficits. First, growth of algae periodically causes missegmentation of individual small pieces of the whitish perlite in the growth medium (Figure 3.3B, far right). Classification of perlite pixels as “plant” can in some cases dramatically alter the shape of the convex hull and related shape measurements for small plants (Figure 3.5). Adding a step to “fill” small objects will likely resolve this issue. Second, disconnected leaf contours occasionally fall outside of the circular ROI, resulting in erroneously low projected leaf area and shape descriptor measurements. Directly drawing shape descriptors on output images facilitates visual detection of such errors (Figure 3.5). For timecourses subsets

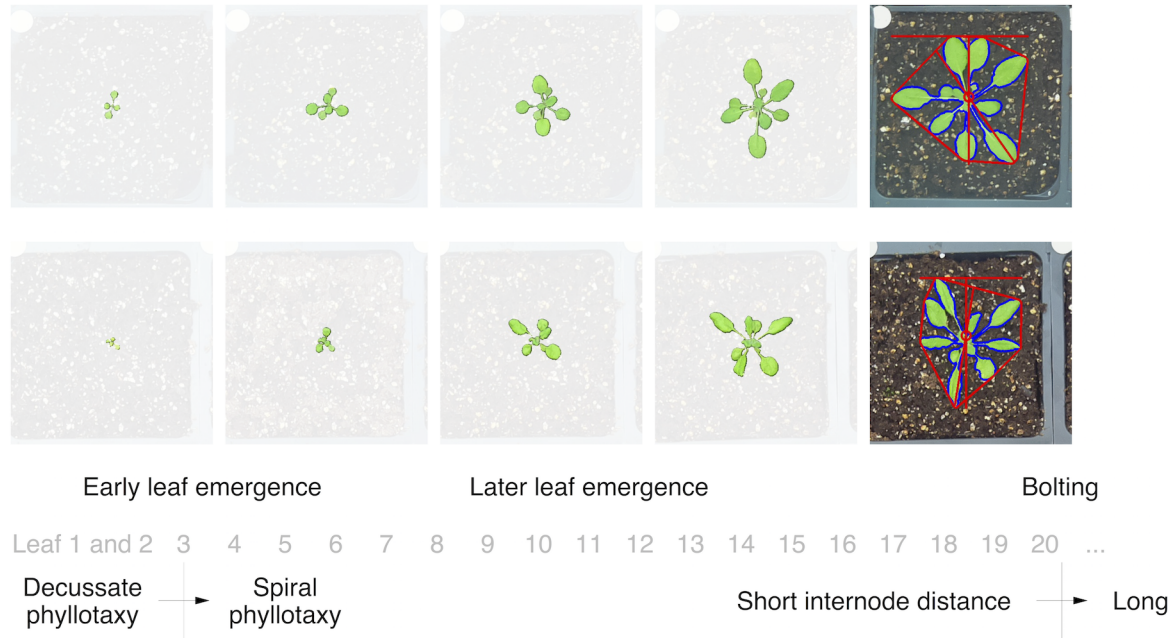


Figure 3.4: Top-down view of the course of vegetative development for plants showing wild-type rosette morphology (upper panels) and the downward-curved-leaf *ago7* mutant phenotype. Cropped images from 12, 17, 22, 27, and 32 days post-stratification are shown. Four timepoint panels show pixels above the a^* threshold applied, by applying semitransparency to background pixels. The last timepoint panels outline the identified object in blue and indicate convex hull, centroid position, x- and y-extent, and maximum dimension in red; these are the default colors used in PlantCV output, rather than the magenta, as used in Figures 3.3 and 3.5. Our computational pipeline writes both types of cropped image into subdirectories by plant and by timepoint, to facilitate assessment of plant segmentation.

(e.g. one photo per day) it is feasible to inspect output images for all plants and exclude timepoints with segmentation errors. Alternatively, for experiments with reasonable sample sizes, simply averaging can smooth out the noise introduced by occasional segmentation errors. Methodology for segmenting rosettes (a multi-instance segmentation problem) was not the main interest here; comparison to two other methods is provided below, in subsection 3.5.1.

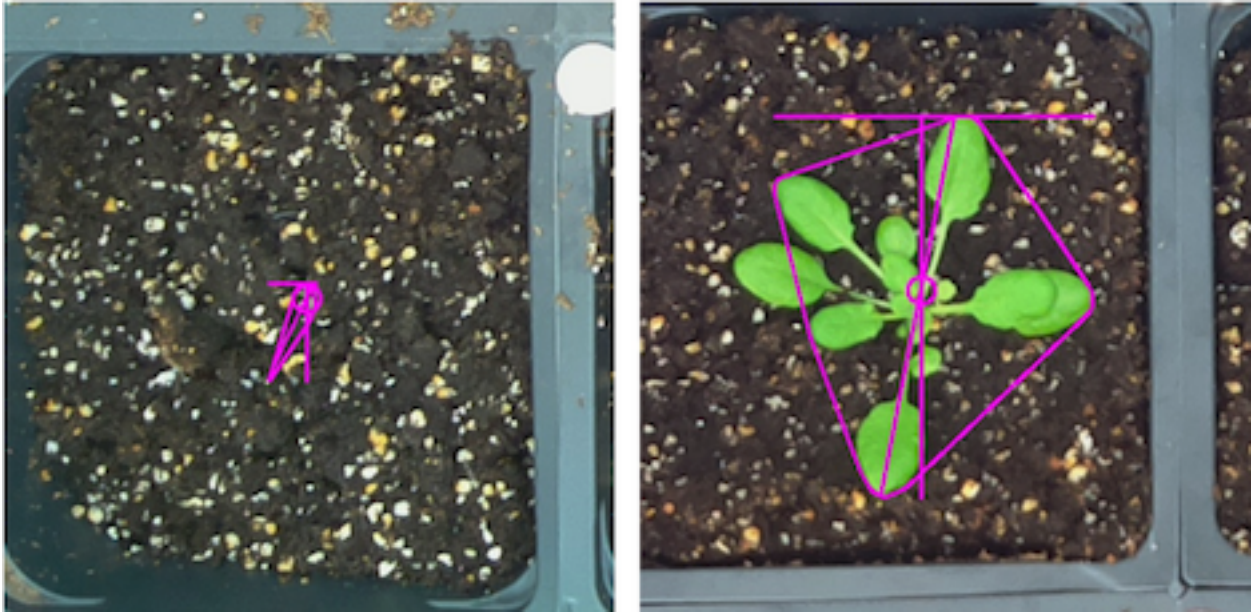


Figure 3.5: Examples of two types of rosette missegmentation and effect on calculated shape descriptors.

Left: Perlite pixels incorrectly included in a plant object, because they exceed the a^* threshold and are present in the circular region of interest for the given plant. This results in an incorrectly large convex hull, longest dimension, etc. (magenta lines).

Right: Leaf pixels incorrectly omitted from a plant object, because the relevant leaf blade contour is disconnected from the rest of the plant and entirely outside of the defined circular region of interest.

The fluorescent growth lights in our growth chamber have yielded very bright and even illumination. We have relied on automatically determined camera exposure and color-balance settings, for simplicity. Our thresholding results indicate that this approach works well for morphological analysis of healthy plants. Comparison of color traits measured with multiple camera sensors

over the course of an experiment will clearly benefit from optimization of both acquisition and post-processing settings, as discussed further below.

3.4.3 Rosette size and shape dynamics in *ago7* mutants

We hypothesized that the shape metrics described above could capture aspects of the *ago7* mutant phenotype over the course of the experiment described in the last chapter (section 2.3.3), and might provide information on when the difference between mutant and wild-type plants (transformed with empty vector) was greatest. As expected, the downward curling of leaves of *ago7* mutants resulted in lower area in two dimensional view from above (projected area; Figure 3.6A). Several metrics indicated that wild-type plants were larger on average in this experiment (Figure 3.6, panels A, B, D, E, G, and H). Dividing projected leaf area by convex hull area (to calculate solidity) has the effect of normalizing some differences in size (Figure 3.6C); solidity was initially comparable for the two reference genotypes but lower for the mutant plants at later timepoints, reflecting the thinness of mutant leaves as viewed from above.

The number of vertices in the convex hull for wild-type plants likely because the “pointy” leaves of *ago7* mutants often only defined a single vertex whereas the rounder wild-type leaves often defined several (Figure 3.4). Rosettes were perhaps slightly more oblong for mutant plants relative to wild-type, as assessed by the unstable metric of minimum bounding ellipse eccentricity (Figure 3.6I). Average values for the six mutant lines transformed with 3xHA-AGO7 constructs were closer to the average for wild-type group than to mutant plants for all nine metrics (not shown).

Many caveats apply to this analysis, as to any timecourse analysis of shoot growth [237]. Most of these relate to improved normalization and statistical modeling. One concern is that mutant plants were simply smaller than wild-type plants due to slightly later germination or the like,

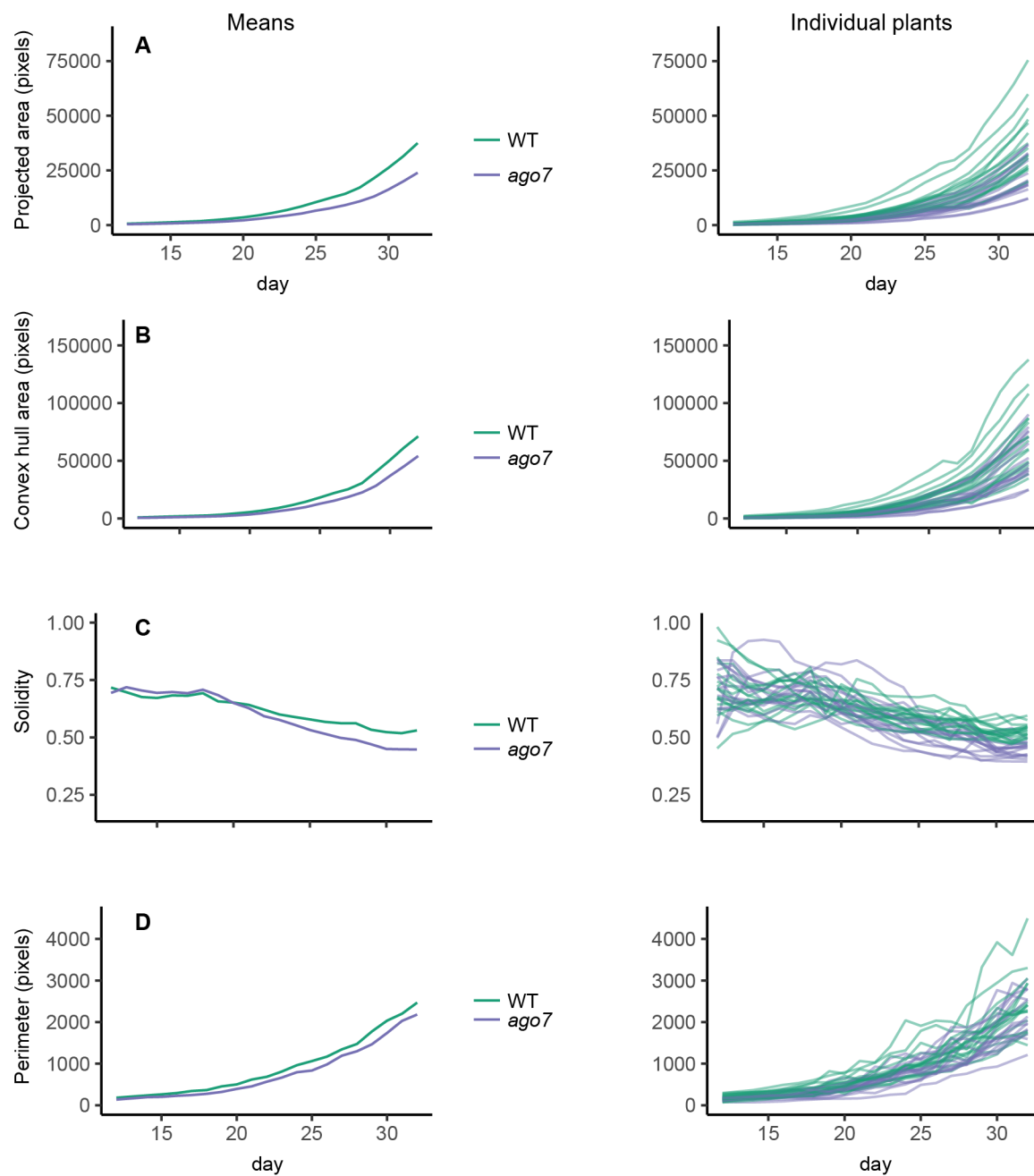
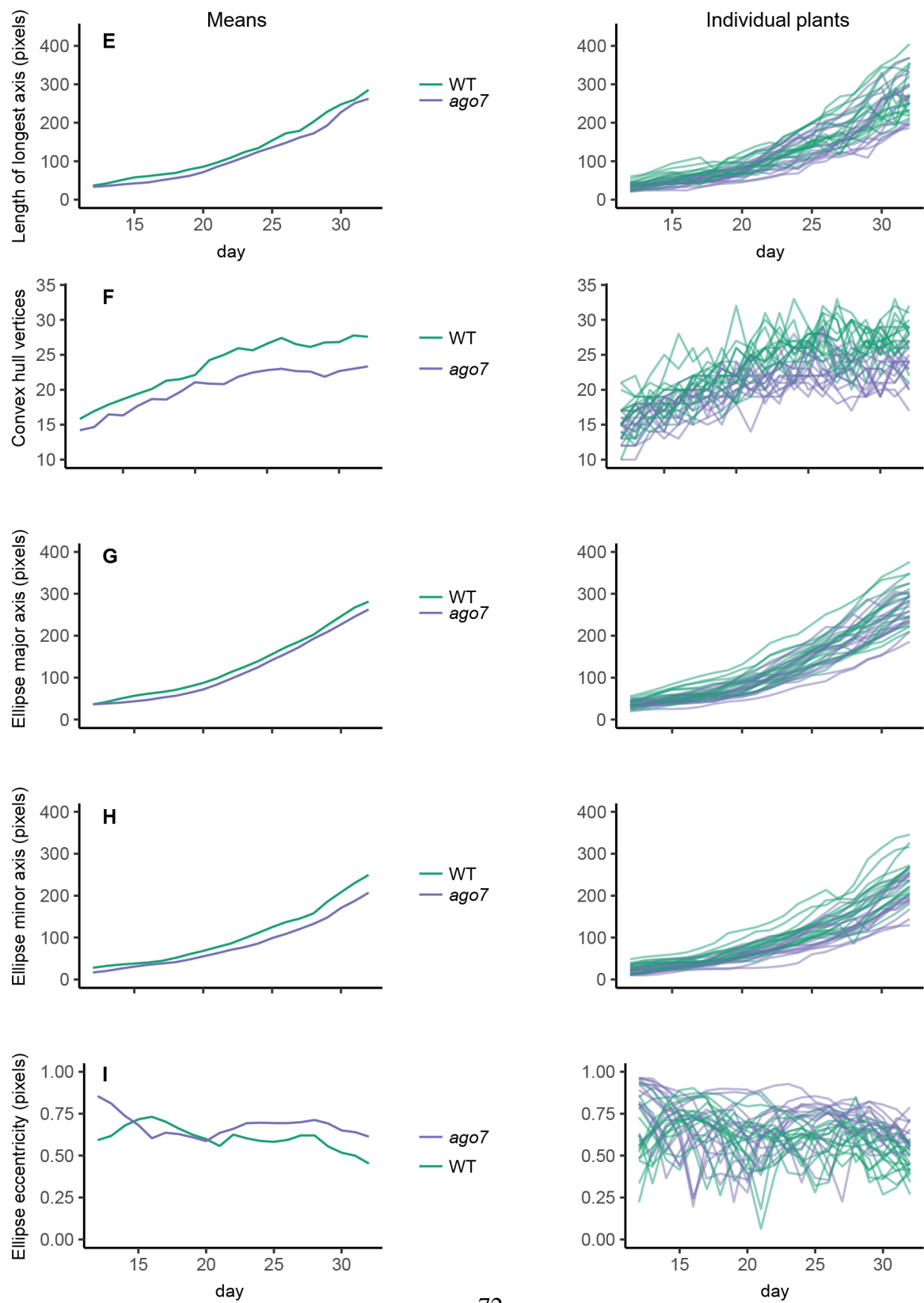


Figure 3.6: Shape and size of wild-type and *ago7* mutant rosettes over time. Left panels show averages for rosette shape descriptors by genotype from days 12 to 32 days post-stratification. Right panels show values for individual plants, colored by genotype. Graphs continue on the next page.



owing to factors other than their genetic difference. Comparing two staggered time series can result in systematic errors [237], and we have not analyzed relative growth rates or other rates of change. Many possibilities exist to improve effect size estimates using on improved modeling, (to account for the effect of plant position, light intensity, etc.) and combination of multiple metrics into higher-dimensional summaries (e.g. through principal component analysis). Integration of information from different experiments will allow estimation of how much of the plant-to-plant variation within genotype groups might be due to technical variability. I have a considerable amount of image data from highly replicated experiments is in hand for this purpose. There are of course also many opportunities for improvement of the image processing methods used, as described in the next section. Values for some size-related metrics did not monotonically increase as they logically should, due to the rare but significant types of segmentation error described above. Fortunately, the fine-grained information in hand will allow continual extension of the analysis presented here, both visually and statistically, as algorithms improve.

To explore the use of the top-down photos for leaf-level measurements, I assessed correlation with flatbed-scanner-based measurements described in the last chapter (Figures 2.15 and 2.13) by identifying leaf blade tip and base coordinates for upper-level plants at the appropriate photo time-point for each. The distance between these points (in pixels) represents leaf blade length. As expected, leaf blade lengths measured with these two complementary methods were well correlated (Figure 3.7). The correlation was not quite as good for petioles, presumably because they are short and the relative measurement variability is proportionally larger. In addition to such variation, differences in leaf angle and leaf curvature are well-known to cause deviations from linearity [238]. Collecting the landmark data required for this type of analysis is time-consuming, but valuable for benchmarking automated methods. Collecting landmark data from relevant time-points would enable similar correlation analysis for unflattened leaves using the manual caliper measurements in hand (Figure 2.14).

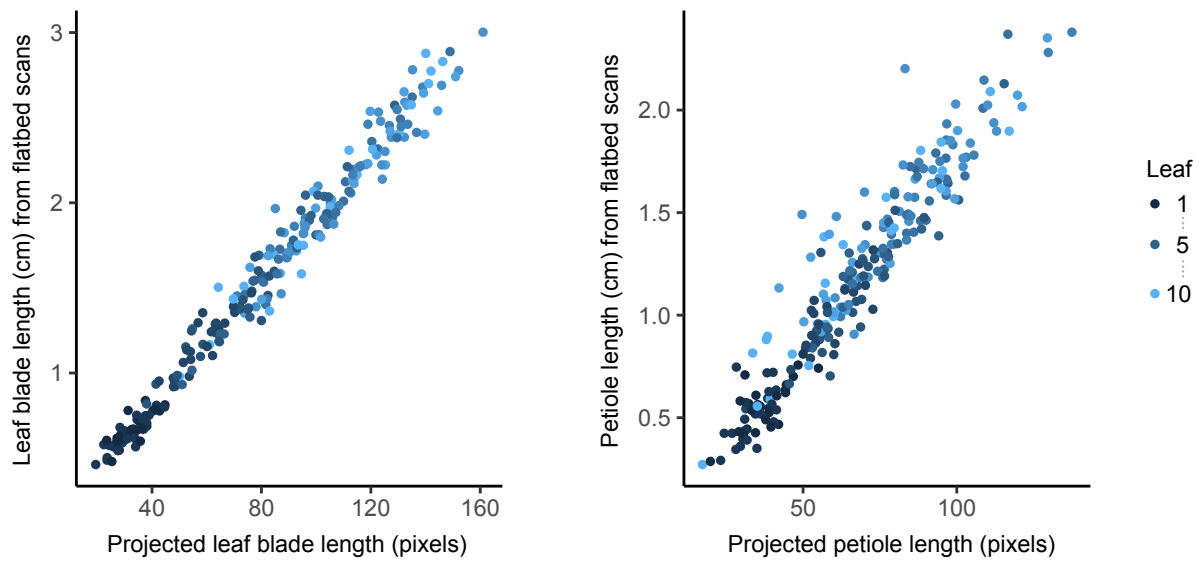


Figure 3.7: Correlation between leaf-level measurements in flatbed scans and top-down photos: leaf blade length (left) and petiole length (right). Flatbed scanner measurements (Y axes) were described in the last chapter. For each of the 24 plants, coordinates for approximate position of the plant center and the tip and base of each leaf blade were identified with the ImageJ point tool in the top-down photo taken immediately before each plant was removed from the growth chamber.

3.5 Possible improvements and related work

High-throughput phenotyping is in its infancy; the next few years will see dramatic improvements. Low-cost custom-built systems currently lack many features and conveniences present in commercial systems. Adding such features will require significant effort, but should also create opportunities for increased transparency, flexibility, and innovation.

3.5.1 Phenotiki resources: comparison and priorities for benchmarking

Minervini et al. [239] summarized the principal advantages of fixed-camera Raspberry Pi systems: low cost and lack of moving parts. Our work differs from their proof-of-concept experiment in that we use twelve Pi/Camera rigs, higher image capture frequency, and the shorter camera-plant distance necessitated by our growth chamber. The shorter camera-plant distance used yields higher spatial resolution, but requires manual focusing of lenses and reduces the size of the field of view. Some advice related to these considerations and others is provided in Appendix B.

Several methods for segmenting individual plants have been described, including methods with a higher degree of automation than the circular-region-of-interest approach described above. The k -means clustering localization method described in reference 240 works with arbitrary arrangements of plants, but requires the user to specify how many plants are present in the image.²⁴ In practice, the majority of experiments use plants grown in grids, with a constant number of plants in each field of view; this geometric information can be exploited. An initial grid-based method for grouping contours was recently added to PlantCV—see Appendix D and the online documentation. I prefer

²⁴ Reference 240 describes how segmentation of individual leaves can be used to iteratively refine plant masks via updating of appearance models. The developers of the Phenotiki App authored that work, but have not yet been integrated leaf-level analysis into their application, which emphasizes rosette-level measurements and leaf counting [239].

to use circular regions of interest, which better fit the profile of the *A. thaliana* rosette. Using closer-fitting regions of interest delays “intrusion” of plants into their neighbor’s areas, and thus helps with segmentation of groups of larger plants, making this method flexible for long-running experiments. We could in principle continue to follow plants and individual leaves after they start to touch each other; methods for modeling object occlusion are a major area of video analysis research.

Segmentation of individual *A. thaliana* leaves is a challenging problem [241] but should eventually allow dynamic measurement of leaf size and shape, as discussed further below. Such measurements will require a significant amount of validation and calibration, as in Figure 3.7. Researchers are often interested in shoot biomass (fresh and/or dry weight) which is well-correlated with projected rosette area [238]. Explicitly modeling leaf occlusion and growth over time can improve the accuracy of such projected-area-based estimates for *A. thaliana* [242]. Simultaneously scanning and collecting dry weight measurements for individual leaves should allow further improved modeling.²⁵ Therefore leaf-level calibration of top-down photography setups should focus on measuring 1) correlation between projected area and flattened (scanned) area and 2) flattened area and biomass, in that order.

Minervini et al. did three useful analyses not considered here [239]. First, their side-by-side comparison indicated that the Raspberry Pi camera module (v1) performed about as well as a single-lens reflex camera for top-down imaging. Secondly, they calibrated a subset of their rosette-level shape metrics with caliper measurements. Thirdly, Minervini et al. demonstrated that they could detect an extreme coloration mutant phenotype based on average hue.

²⁵ Combining fresh weight measurements (for single leaves) with flat-bed scanning is more difficult, because the leaves begin drying out so rapidly after dissection. I thank V. Coneva for insightful discussions of this subject.

3.5.2 Stereo photography, distortion correction, and color accuracy

Color analysis, size calibration, and distortion correction all present major possibilities for improvement. A certain amount of white balance correction and size calibration can be achieved using objects (such as the white pot labels) in each field of view. We have photographed color standard cards at the start and end of experiments (Figure 3.1), but did not leave these in place (for more thorough correction over time), because of the expense and field-of-view space required. Concern about camera sensor drift was one reason for the use of automatic capture settings. Manually setting shutter speed and gains (analog, digital, red, and blue) is preferable for color-based applications,²⁶ including quantifying disease phenotypes such as water-soaking or chlorosis [243, 244]. Use of a large color-grid poster allows more sophisticated correction of color and both optical and perspective distortion [245, 246], beyond what can be done by joining and orthorectifying photos from overlapping fields of view into a single stitched image. Stereoscopic measurement using overlapping fields of view from multiple nearby cameras can also be used to gain depth information, to improve segmentation or for applications such as measuring circadian movements of leaves [246].

3.5.3 Metadata standards, monitoring, and experimental design

As noted above, management of metadata from these experiments has been based on folder and filename structure together with the Exif metadata embedded in each JPEG image file.²⁷ Automatic capture of additional information in machine-readable form would further improve the reusability of these data. The recent “minimal information about a plant phenotyping experiment” (MIAPPE) standard [247] is so far the most detailed schema developed for this purpose. The “e!DAL” software

²⁶ Unfortunately, as Dave Jones has noted, “It can be difficult to know what appropriate values might be for these attributes.” See <https://picamera.readthedocs.io/en/latest/recipes1.html>

²⁷ I captured photos in JPEG format (instead of PNG, which uses lossless compression) specifically to take advantage of this method for preserving metadata.

underlying the Plant Genomics and Phenomics Research Data Repository [248, 249] supports this scheme and has several other nice features, but it remains to be seen whether it will be deployed at other institutions.

Tracking metadata at the plant and individual leaf level could also be automated further to increase reliability and allow alternative designs. Automatic detection and readout of Quick Response (QR) codes is one option, and can facilitate repeated shuffling of plant positions throughout each experiment [245]. Such plant rearrangement can be desirable to reduce systematic effects from microclimate (differences in light intensity, temperature, airflow, etc.) but was avoided here for simplicity and to facilitate computational object tracking. Brien et al. [250] have argued that statistical control for the effects of such variation is preferable to plant rearrangement, even for systems that include conveyor belts for this purpose. Conveyor belt systems have a clear advantage for controlled watering during experiments, as for example in reference 234; carefully controlled watering is certainly possible in top-down imaging experiment, but is laborious, error-prone, and often unnecessary.

While my experience has been that Pi/Camera systems have been very reliable, more refined alerting methods will be advantageous when the inevitable hardware failures start. As image analysis methods mature, one can imagine setting standard analyses to be run automatically, for daily or real-time generation of results. Monitoring such results can help with identifying unexpected growth problems and may become convenient enough to be worthwhile for routine quality control of growth. The formal planning required for this type of experiment and the detailed records produced facilitate blinding and randomization and may provide an opportunity for improving statistical design of experiments.

3.5.4 Similar datasets, video analysis, and leaf phyllotactic order

Most work on leaf segmentation has focused on single timepoint snapshots [241], but integration of methods that use temporal information and appearance models should greatly improve the utility of analysis tools such as PlantCV. A small number of publicly available datasets (described below) include top-down photos from multiple experimental timepoints, which are a prerequisite for building and testing such tools. Data generated by the International Plant Phenotyping Network [239, 251] has been particularly useful;²⁸ this well-annotated data has stimulated great progress in methods for segmenting individual leaves [241]. The main difficulty in segmenting rosette leaves is that later leaves occlude early ones, resulting in “overlap” within images (Figure 3.8). Minervini et al. [252] have developed, shared, and used semiautomated interactive tools to efficiently generate “ground truth” color-label images for testing and validating leaf-level tools (Figure 3.8). Despite progress, the automated methods described thus far are not flexible enough for measurement of leaf heteroblasty, because they do not attempt to use or record the phyllotactic order of the leaves, resulting in loss of geometric information and developmental context.

At least two other publicly available datasets include top-down photos with multiple *A. thaliana* plants in each field of view. The first of these, the ‘MSU-PID’ dataset [254] includes four types of images, including depth sensor data. Yin et al. have also produced excellent work framing semiparametric leaf-level video analysis as a set of related segmentation, shape alignment, and tracking problems [255–257]. Cruz, Yin et al. [254] developed a semiautomated tool that allows tracking individual leaves, but loses all information about phyllotactic order; leaves are instead colored and numbered by tip angle relative to an arbitrary reference point.

²⁸ See <http://www.plant-phenotyping.org/datasets>

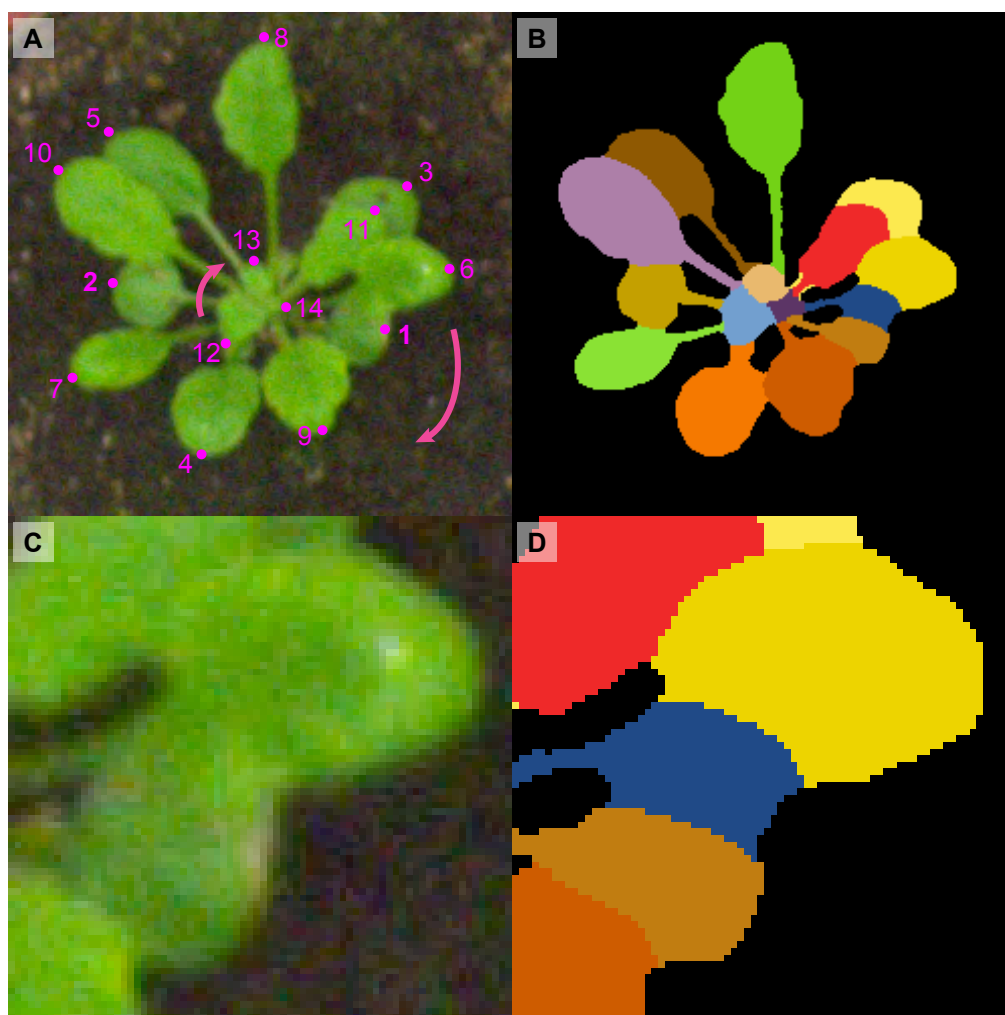


Figure 3.8: Example illustrating leaf-level segmentation and misannotation.

A. Cropped 250 by 250 pixel view of wild-type Col-0 plant #1 from the Leaf Segmentation Challenge dataset A1 [251]. Numbers indicate the order in which leaves emerged (phyllotactic order). Dots indicate approximate leaf tip positions and arrows indicate the direction in which leaves emerged (clockwise). Note occlusion of true leaf 3 by leaf 11. Cotyledons (embryonic leaves) are partly visible, adjacent to the tips of leaves 13 and 12.

B. Reference leaf annotation provided with dataset. This image represents the best human-curated segmentation currently available, and was used as an example in references 252, 253, and 241. Colors are unordered, as discussed in the text. True leaves 1 and 6 are missegmented in the annotation shown: an extra leaf object has been added (dark blue). The boundary between true leaf 1 and true leaf 6 is ambiguous in this timepoint, but can be clearly seen by comparison of other timepoints (not shown). This image represents the second timepoint in this series; geometric information from simpler rosette growth stages is not available for algorithmic processing. (Cf. Figure 3.4.)

C and D. Further-cropped view of leaves #1 and 6 in original image (75 by 75 pixels) and corresponding leaf annotations. Note brown pixels at the tip and near the base of the leaf blade.

By contrast, Bell and Dee [258] have appreciated the importance of leaf position and order. For their ‘Aberystwyth’ timecourse dataset, they annotated segmented leaves in phyllotactic order, using a red-to-yellow coloration method. Their method is somewhat complicated in that it incorporates uncertainty about order; independent annotation and estimates of error rates would increase the utility of this dataset further. Bell and Dee additionally used a flatbed scanner to measure area of dissected rosette leaves, similar to the method used here.

The ‘Aberystwyth’ dataset was collected with a conveyor belt system, and other *A. thaliana* datasets from conveyor belt systems (with one plant per field of view) are available [259, 260]. We suggest that fixed-camera systems will be most cost-effective for simple visible-light imaging. Mobile camera, conveyor belt and manually-loaded systems will remain useful for systems with more sophisticated sensors or optical systems. Tools that can combine information from both A) high-throughput measurement of large numbers of plants and B) higher-resolution and/or better isolated measurements of fewer plants will likely yield the most useful models of *A. thaliana* growth. Dissecting and scanning of leaves will remain an important isolation method, to simplify processing and generate leaf area validation data. We anticipate that the availability of this type of data will dramatically increase in the near future, and be creatively combined with other types of measurements to address diverse biological questions.

The *A. thaliana* community seems well-positioned to benefit from statistical machine learning methods once sufficient high-quality training and test data are amassed [202]. Plant biologists have an important role to play in this work, both in optimizing data collection and applying expert knowledge to annotation of images [261, 262]. Using human perception of leaf phyllotactic order (Figure 3.8) is one example of how biological knowledge can be used to guide annotation.²⁹ It is

²⁹ My description of the missegmentation in Figure 3.8 is not intended as a criticism of any of the relevant papers cited. As emphasized above, I think release of this annotated dataset has driven great progress. I have alerted S. Tsaftaris and H. Scharr of the issue illustrated in the figure.

difficult to know just how much test data will be needed for different types of inference, and how robust different methods will be to errors in labeling of data for training and testing. Carefully coordinated experiments suggest that inter-laboratory variation will be substantial [263]. Out-of-sample error will always be a large and difficult-to-assess concern, and thus there will always be a need for “a human in the loop”. Multiple groups have convincingly argued that statistical approaches (based, for example, on neural network models) will eventually be the dominant approach to rosette image analysis, as they have in the field of computer vision more broadly [202, 262]. However, uncritical application of these methods can lead to misinterpretation, adding unnecessary high-dimensional complexity and compounding errors over the long term, and thus creating substantial technical debt [264]. Semiautomated heuristic methods such as the ones emphasized above will remain important [262], not least because their interpretability aids experimenters in optimizing image capture conditions.

3.5.5 Historical note and personal reflections (part 1)

Despite widespread enthusiasm for automated plant phenotyping, I was initially reluctant to start the experimentation described here. I expected that multiple descriptions of Raspberry Pi imaging would be published in short order, and I did not want to reinvent the wheel. Publications describing this approach (notably reference 239) have indeed started to appear, and at least one company has started to offer commercial support.³⁰ It may be that camera systems will become a standard component of controlled-environment growth chambers. Alternatively, the technological trends emphasized here may drive increased customization of hardware for environmental control and monitoring; researchers and hobbyists may increasingly construct their own growth chambers [265–267]; particularly if light-emitting diode grow light costs continue to drop. The needs

³⁰See <https://cropscore.com>

of different experimenters vary quite a bit, so diverse approaches will always be required and independent implementations that enable comparison will likely have great value. Concrete details and hands-on experience are both extremely valuable.

Video systems have been used for analysis of *A. thaliana* growth for at least twenty-five years [268]. Why has this type of analysis not become more common? Cost of cameras is likely one factor. Relatively low-cost digital cameras have been used since at least 2009 [269], but Raspberry Pi camera modules represent a substantial improvement in this regard. The effort required to program specialized cameras may be another factor, and here the fact that the Raspberry Pi is a general purpose computer may help with usability. My preferred data acquisition method (Appendix C) requires developing familiarity with a command-line interface, but other methods for configuration (including via web browser; reference 239) are available. Data processing is clearly another bottleneck, and involves similar tradeoffs in terms of degree of automation, interactivity, and user-friendliness. ImageJ is an obvious framework to build on for graphical user interfaces, including for rosette analysis [270–273]. Interactive systems take advantage of the powerful human visual system and I have found ImageJ extremely useful, as described above, but interactive operation is not an unalloyed good—see Appendix D.

This work represents incremental progress, not a definitive standardized methodology that will find immediate widespread use. It should be frankly admitted that most important recent work on leaf heteroblasty (e.g. references 132, 133, and 155) have been based on careful observation and manual measurements rather than time-lapse imaging; related image analysis has mostly been based on scans of flat leaves [89], including one recent exceptional topdown-photo-augmented study of *A. thaliana* [274]. Nevertheless, I find it plausible that before too long many groups will find it more convenient than not to capture photos, particularly because they enable “rewinding of time” for exploratory measurement of traits that were not of *a priori* interest.

In describing computational work it is difficult to find a balance between focusing on high-level concepts and including sufficient low-level detail for the work to be useful. This tension is very relevant to plant image analysis, and has been described well by Pridmore et al. [275]. The interdisciplinary nature of this project has created both challenges and opportunities. My perspective on plant image analysis may seem unusual because my interest in time-lapse imaging was so strongly motivated by a desire to quickly and easily figure out the phyllotactic order of rosette leaves. Effort may have been better invested in rosette-level analysis and measurement rather than at the much more challenging leaf level; hopefully some of the choices I made about analysis software will yield long-term benefits (Appendix D). I have tried to show here how applying expert knowledge of rosette morphology may prove useful for future image processing algorithms. I reflect further on some of these issues in section 4.5, as part of describing the broader scientific context in which this project fits.

Chapter 4

Prospects and conclusion

The work described here leveraged recent improvements to two automated methods for understanding development at the levels of 1) protein-DNA interaction and 2) whole-plant growth. These efforts yielded a model for control of the key conserved regulator *AGO7* (Figure 2.17) and practical improvements to low-cost imaging methodology. This work represents progress in linking molecular events to the development of whole organisms, and I have explored how improvements in automated measurement might drive further progress. Vegetative growth of *A. thaliana* will continue to be an important experimental system.

As described in the introduction, forward genetic studies of heterochronic mutants, particularly in *C. elegans* and *A. thaliana*, have been enormously productive. The initial curiosity-driven studies were designed to clarify the logic of development and evolution [134, 276–279], and yielded insights about post-transcriptional regulation that could not have been directly anticipated. *MIRNA* genes may have long eluded discovery because of redundancy, small genetic footprint, subtle phenotypic effects, and/or absence from models such as budding yeast [12]. Identification of small RNA functions is a case in which complementary studies of plant and animal systems accelerated discovery of a broad principle. In terms of our understanding of fundamental molecular mechanisms, these studies drove a paradigm shift at least as large as the celebrated screens for

disrupted patterning in *Drosophila* larvae [280, 281] and studies of homeotic mutants in flies and plants [282–284]. Analysis of heterochronic and homeotic transformations uncovered analogous control of gene expression in time and in space, respectively [285]. Overall, studies of small RNAs and TFs reveal commonality, in which master regulators drive progressive elaboration of patterning. Conservation of regulators and their binding sites preserves some of the history of how developmental mechanisms evolved [2, 286].

Chapter 2 presented evidence for a new direct connection between three deeply conserved small RNA-TF circuits. Possible functions of this direct connection are still largely unclear, but fit the theme of dense interconnection between AGO-small RNA-TF circuits. The next three sections provide broader context by outlining other recent progress and outstanding questions.

4.1 How do *TAS3* tasiRNAs function in patterning?

As emphasized in the introduction, special aspects of *TAS3* tasiRNA biogenesis may be important for their patterning function. This proposed function was a major reason we sought to understand upstream control of polar *AGO7* transcription. The predominant model in the literature can be summarized as follows:

1. Cell-layer-limited AGO7-miR390 action sets up localized production of *TAS3* tasiRNAs.
2. tasiRNAs move between cells, establishing an adaxial-abaxial gradient in developing leaf primordia.
3. tasiRNA targeting creates a sharp boundary of *ARF* protein accumulation, which contributes to robust maintenance of polarity.

This model fits the steps of the morphogen concept: 1) local chemical signal production, 2) gradient formation, and 3) direct gradient interpretation. Definitions of the word “morphogen” vary: most involve movement or graded distribution, and some require readout of at least two thresholds (per the “French flag model” described in reference 287), as is common in animal development [288, 289]. The word itself was introduced by Alan Turing [290], but is used broadly—it does not only refer to the specific two-component chemical model introduced and mathematically modeled in his paper.³¹ Most discussions of the possible relevance of morphogen gradient concept in plants have focused on auxin signaling [292, 293], but similarities between *TAS3* tasiRNA function and morphogen action in animal systems have been noted [207].

Chapter 2 focused on the first step in the model outlined above (polarization via restricted transcription), but all three must be considered. Unfortunately, our understanding is built mostly on indirect evidence.

1. AGO7-miR390 complexes have not been physically localized across leaves; patterns of action have been inferred based on transcriptional reporters [50, 117].
2. *TAS3* tasiRNA movement has not been measured directly; movement has been inferred based on qualitative discordance between patterns of RNA accumulation and reporter gene signal [117, 118].
3. The hypothesized sharp boundary of ARF protein levels has not been measured directly; this concept was postulated based on a mathematical model that depends on the unproven assumption that AGO1-tasiRNA complexes function primarily in noncatalytic modes [294].

As discussed below, the evidence for noncatalytic (“translational repression”) functions for

³¹ Turing was concerned with the spontaneous stochastic generation of *initial* patterns. Despite suggestive examples, the direct relevance of Turing patterns *per se* to living systems remains to be demonstrated [289]. The ideas of position-dependent cell differentiation and mobile signals have an older history—see reference 291.

plant AGOs is mixed. Additionally, cleavage of *ARF* transcripts is readily detected [48, 204], indicating some degree of catalytic turnover.

Overall, it is unclear how fast and how far *TAS3* tasiRNAs move, and how they affect *ARF* mRNA and protein levels, largely because of technical challenges associated with measuring localization of protein-RNA complexes. No mechanism for active transport of small RNAs has been demonstrated and the extent of passive diffusion is also unknown. Consequently, short-distance movement of small RNAs has been more difficult to demonstrate than long-distance movement. It may be possible to formally prove small RNA movement within leaves using mosaic plants with periclinal chimerism or local sectoring [295]. Alternatively, direct imaging of tasiRNA movement between cells may become possible, and would be very useful for modeling and mechanistic studies. Ultimately, methods for precise manipulation of tasiRNA gradients will be necessary if we are to truly test their proposed functions.

An intermediate area in which the juxtaposition of abaxial and adaxial domains leads to expression *WUSCHEL*-related homeobox genes was recently discovered [296]. The presence of this intermediate zone, which appears critical for outgrowth from “marginal meristems” [165], can be considered evidence against the idea of a sharp dorsoventral boundary in expanding leaves. Recent data suggest that *ARF* repressors directly contribute to establishment of this middle domain, together with the *ARF* activator *MONOPTEROS* [297]. Furthermore, the question of whether *TAS3* tasiRNAs create a sharp *ARF* protein boundary is less important than the downstream question of how *ARF* levels control transcriptional changes. A continuous gradient of *ARF* protein levels could certainly contribute to precise definition of two or more transcriptional states, possibly through a mechanism similar to readout of the nuclear Dorsal gradient in early *Drosophila* embryos [288]. Understanding *ARF* repressor function will require considering both timing and patterning.

4.2 Inputs and outputs of SPL and ARF circuits

As noted above, our knowledge of plant TF DNA binding specificity has expanded dramatically in the past few years. Nevertheless, measurements of genome-wide binding and transcriptional effects should provide several types of insight into ARF and SPL function.

Understanding how ARF3 and ARF4 targets act at the cellular level should clarify how polarity establishment and shape changes are connected, and thus clarify the function of AGO7. ARF3 action in flowers and floral meristems was recently analyzed with RNA-seq and ChIP-seq [120]. This analysis did not consider regulation via the *TAS3* pathway, but nevertheless revealed considerable complexity, including feedback into all levels of auxin signaling and related interactions with TFs such as INDEHISCENT and REPLUMLESS. However, global expression analysis of *ago7* mutants has thus far identified very few misexpressed genes [45, 175]. It is not possible, with our current level of understanding, to say whether downward leaf-curling in *ago7* mutants is a consequence of subtle polarity defects (abaxialization), premature shape changes, or some combination of the two. The heterochronic shift interpretation was initially favored because of related changes in trichome and leaf serration phenotypes [113]. Interpretation in terms of polarity came later, once the genetic connection to ARF3 and ARF4 was clearer [128, 175, 205, 206] and polar expression patterns were identified [117, 118, 175]. The idea that limited spatial action of AGO7-miR390 complexes is functionally important for leaf curling phenotypes (Point 1 in the model described above) has been an *inference* rather than a directly tested hypothesis. Tentative evidence from the work described here suggests that AGO7 spatial transcription pattern and genetic function in leaf curling are separable in *A. thaliana* (Figure 2.17), and thus raises questions about the relevance the mobile *TAS3* tasiRNA model. Differences in phenotype interpretation have also arisen in the intensive studies of AGO7 in maize [208, 209, 298], perhaps due to effects from genetic background.

It is plausible that polarity specification and heteroblasty are inextricably linked. If this is the case, it should be useful to define this link in cellular terms in plants with diverse leaf morphologies.

A related issue is that we have little information about how SPL and ARF activities converge to affect cellular processes. Identifying SPL targets is essential, because studies of in diverse plants have shown that they are the primary TFs that drive heteroblasty [88]; other groups of TFs such as TCPs and ARFs appear to function mainly as modulators of the developmental outputs resulting from increasing SPL levels. Identifying direct targets should help us evaluate the abstract threshold model for heteroblastic changes [128, 134; see Figure 4.1] and could reveal regulatory logic for activation of *AGO7* by SPLs.

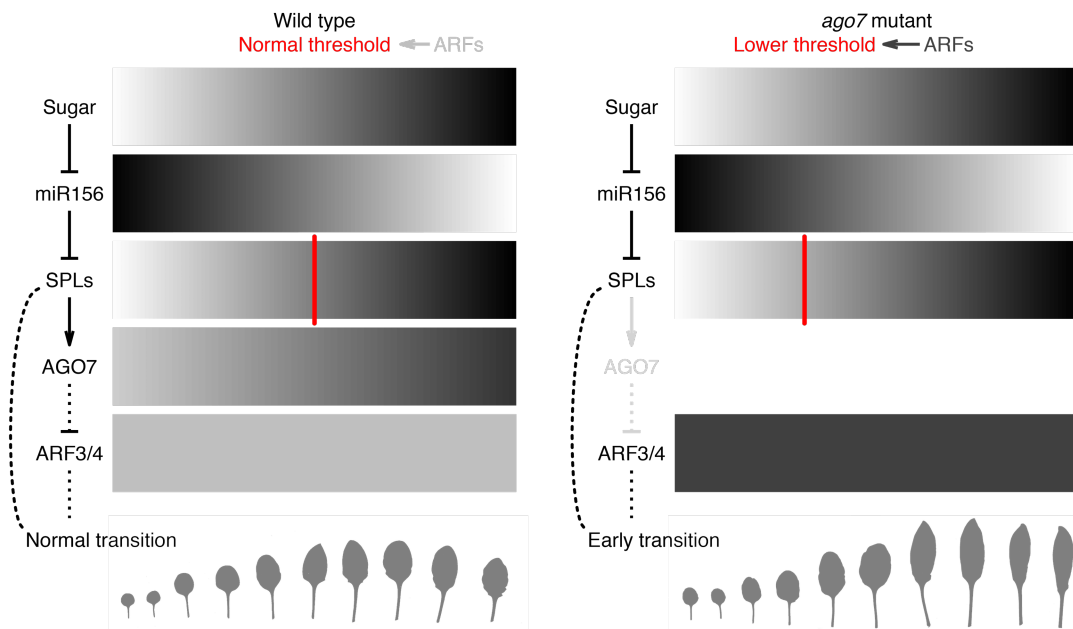


Figure 4.1: Revised model for control of heteroblasty in *A. thaliana*, incorporating the likely activation of *AGO7* by SPLs. SPL levels gradually increase until they reach a hypothetical activity threshold, controlled in part by ARFs, and trigger changes in leaf characteristics. When the *AGO7* pathway is disrupted, ARF levels go up, lowering the threshold for transition such that it is reached earlier. This perturbation results in leaves that are thinner, longer, and more curled in *ago7* mutants. Indirect repression of ARFs via *AGO7* and *TAS3* tasiRNAs may have significance for feedback control of SPL activity.

ARFs and SPLs presumably share downstream targets, perhaps separated by one or two layers of regulation (Figure 4.1). While SPLs and ARFs both promote phase change, they may have opposing transcriptional effects, given that SPLs generally function as activators and ARF3 and ARF4 as repressors [299]. Understanding how miR156-SPL and auxin pathways intersect in control of leaf primordium initiation is a related and important goal. The role of auxin depletion in primordium initiation is well established; emergence of leaves at auxin maxima providing a convincing model for spiral phyllotaxy [300, 301]. Evidence from live-cell imaging increasingly suggests that local auxin depletion also prepatterns the abaxial-adaxial axis prior to leaf primordium emergence [297, 302–305], counter to the mobile signal model described in section 1.2.1.

4.3 Methods for dissecting enhancer regions in plants

It is fair to say that our knowledge of transcription in plants lags behind our understanding of yeast and animal systems, and that is not yet clear which concepts translate well across kingdoms. Comparison to patterning of the *Drosophila* embryo, in which gradients of maternally-derived factors are translated into precise expression boundaries and repeated metameric segments may be instructive. Studies of anterior-posterior specification revealed how a cluster of genes encoding homeodomain proteins have analogous patterning functions in all bilaterians [286].

A central concept in transcriptional control of animal body patterning is that clusters of TF binding sites are organized in modular enhancer units [306]. Application of this concept is surprisingly difficult in plants. Enhancers are clearly relevant at a broad conceptual level [307], but major uncertainties remain about action at a distance, which seems to be rare in plants³² but is very

³² I am aware of only one well-studied plant example: enhancement of *teosinte branched 1* transcription by a distantly inserted transposon [308].

common in animals [309]. It is possible that this reflects fundamental differences in “promoter architecture” between plants and animals³³, possibly associated with differences in insulator gene family content [310]. Yeast promoters certainly seem compact [311], a fact explicitly considered by our collaborators during development of their initial Y1H strategy [179]. Genome-wide studies of chromatin packing may shed light on this issue [310, 312], including possible peculiarities associated with the recently reduced size of the *A. thaliana* genome [313, 314].

It will be methodologically useful to assess the extent to which Y1H methods can identify binding sites that are *weak but functionally significant*, overcoming an important type of ascertainment bias. It is increasingly recognized that such low-affinity sites, particularly in clusters, can both reinforce requirements for cooperativity and promote tissue-specificity [315]. As noted in several places, the Y1H assay is a powerful method for circumventing genetic redundancy, both between structurally related TFs and among physically linked binding sites of unrelated TFs. Identifying such redundancy remains challenging, but should eventually become predictable, which will be important because redundant sets of enhancers are increasingly thought to contribute to developmental and/or evolutionary robustness [306, 316–318]. Similar to animal miRNA with presumed fine-tuning roles, such enhancers can be difficult to identify because they are individually dispensable under favorable growth conditions [319–321]. It remains possible that the SPL and TCP binding sites described above have roles in fine tuning or canalization. Testing function under temperature stress may be particularly interesting because of the influence of the plant “thermal clock” on heteroblasty [322]. Thorough time-course measurement in sensitized genetic backgrounds and in diverse conditions will remain important.

CRISPR mutagenesis should enable much more thorough dissection of elements upstream of *AGO7* and other genes in their natural genomic context. Saturating mutagenesis is feasible

³³ Animal molecular geneticists generally use the word “promoter” to refer specifically to the proximal region bound by core transcriptional machinery.

[323] and targeted mutagenesis (based, for example, on Y1H and/or PWM scan results) avoids variability introduced by position effects associated with random insertion of transgenes.³⁴ Careful use of sensitized backgrounds and crossing designs can increase the power of this approach further, including in crops [324]. Homologous-recombination-based replacement of genes and promoter elements will reduce technical challenges that have long plagued studies of TF binding and transcriptional silencing [325]. CRISPR mutagenesis can be productively combined with both chromatin immunoprecipitation (to biochemically verify altered binding by specific factors) and Y1H screening (for prioritizing specific *cis*-element-dependent hits).³⁵ Successful application of a related transgene-based approach by Li et al. [326] serves as an instructive model.

MIR156 promoters will be particularly interesting targets for this type of analysis because of their central role in control of timing. The molecular events necessary for repression of *MIR156* genes are starting to become clear [6, 327–331]. However, this gradual repression but have not yet been directly linked to transduction of signals such as or temperature stress [332] or sugars (per the model illustrated in Figure 4.1; references 132, 133, and 333). Improved understanding of *MIR156* should help connect models of developmental control to physiology, circadian regulation, and metabolism [334], including likely connections to trehalose-6-phosphate signaling [335, 336].

The general approaches described above could also be productively applied to other *AGO* genes, including dissection of highly localized expression in reproductive tissues [26, 40, 108, 337]. Most of these AGOs are presumed to function in germline defense via RNA-directed DNA methylation directed by mobile small RNAs, but mechanistic differences among them are largely obscure; these AGOs have similar structures and siRNA binding preferences (Figure 1.4).

³⁴ I suspect that hygromycin selection was a much larger source of developmental variability than position effects in my experiments. Avoiding antibiotic selection by characterizing a sufficiently large number of stable homozygous lines (generated with CRISPR-based strategies or otherwise) is clearly preferable, when feasible.

³⁵ Large-scale Y1H assays also allow initial assessment of secondary effects: demonstration that mutation of a bait sequence affects only a small number of TF hits can provide preliminary evidence that the effect of a corresponding *in vivo* mutation is relatively specific. See for example reference 326.

4.4 Molecular functions of AGOs in development

AGO function depends on *cis* elements and physical protein structure. This dissertation has focused on upstream regulators of AGOs, but there are also many open questions regarding molecular mechanisms of AGO action [6, 338, 339]. Particularly relevant to work described here are questions about possible *noncatalytic* action of plant AGO complexes and the related issue of how specific complexes trigger phased siRNA production.

Initial studies suggested that plant miRNA function by cleaving their targets and that animal miRNA inhibit translation. Further work has indicated that animal miRNA can promote transcript turnover by deadenylation and decapping but may not significantly inhibit translation initiation or elongation *per se* [340–344]. Simultaneously, reports of apparent translational repression have complicated our understanding of targeting in plants. The most notable examples are copper-related miR398 targets [345–348] and miR156-targeted SPLs [147, 349, 350].³⁶ Initial observations were puzzling, because both miRNA are highly complementary to their targets, and cleavage products can be detected for both. The “seed”-type matches predominant in animal miRNA targeting do not appear to function for either cleavage or translational repression in plants [351]. However, discovery that the Glycine-Tryptophan motif (GW) protein SUO is required for translation inhibition appears to have revealed a mechanistic similarity to target repression in animal systems [350]. SUO and other plant GW proteins may have eluded biochemical detection because of differences in structure and affinity [352]. Further study of SUO may help clarify noncatalytic action of AGOs, with important implications for two hypothetical properties: sharpening of boundaries by mobile small RNAs [294] and rapid reversal of repression [353, 354].

³⁶ AP2 targets of miR172 were suggested to be translationally repressed, but this effect seems to have been confounded by feedback regulation [148].

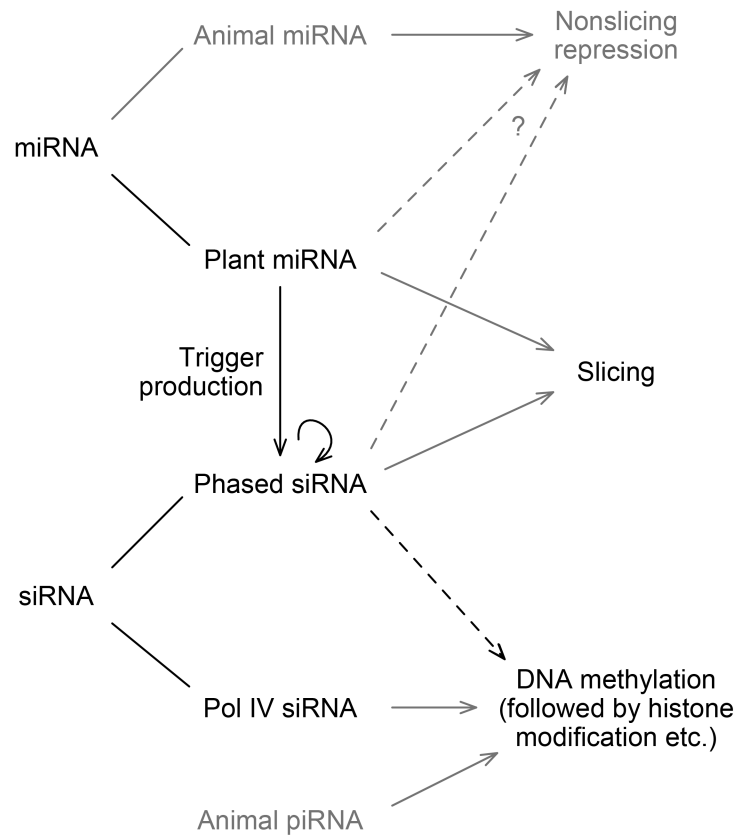


Figure 4.2: Schematic illustrating possible small RNA targeting outcomes. As discussed in section 1.1.1, the predominant targeting modes of plant and animal miRNA were initially thought to differ. More recent models described in the text indicate greater similarity, particularly in terms of endonucleolytic cleavage vs. noncatalytic action. Genome-wide studies of diverse tissues, timepoints, and conditions will be required to understand the prevalence of these modes. Recent work has also clarified determinants of amplification (AGO-triggered secondary siRNA biogenesis), apparent threshold-based connections between amplification and transcriptional silencing, and competition between routing pathways. See also Figure 1.3 and Figure 1.5 for schematic molecular representations.

Determinants of slicing vs. nonslicing action are also relevant to phased siRNA production, including from *TAS3* transcripts. Understanding this biogenesis mechanism is essential because it appears to be an important determinant of small RNA mobility [355]. Recent work has shown that slicing is not required for tasiRNA biogenesis [356], consistent with the possibility that AGO-miRNA complexes may recruit the stabilizing factor SGS3 [357]. SGS3 and/or AGO complexes

may also recruit RDR6, either through protein-protein interactions or by passively enabling it to recognize transcript features such as the lack of a poly(A) tail [358, 359]. In both scenarios, subcellular localization could be an important determinant. AGO7, SGS3, and RDR6 appear to be present in specific cytoplasmic membrane-associated bodies [360, 361], which may be associated with ribosomes [362, 363].

In spite of this progress, it remains unclear why a few complexes trigger biogenesis and others do not. One clear point, learned mostly from studies of the AGO7-miR390 complex, is that dual targeting is often sufficient [49, 356]. A second point is that small RNA duplex structure can affect competence of AGO1, possibly because complexes are “programmed” during loading or because specific long mature miRNA strands sterically alter AGO1 conformation [364–366].³⁷ AGO-intrinsic features must also be important, given that programming AGO2 with 5'-nucleotide-swapped variants of the same miRNA does not suffice to trigger siRNA production [217]. Further structural, biochemical, and high-resolution microscopy work will be necessary to clarify mechanisms [338, 339]. Grass anthers could be a useful system for such work, because they contain a vast array of cell-type and stage-specific phased siRNAs [27].

Similar studies may also clarify triggers for amplification of antiviral silencing, which has similar genetic requirements, as illustrated in Figure 1.3. Sensing molecular signatures of invasion may use a threshold mechanism, in which one Dicer becomes saturated, allowing a second to produce 22-nt siRNAs that confer amplification-triggering properties to AGO complexes. Recent data suggest that a similar mechanism contributes to initiation of *de novo* DNA methylation via AGO6 after ectopic or endogenous RNA levels reach a certain threshold [370, 371]. Similarly, miRNA targeting can trigger secondary siRNA biogenesis and silencing when derepressed transposon transcripts reach

³⁷ Little is known about plant AGO loading complexes beyond a requirement for cyclophilin 40 and heat-shock protein 90 [367]. As emphasized in Figure 1.4, 5' nucleotide is a small RNA loading determinant, controlled through interaction with AGO MID domains [368]. Loading of AGO7 also depends on other parts of the miR390/miR390* duplex; the mechanism for this selectivity is unclear [369].

high levels [372], and there is some evidence that miRNA can trigger methylation directly [54, 373]. These data suggest a continuous spectrum from post-transcriptional to transcriptional silencing. Understanding routing and threshold mechanisms may help researchers evaluate of the possibility that large networks of phased siRNAs “buffer” the rapid expansion and divergence of gene families such as immune receptors [47, 374].

Overall, questions remain about several fundamental aspects of AGO action, including the interrelated issues of small RNA loading, protein-protein interactions, RNA target spectrum, and subcellular localization. Understanding transcriptional specialization of *AGO* genes is a prerequisite for deep understanding of all of these processes, and will therefore remain an important research priority.

4.5 Personal reflections (part 2) and coda

The purpose of this dissertation is to document my accomplishments in the fields of computational and systems biology, so it seems appropriate to record some thoughts on those subjects. Computation has become so pervasive that the phrase “computational biology” feels redundant, though drops in costs that have made computer sensors effectively disposable are a new development. Systems biology is a bit more specialized, because of the direct analogy to the discipline of systems engineering [375]. One description that I like [376] emphasizes comprehensiveness of component identification and measurement.³⁸ Our “parts lists” are far from sufficient for most modeling purposes, especially for approaches that seek to consider multiple scales. Identifying

³⁸ The interview cited provides an accessible overview of the problem of understanding genome-wide enhancer function (described above), with an emphasis on research in fruit flies and sea squirts.

targets of conserved plant *MIRNA* families (Table 1.1) is one bright spot where our catalog of (abstract) miRNA-target pairs seems near-complete, providing a framework for analyzing divergence of the relevant genes throughout the plant kingdom. I hope it has been clear how the two main methodologies used in this thesis advance us toward comprehensive profiling of vegetative growth and identification of direct upstream regulators.

Functional dissection of promoter regions has been a powerful method for learning fundamental facts about biology, including in *A. thaliana*. Some of my favorite examples from the plant literature include references 377, 378, and 326. I found it challenging to abstract principles for effective “promoter bashing” from these papers, probably because of some combination of misallocated effort and fundamental limitations in our current understanding of transcription. Molecular cloning was a major bottleneck throughout this project, unfortunately. DNA synthesis technology is improving rapidly, but is still rather expensive for regions such as the *AGO7* promoter that are repetitive and contain long homopolymer stretches. I am optimistic that massively-parallel reporter gene assays will find frequent use in plant cell research, as they have in the community working in animal systems [379]. Other logistical challenges have included recovering the large number of independent transformants required for conclusive transgenic experiments and limitations of tools for management and integration of image data.

A central idea driving much work described here and elsewhere is that small RNAs have nontrivial unique properties, possibly related to movement, speed, or reversible action. The nature of these properties makes it difficult to directly test this concept: we cannot falsify the alternative possibility that gene regulation by small RNAs emerged by chance and later become “locked in” as a constraint on further evolution. Researchers have amassed many examples of small RNA regulatory interactions in diverse lineages and contexts. Deriving coherent theory from this body of knowledge will require mathematical approaches at appropriate levels of abstraction. This dissertation has

illustrated that uncertainty about how dynamic TF complexes interact with DNA elements in the context of chromatin is a limiting factor for our understanding of silencing factor specialization, as it is for many other areas of biology.

Appendix A

Roles and programming of *Arabidopsis* ARGONAUTE proteins during *Turnip* *mosaic virus* infection

A.1 Preface

The body of this appendix is published work, cited above as reference 194. This preface describes my contributions to the work described (lead by Hernan Garcia-Ruiz), and its relation to the rest of this thesis. As schematically illustrated in Figure 7, this paper focuses primarily on AGO2, AGO10, and AGO1, including differing requirements for A) restriction of infection in inoculated rosette leaves and B) systemic movement to noninoculated cauline leaves and flowers.

I contributed to the drafting and revision of the manuscript, and also made two contributions to the experimental work. First, I assisted with the genetic analysis (Figures 1 and 2), both in generating and genotyping *ago* multimutant strains and in characterizing their susceptibility phenotypes. Secondly, I made tagged 3xHA-AGO10 constructs (mutant and catalytically normal forms of a

large genomic fragment; see Materials and Methods section), transformed plants, and propagated the resulting lines. These lines were used for co-immunoprecipitation experiments (Figure 5), which confirmed that AGO10 preferentially binds miR166 [100] and also showed that AGO10 can bind virus-derived small RNAs.

We included *ago10* mutants in our genetic analysis, but the manuscript does not emphasize their developmental defects. Shoot apical meristem “*pinhead/zwille*” defects are pronounced in a Landsberg *erecta* genetic background [380, 381], but are mild in Col-0 [382–384], which was used for this study.³⁹ Col-0 T-DNA lines have petal number defects and occasional stem fasciation, and HA-AGO constructs complement these phenotypes. Catalytic mutant construct lines showed upward-curved leaves at high frequency in both wild-type and mutant backgrounds (11 of 54 and 7 of 11 primary transformants, respectively). A similar phenotype is observed when AGO10 misexpression is driven by the abaxial *FIL* promoter [387], and presumably reflects adaxialization caused by increased *HD-ZIP* mRNA levels (Figure A.1).⁴⁰

Use of suppressor-deficient viruses have been key for uncovering normal roles of host silencing factors, including in this work and the previous study it built on [389]. Challenging *ago7* mutants with suppressor deficient turnip mosaic virus (TuMV AS9) indicates that AGO7 makes little or no direct contribution to normal TuMV resistance.⁴¹ By contrast, AGO7 appears to make a larger

³⁹ The basis for the effect of genetic background on *ago10* mutant phenotype penetrance is largely unclear. The cyclophilin 40 gene *SQUINT* emerged as a candidate modifier of *AGO10* action based on QTL mapping [384] and affects development via its important role in AGO loading [367, 385]. However, *SQUINT* accounts for only a small proportion of the variability in frequency (penetrance) of *pinhead*-type mutant meristem defects between a Landsberg (or Fe-1) background vs. Col-0 [384]. The mutant *erecta* allele (retained in the reference *Ler* background to promote compact growth) was an obvious candidate because of its large pleiotropic effects on vegetative development [386], but does not appear to directly modify *AGO10* genetic function [384].

⁴⁰ *AGO10* promotes *HD-ZIP* transcript accumulation, by sequestering miR166 away from AGO1 [100]. This molecular observation, however, does not imply that AGO10 cannot cleave *HD-ZIP* transcripts at some rate. It is therefore possible that substitution of AGO10 catalytic residues may slightly increase *HD-ZIP* mRNA accumulation. As alluded to above, it seems plausible that AGO10 enhances small RNA turnover through recruitment of endonucleases [101, 388].

⁴¹ See Figure 1 and Table 1. Mutant *ago7* alleles do not enhance the susceptibility phenotype of *ago2* mutants (Figure 2A, panel I).

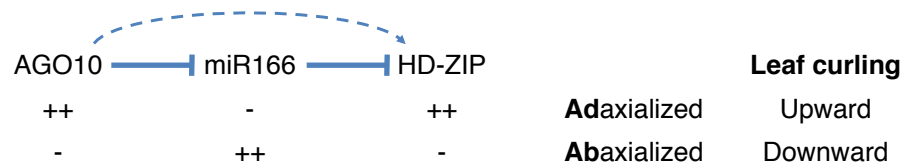


Figure A.1: Summary of polarity phenotypes resulting from altered *AGO10* levels. Increased AGO10 levels in transgenic lines increase sequestration of miR166, and thus have a net positive effect on HD-ZIP levels. In the absence of AGO10 protein, miR166 activity (in complex with AGO1) increases, resulting in decreased HD-ZIP levels and abaxialization.

contribution to resistance against suppressor-deficient turnip crinkle virus [390]. The mechanistic basis for this role is unknown, and could depend on specialized expression or structural features. As noted in section 4.4, it is not known which AGO proteins may trigger amplification of viral RNA silencing responses, nor if the resulting secondary siRNAs are loaded into specific AGOs.⁴² Both domain swap transgenes (as in reference 383) and promoter swap transgenes may be useful for comparative analysis.

Our genetic analysis only identified a small role for AGO5, which can bind small RNAs derived from cucumber mosaic virus [382] and also functions in defense against potato virus X [392]. AGO5 promotes megagametogenesis through largely unclear mechanisms that may be related to its tissue-specific expression patterns [337, 393]. *AGO5* genes are most closely related to *AGO1* and *AGO10* (Figure 1.4). As with AGO10, it is not clear if antiviral action is a primary molecular function of AGO5 (under long-term positive selection) or a side effect from more important roles in endogenous gene regulation.

As outlined in the Discussion section of this paper, understanding the relative contributions of different AGOs is complicated by the potential for indirect effects on endogenous gene regulation. Indirect effects via perturbation of non-silencing defense pathways and/or homeostatic silencing cross-regulation seem particularly likely. Reducing AGO1 activity can increase AGO2 levels,

⁴² An unusual class of endogenous “virus-activated siRNAs” appear to be specifically loaded into AGO2 [391].

due to relief of repression by miR403. The resulting upregulation of AGO2 been proposed as a mechanism for sensing inhibition of silencing [394]. Similarly, reductions in the activity of AGO10 and/or other AGOs may lead to disrupted miR168 targeting of *AGO1* transcripts [383]. Activities of AGO1 and AGO10 appear synergistic in some contexts [103, 387] and antagonistic in others [100, 383]. Disentangling possible indirect effects from cross-regulation and competition for small RNA loading is an important goal for future work [33].

RESEARCH ARTICLE

Roles and Programming of Arabidopsis ARGONAUTE Proteins during *Turnip Mosaic Virus* Infection

Hernan Garcia-Ruiz^{1,2#a}, Alberto Carbonell^{1,2}, J. Steen Hoyer^{1,2,3}, Noah Fahlgren^{1,2}, Kerrigan B. Gilbert¹, Atsushi Takeda^{2#b}, Annalisa Giampetruzzi^{2#c}, Mayra T. Garcia Ruiz^{1#d}, Michaela G. McGinn^{1#e}, Nicholas Lowery^{2#f}, Maria T. Martinez Baladejo¹, James C. Carrington^{1,2*}

1 Donald Danforth Plant Science Center, St. Louis, Missouri, United States of America, **2** Center for Genome Research and Biocomputing, Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, United States of America, **3** Computational and Systems Biology Program, Washington University in St. Louis, St. Louis, Missouri, United States of America

#a Current address: Department of Plant Pathology and Nebraska Center for Virology, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America

#b Current address: Department of Biotechnology, Graduate School of Life Sciences, Ritsumeikan University, Noji-Higashi, Kusatsu, Shiga, Japan

#c Current address: Istituto per la Protezione Sostenibile de la Pianta del CNR, Bari, Italy

#d Current address: Recursos Genéticos y Productividad, Fisiología Vegetal, Colegio de Postgraduados, Montecillo, México

#e Current address: Molecular and Cellular Biology Program, School of Biological Sciences, Illinois State University, Normal, Illinois, United States of America

#f Current address: Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom

* jcarrington@danforthcenter.org



OPEN ACCESS

Citation: Garcia-Ruiz H, Carbonell A, Hoyer JS, Fahlgren N, Gilbert KB, Takeda A, et al. (2015) Roles and Programming of Arabidopsis ARGONAUTE Proteins during *Turnip Mosaic Virus* Infection. PLoS Pathog 11(3): e1004755. doi:10.1371/journal.ppat.1004755

Editor: Biao Ding, The Ohio State University, UNITED STATES

Received: December 9, 2014

Accepted: February 19, 2015

Published: March 25, 2015

Copyright: © 2015 Garcia-Ruiz et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and Supporting Information Files. Sequence data from this article can be found in the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession number GSE64911.

Funding: National Institutes of Health (www.nih.gov) grant AI43288 to JCC. National Science Foundation (www.nsf.gov) grant MCB-0956526 to JCC. Helen Hay Whitney (www.hwhf.org) Post-Doctoral fellowship (F-972) to HGR. USDA AFRI NIFA (www.csrees.usda.gov) Postdoctoral Fellowship (MOW-2012-01361) to NF. NSF (www.nsf.gov) Graduate

Abstract

In eukaryotes, ARGONAUTE proteins (AGOs) associate with microRNAs (miRNAs), short interfering RNAs (siRNAs), and other classes of small RNAs to regulate target RNA or target loci. Viral infection in plants induces a potent and highly specific antiviral RNA silencing response characterized by the formation of virus-derived siRNAs. *Arabidopsis thaliana* has ten AGO genes of which AGO1, AGO2, and AGO7 have been shown to play roles in antiviral defense. A genetic analysis was used to identify and characterize the roles of AGO proteins in antiviral defense against *Turnip mosaic virus* (TuMV) in Arabidopsis. AGO1, AGO2 and AGO10 promoted anti-TuMV defense in a modular way in various organs, with AGO2 providing a prominent antiviral role in leaves. AGO5, AGO7 and AGO10 had minor effects in leaves. AGO1 and AGO10 had overlapping antiviral functions in inflorescence tissues after systemic movement of the virus, although the roles of AGO1 and AGO10 accounted for only a minor amount of the overall antiviral activity. By combining AGO protein immunoprecipitation with high-throughput sequencing of associated small RNAs, AGO2, AGO10, and to a lesser extent AGO1 were shown to associate with siRNAs derived from silencing suppressor (HC-Pro)-deficient TuMV-AS9, but not with siRNAs derived from wild-type TuMV. Co-immunoprecipitation and small RNA sequencing revealed that viral siRNAs broadly associated with wild-type HC-Pro during TuMV infection. These results support the hypothesis that suppression of antiviral silencing during TuMV infection, at least in part,

Research Fellowship (DGE-1143954) to JSH Japan Society for the Promotion of Science (www.jsps.go.jp) Postdoctoral Fellowship to AT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

occurs through sequestration of virus-derived siRNAs away from antiviral AGO proteins by HC-Pro. These findings indicate that distinct AGO proteins function as antiviral modules, and provide a molecular explanation for the silencing suppressor activity of HC-Pro.

Author Summary

RNA silencing is a primary, adaptive defense system against viruses in plants. Viruses have evolved counter-defensive mechanisms that inhibit RNA silencing through the activity of silencing suppressor proteins. Understanding how antiviral silencing is controlled, and how suppressor proteins function, is essential for understanding how plants normally resist viruses, why some viruses are highly virulent in different hosts, and how sustainable antiviral resistance strategies can be deployed in agricultural settings. We used a mutant version of *Turnip mosaic virus* lacking a functional silencing suppressor (HC-Pro) to understand the genetic requirements for resistance in the model plant *Arabidopsis thaliana*. We focused on ARGONAUTE proteins, which have long been hypothesized to bind short interfering RNAs (siRNAs) derived from virus genomes for use as sequence-specific guides to recognize and target viral RNA for degradation or repression. We demonstrated specialized antiviral roles for specific ARGONAUTES and showed that several can bind viral siRNAs from across the entire viral genome. However, ARGONAUTE proteins are only loaded with virus-derived siRNAs in the absence of HC-Pro, which we showed binds siRNAs from the viral genome. This indicates that several AGO proteins, which collectively are necessary for full anti-TuMV defense, need to properly load virus-derived siRNAs to execute their antiviral roles.

Introduction

In plants, RNA silencing is a highly specific and adaptive defense mechanism against viruses [1, 2]. Factors involved in antiviral silencing overlap with those of endogenous small RNA pathways, and include i) small RNA biogenesis components such as Dicer-like ribonucleases (DCLs), RNA-dependent RNA polymerases (RDRs), and double-stranded RNA (dsRNA) binding proteins, and ii) ARGONAUTE (AGO) proteins, which function as small RNA-binding effectors [3–6].

RNA-based silencing is triggered by dsRNA that is processed by DCLs into 21- to 24-nt short interfering RNAs (siRNAs), which subsequently associated with AGO proteins to form the RNA-induced silencing complex (RISC) [7, 8]. Inhibition of target RNA can occur by endonucleolytic cleavage (“slicing”), translational repression, or delivery of chromatin-modifying complexes to a locus [9–11, 12]. In some cases, amplification of the silencing response occurs by triggering dsRNA synthesis and secondary siRNA accumulation [13].

Viruses are inducers of RNA silencing; infected plants accumulate large amounts of siRNAs derived from viral RNAs [1]. Most plant viruses encode one or more silencing suppressor proteins that interfere with antiviral RNA silencing [13, 14]. One mechanism of silencing suppression by viral suppressors is through sequestration of siRNA duplexes [1], preventing assembly of the RISC effector complex. Other viral silencing suppressors promote AGO degradation [15–19], prevent slicing or degradation of target RNAs by associating with AGOs [20, 21], or use other mechanisms (for a recent review see Nakahara and Masuta 2014 [22]). In effect, viral

suppressors mask the effects of antiviral silencing, making genetic analysis of antiviral silencing factors in host plants dependent on the use of suppressor-deficient viruses [3, 4, 6, 23].

A. thaliana has ten AGO genes [24], of which AGO1, AGO2 and AGO7 have been implicated in antiviral defense against various viruses by genetic and biochemical criteria [6, 25–31]. Antiviral roles for AGO3 and AGO5 have also been suggested based on virus-derived siRNA association and/or *in vitro* analyses [8, 32]. One model for AGO antiviral activity states that AGO proteins bind virus-derived siRNAs and directly repress viral RNA through slicing, translational repression, or other mechanisms [2, 8, 33]. Given that AGO-dependent regulation of gene expression affects numerous biological processes, including DNA repair [34], AGO proteins might also affect virus replication indirectly through regulation of genes with roles in defense. For example, AGO2-miR393* complexes regulate the expression of *MEMBRIN 12* (*MEMB12*), which is required for resistance to *Pseudomonas syringae* in *A. thaliana* [35]. Moreover, some AGO proteins are known to modulate the activity of other AGO proteins [36, 37], which could affect AGOs with roles in antiviral defense.

Potyviral HC-Pro is a suppressor of RNA silencing. As shown using potyviruses like *Turnip mosaic virus* (TuMV) [23, 38], the counter-defensive function of HC-Pro is necessary for establishment of infection or systemic spread. HC-Pro has been proposed to function through sequestration of virus-derived siRNAs [39–44]. HC-Pro may also function through physical interaction with factors like the transcription factor RAV2 [45], translation initiation factors eIF(iso)4E and eIF4E [46], calmodulin-related protein (CaM) [47], auxiliary proteins like Heat Shock Protein 90 (HSP90) [48], and/or through effects on downstream defense or silencing factors [49, 50]. Here, the role of several *A. thaliana* AGOs in antiviral defense against TuMV was analyzed in various organs of systemically infected plants. The impact of HC-Pro on the loading of antiviral AGOs with virus-derived siRNAs was also studied.

Results

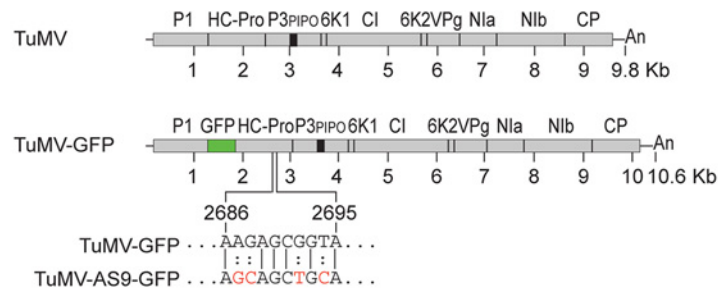
AGO2 has a strong antiviral effect in leaves

Three of the ten *A. thaliana* AGO genes have been implicated in antiviral defense: AGO1 against *Cucumber mosaic virus* (CMV) [25], *Turnip crinkle virus* (TCV) [6, 33], and *Brome mosaic virus* (BMV) [30]; AGO2 against TCV [26], *Potato virus X* (PVX) [27], CMV [26, 28, 29], and TuMV [31]; and AGO7 against TCV [6]. To identify the complete set of AGOs required for antiviral defense against TuMV in *A. thaliana*, single, double, and triple *ago* mutants were inoculated with a GFP-expressing form of parental TuMV (TuMV-GFP) and HC-Pro-deficient TuMV-AS9-GFP [23]. The GFP sequence was inserted between P1 and HC-Pro sequences (Fig. 1A). Both TuMV and TuMV-GFP require translation factor eIF(iso)4E [51], and lead to similar virus-derived siRNA profiles in wild-type and *dicer-like* mutant *A. thaliana* [23]. To determine if AGOs have spatially distinct functions, TuMV-GFP and TuMV-AS9-GFP accumulation was analyzed in inoculated rosette leaves, and in noninoculated cauline leaves and inflorescences. Establishment of local and systemic infection was monitored using GFP fluorescence, and virus accumulation in inoculated and noninoculated tissues was measured by immunoblotting assays (coat protein) as described [23].

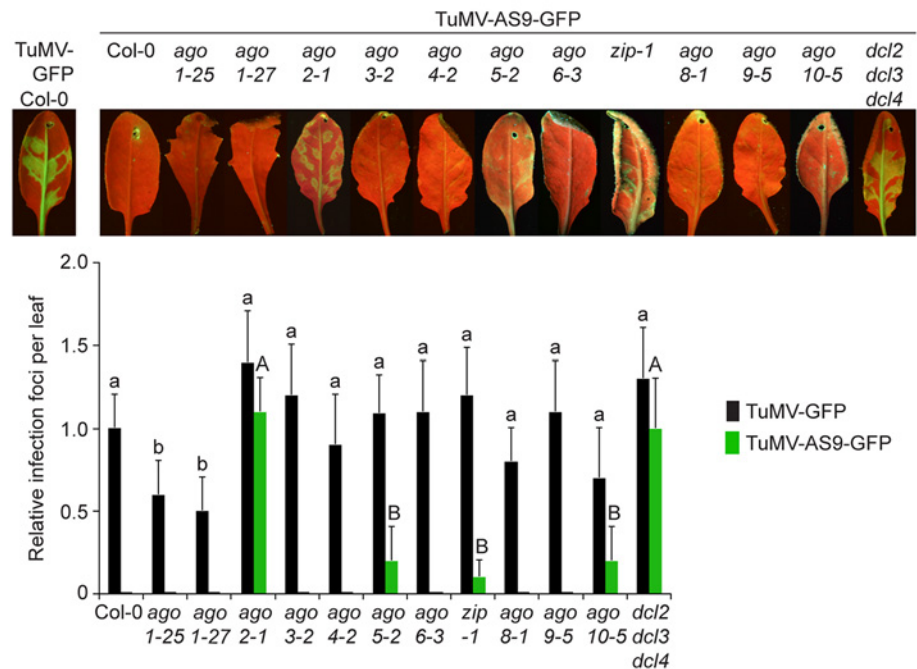
Parental TuMV-GFP was detected in inoculated leaves and noninoculated inflorescences of all single *ago* mutants analyzed (Table 1 and Fig. 1B). Local infection of single *ago1* mutants was significantly lower than that of wild-type Col-0 (Fig. 1B), but this was likely due to the difficulty of inoculating the smaller leaves of hypomorphic mutants containing *ago1* alleles.

As described for *A. thaliana rdr* and *dcl* mutants [23], suppressor-deficient TuMV-AS9-GFP was expected to infect only those plants lacking one or more AGOs with a role in antiviral defense. No infection foci were detected in wild-type Col-0 plants (Fig. 1B and Table 1). Local

A Schematic representation of TuMV and TuMV-GFP genomes



B Local infection of single mutants by TuMV-GFP and TuMV-AS9-GFP at 7 dpi



C Systemic infection by TuMV-AS9-GFP at 15 dpi

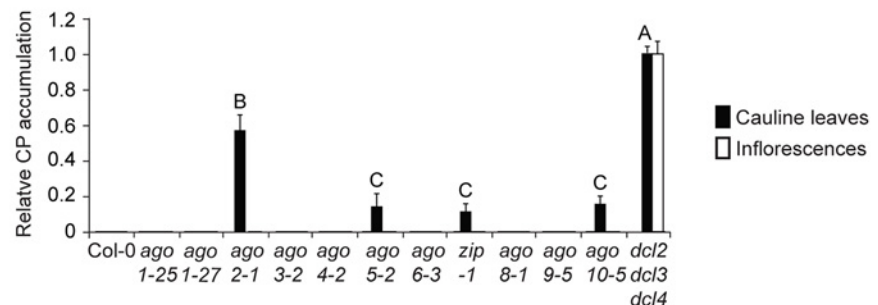


Fig 1. Local and systemic infection of *A. thaliana* single *ago* mutants by TuMV-GFP and TuMV-AS9-GFP. (A) Schematic representation of the TuMV and TuMV-GFP genomes showing insertion of GFP between P1 and HC-Pro, and the AS9 mutation on HC-Pro. (B) Visualization of local infection of inoculated rosette leaves. Pictures were taken at 7 days post inoculation (dpi). Col-0 infected by TuMV-GFP is shown for comparison. The histogram shows average (+ SE) infection efficiency of 14 plants, each with four inoculated leaves. Infection efficiency by TuMV-GFP or TuMV-AS9-GFP is expressed relative to Col-0

(9.8 ± 2 foci per leaf) or to *dcl2-1 dcl3-1 dcl4-2* (3.5 ± 1.4 foci per leaf), respectively. For each virus, bars with the same letter are not statistically different (Tukey's test with $\alpha = 0.05$). (C) TuMV-AS9-GFP coat protein (CP) accumulation in noninoculated cauline leaves and in inflorescence at 15 dpi is determined by immunoblotting and expressed relative to *dcl2-1 dcl3-1 dcl4-2*. The histogram shows average (\pm SE) of four biological replicates. Bars with the same letter are not statistically different (Tukey's test with $\alpha = 0.05$). The experiment was repeated twice with similar results.

doi:10.1371/journal.ppat.1004755.g001

infection foci of suppressor-deficient TuMV-AS9-GFP were readily visible at 7 days post inoculation (dpi) in *ago2-1* mutant plants (Fig. 1B and Table 1), and infection efficiency was not significantly different than that of the *dcl2-1 dcl3-1 dcl4-2* triple mutant, which served as the hypersusceptible, silencing-deficient control (Fig. 1B) [23]. Low numbers of infection foci were also detected in single *ago5-2*, *zip-1* (*ago7*), and *ago10-5* mutant plants (Fig. 1B and Table 1). Systemic movement of TuMV-AS9-GFP into cauline leaves was detected at 15 dpi in *ago2-1* plants, and also in *ago5-2*, *zip-1*, and *ago10-5* plants though at significantly lower levels (Fig. 1C

Table 1. TuMV-GFP and TuMV-AS9-GFP infection in single *ago* mutants ^a.

Virus	Arabidopsis genotype	Plants inoculated	Local infection	Cauline leaves	Inflorescence
TuMV-GFP					
	Col-0	14	14	14	14
	<i>ago1-25</i>	14	14	14	14
	<i>ago1-27</i>	14	14	14	14
	<i>ago2-1</i>	14	14	14	14
	<i>ago3-2</i>	14	14	14	14
	<i>ago4-2</i>	14	14	14	14
	<i>ago5-2</i>	14	14	14	14
	<i>ago6-3</i>	14	14	14	14
	<i>zip-1</i>	14	14	14	14
	<i>ago8-1</i>	14	14	14	14
	<i>ago9-5</i>	14	14	14	14
	<i>ago10-5</i>	14	14	14	14
	<i>dcl2-1 dcl3-1 dcl4-2</i>	14	14	14	14
TuMV-AS9-GFP					
	Col-0	14	0	0	0
	<i>ago1-25</i>	14	0	0	0
	<i>ago1-27</i>	14	0	0	0
	<i>ago2-1</i>	14	14	14	0
	<i>ago3-2</i>	14	0	0	0
	<i>ago4-2</i>	14	0	0	0
	<i>ago5-2</i>	14	6	6	0
	<i>ago6-3</i>	14	0	0	0
	<i>zip-1</i>	14	5	5	0
	<i>ago8-1</i>	14	0	0	0
	<i>ago9-5</i>	14	0	0	0
	<i>ago10-5</i>	14	7	7	0
	<i>dcl2-1 dcl3-1 dcl4-2</i>	14	14	14	14

^a Number of plants showing local and systemic infections were scored by GFP fluorescence under UV illumination. Local infection foci were counted at 7 days post-inoculation (dpi). All other data is from plants at 15 dpi.

doi:10.1371/journal.ppat.1004755.t001

and Table 1). In cauline leaves from single *ago2-1* mutant plants, TuMV-AS9-GFP accumulated to approximately 60% of the level measured in *dcl2-1 dcl3-1 dcl4-2* plants, while *ago5-2*, *zip-1*, and *ago10-5* plants accumulated TuMV-AS9-GFP to approximately 10% of the levels measured in the hypersusceptible control (Fig. 1C). In contrast to *dcl2-1 dcl3-1 dcl4-2* plants, systemic infection by TuMV-AS9-GFP did not reach inflorescence tissues in any of the single *ago* mutant or Col-0 plants (Fig. 1C and Table 1). Systemic infection did not reach cauline leaves in any of the other single *ago* mutants or Col-0 plants (Fig. 1C and Table 1).

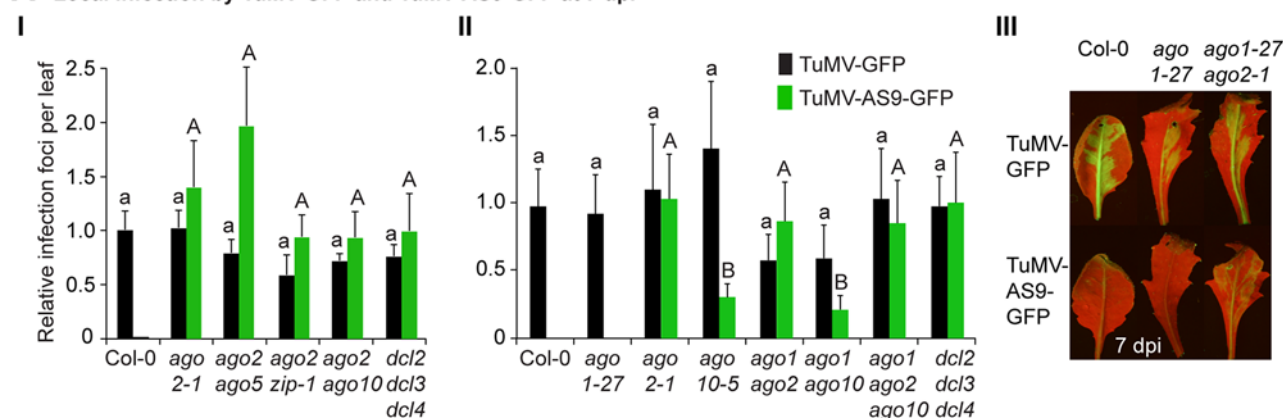
AGO1 and AGO10 have modest antiviral effects in inflorescences

To determine if the major effect of AGO2 was additive with the minor effects of AGO5, AGO7 and AGO10, and to examine if AGO1 possessed redundant or masked activities, double and triple *ago* mutant plants were inoculated with TuMV-GFP or TuMV-AS9-GFP, and virus accumulation was measured in inoculated and noninoculated organs as described above. To reduce the effect of differences in leaf size, we planted mutant lines with the *ago1-27* allele one week earlier than the other mutant lines inoculated at the same time. Parental TuMV-GFP infected locally (Fig. 2A panels I and II) and moved systemically into the inflorescence of all double and triple *ago* mutants analyzed (Tables 2 and 3), with no significant differences in infection efficiency.

In double mutants harboring the *ago2-1* allele and one of *ago5-2*, *zip-1*, or *ago10-5* alleles, no significant differences in number of infection foci were detected at 7 dpi in rosette leaves inoculated with TuMV-AS9-GFP (Fig. 2A panel I and Table 2). Similarly, no significant differences were detected in TuMV-AS9-GFP coat protein accumulation in cauline leaves at 15 dpi (Fig. 2B panel I). As observed for the *ago* single mutants, TuMV-AS9-GFP was not detected in inflorescences from double mutant plants containing the *ago2-1* allele (Fig. 2B panel I). These results indicate that the minor activities of AGO5, AGO7 and AGO10 are not additive with the major antiviral activity of AGO2. Double and triple mutants harboring the *ago1-27* allele were generated and inoculated with parental TuMV-GFP or suppressor-deficient TuMV-AS9-GFP. Col-0 plants and *ago1-27*, *ago2-1* and *ago10-5* single mutant lines were included as controls. Local TuMV-AS9-GFP infection foci were observed in inoculated rosette leaves, and virus was detected in noninoculated cauline leaves, from *ago1-27 ago2-1* double mutant plants, but *ago1-27* had no enhancing or suppressing effects when combined with *ago2-1* (panel II in Fig. 2A and 2B, Table 3). Combining *ago1-27* with *ago10-5*, or with *ago2-1* and *ago10-5* in a triple mutant, had no effects on local TuMV-AS9-GFP infection foci (Fig. 2A panel II) or accumulation in cauline leaves beyond those measured in the single *ago2* or double *ago2 ago10* mutants (Fig. 2B panels I and II, and Table 3). However, combining *ago1-27* with *ago10-5* resulted in an increase in TuMV-AS9-GFP CP accumulation in cauline leaves relative to single *ago10-5* mutants (Fig. 2B panel II). Infection efficiency of *ago1* single, double or triple mutants by TuMV-GFP was similar to that of wild type plants (Fig. 2A panels II and III), and infection efficiency of *ago1-27 ago2-1* double and *ago1-27 ago2-1 ago10-5* triple mutants by TuMV-AS9-GFP was similar to that of *dcl2-1 dcl3-1 dcl4-2* plants used as susceptible control (Fig. 2A panel II). Thus, both the lack of TuMV-AS9-GFP infection in single *ago1* mutants and the lack of systemic infection of inflorescence in *ago1-27 ago2-1* double mutants were not due to pleiotropic effects.

Surprisingly, systemic infection of inflorescence tissue was detected in the *ago1-27 ago10-5* double mutant and *ago1-27 ago2-1 ago10-5* triple mutant plants (Fig. 2B panel III and Table 3). Among all single and combination *ago* mutants tested, only those containing both *ago1* and *ago10* defects exhibited movement to, and accumulation in, inflorescences. However, while TuMV-AS9-GFP was detected in all inflorescence clusters of the *dcl2-1 dcl3-1 dcl4-2*

A Local infection by TuMV-GFP and TuMV-AS9-GFP at 7 dpi



B Systemic infection by TuMV-AS9-GFP at 15 dpi

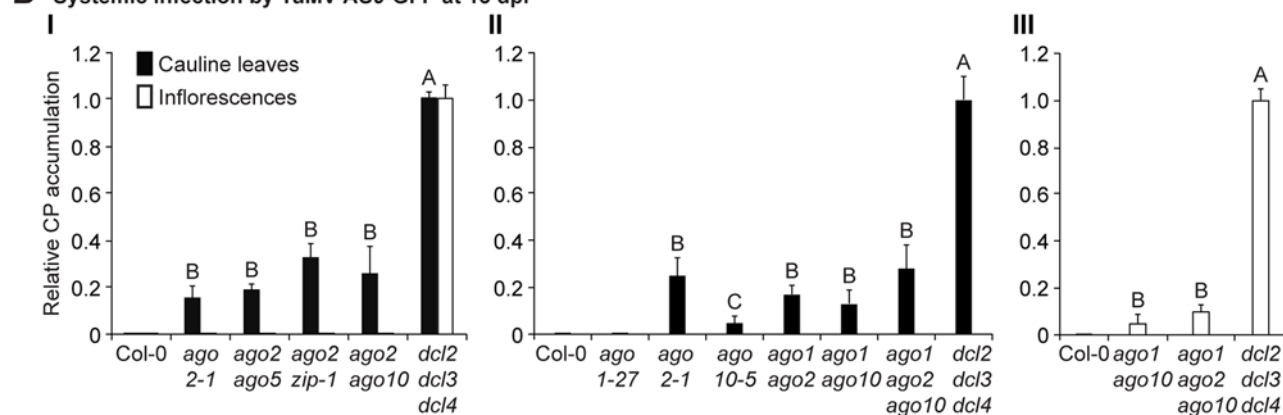


Fig 2. Local and systemic infection of a selected group of double and triple *ago* mutants by TuMV-GFP and TuMV-AS9-GFP. (A) Local infection efficiency. Panel I: infection efficiency of TuMV-GFP or TuMV-AS9-GFP is expressed relative to Col-0 (19.6 ± 3.3 foci per leaf) or to *dcl2-1 dcl3-1 dcl4-2* (2.2 ± 0.7 foci per leaf), respectively. The histogram shows the average (+ SE) of 10 plants, each with four inoculated leaves. Panel II: local infection of inoculated rosette leaves for a selected group of mutants harboring *ago1-27*. The histogram shows average (+ SE) infection efficiency of 14 plants, each with four inoculated leaves. Infection efficiency of TuMV-GFP or TuMV-AS9-GFP is expressed relative to Col-0 (4.1 ± 1.2 foci per leaf) or to *dcl2-1 dcl3-1 dcl4-2* (2.8 ± 1.1 foci per leaf), respectively. Panel III: Representative leaves of *ago1-27* single and *ago1-27 ago2-1* double mutants showing TuMV-GFP local infection foci. *ago1-27 ago2-1*, but not *ago1-27*, was infected by TuMV-AS9-GFP. Col-0 is shown for comparison. Pictures were taken at 7 dpi under UV light. (B) Systemic infection. TuMV-AS9-GFP coat protein accumulation in noninoculated cauline leaves and in inflorescence at 15 dpi. Panel I: double mutants harboring *ago2-1*. Panel II: double and triple mutants harboring *ago1-27* and *ago10-5*. The histograms show average (+ SE) of four biological replicates, expressed relative to *dcl2-1 dcl3-1 dcl4-2*. Bars with the same letter are not statistically different (Tukey's test with $\alpha = 0.05$). Panel III: in double and triple mutants harboring *ago1-27*, inflorescence samples were collected only from clusters showing systemic GFP.

doi:10.1371/journal.ppat.1004755.g002

triple mutant reference, in *ago1-27 ago10-5* and in *ago1-27 ago2-1 ago10-5* TuMV-AS9-GFP was detected only in 4% and 14% of the inflorescence clusters, respectively (Table 3). In inflorescences of *ago1-27 ago10-5* and *ago1-27 ago2-1 ago10-5* plants with visible GFP fluorescence, TuMV-AS9-GFP CP accumulated to 5% and 10% relative to the *dcl2-1 dcl3-1 dcl4-2* triple mutant (Fig 2B panel III).

Collectively, the genetic analysis of local and systemic infection using TuMV-AS9-GFP revealed two sets of AGOs that limit infection. In inoculated rosette and noninoculated cauline leaves, AGO2 plays a major antiviral role, while AGO5, AGO7 and AGO10 play minor roles that are non-additive with AGO2. In noninoculated inflorescence tissues, AGO1 and AGO10 play overlapping or redundant antiviral roles, but these functions likely account for only a fraction of the RNA-mediated antiviral activity. It is possible that other factors, including AGO

Table 2. TuMV-GFP and TuMV-AS9-GFP infection in selected *ago2-1* based double mutants^a.

Virus	Arabidopsis genotype	Plants inoculated	Local infection	Cauline leaves	Inflorescence
TuMV-GFP					
	Col-0	10	10	10	10
	<i>ago2-1</i>	10	10	10	10
	<i>ago2-1 ago5-2</i>	10	10	10	10
	<i>ago2-1 zip-1</i>	10	10	10	10
	<i>ago2-1 ago10-5</i>	10	10	10	10
	<i>dcl2-1 dcl3-1 dcl4-2</i>	10	10	10	10
TuMV-AS9-GFP					
	Col-0	10	0	0	0
	<i>ago2-1</i>	10	10	10	0
	<i>ago2-1 ago5-2</i>	10	10	10	0
	<i>ago2-1 zip-1</i>	10	10	10	0
	<i>ago2-1 ago10-5</i>	10	10	10	0
	<i>dcl2-1 dcl3-1 dcl4-2</i>	10	10	10	10

^a Number of plants showing local and systemic infections were scored by GFP fluorescence under UV illumination. Local infection foci were counted at 7 days post-inoculation (dpi). All other data is from plants at 15 dpi.

doi:10.1371/journal.ppat.1004755.t002

Table 3. TuMV-GFP and TuMV-AS9-GFP infection in selected *ago1-27* based combination mutants^a.

Virus	Arabidopsis genotype	Plants inoculated	Local infection	Cauline leaves	Inflorescence	Percent ^b
TuMV-GFP						
	Col-0	14	14	14	14	100
	<i>ago1-27</i>	14	14	14	14	100
	<i>ago2-1</i>	14	14	14	14	100
	<i>ago10-5</i>	14	14	14	14	100
	<i>ago1-27 ago2-1</i>	14	14	14	14	100
	<i>ago1-27 ago10-5</i>	14	14	14	14	100
	<i>ago1-27 ago2-1 ago10-5</i>	14	14	14	14	100
	<i>dcl2-1 dcl3-1 dcl4-2</i>	14	14	14	14	100
TuMV-AS9-GFP						
	Col-0	14	0	0	0	0
	<i>ago1-27</i>	14	0	0	0	0
	<i>ago2-1</i>	14	14	14	0	0
	<i>ago10-5</i>	14	8	6	0	0
	<i>ago1-27 ago2-1</i>	14	14	14	0	0
	<i>ago1-27 ago10-5</i>	14	14	14	3	4 ±1
	<i>ago1-27 ago2-1 ago10-5</i>	14	14	14	8	14 ±2
	<i>dcl2-1 dcl3-1 dcl4-2</i>	14	14	14	14	100

^a Number of plants showing local and systemic infections were scored by GFP fluorescence under UV illumination. Local infection foci were counted at 7 days post-inoculation. All other data is from plants at 15 dpi.

^b Proportion (%) of inflorescence clusters showing GFP with respect to the total number of clusters on each plant with inflorescence GFP fluorescence.

doi:10.1371/journal.ppat.1004755.t003

proteins not analyzed here, have a role in protecting inflorescence tissue from virus infection. The scope of subsequent AGO analyses was restricted to the functions of AGO1, AGO2 and AGO10 in the presence and absence of functional HC-Pro.

Differential association of AGO2 with viral siRNAs in the presence and absence of functional HC-Pro

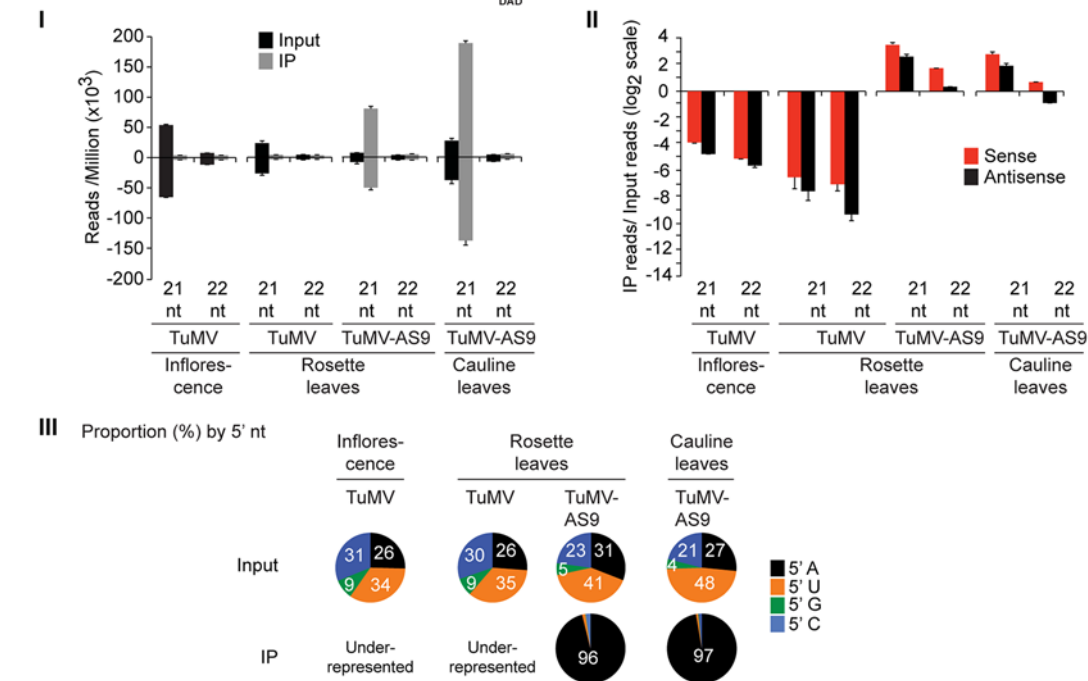
We hypothesized that AGO proteins with anti-TuMV activity associate with TuMV-derived siRNAs. This idea was tested first with epitope-tagged AGO2 in plants inoculated with parental TuMV or HC-Pro-defective TuMV-AS9 (lacking GFP) [23]. AGO2 immunoprecipitation and small RNA sequence analyses were done using transgenic *A. thaliana* expressing a triple-hemagglutinin (HA) epitope-tagged, catalytically inactive form of AGO2 (HA-AGO2_{DAD}). The second of three aspartic acid residues of AGO2 was substituted with alanine; this substitution eliminates antiviral activity of AGO2, but preserves both the siRNA-binding and target RNA-binding functions [31]. These experiments require the use of plants lacking AGO2-mediated antiviral functions, as infection by TuMV-AS9 would otherwise be blocked (Figs. 1 and 2) [31].

Small RNAs from the input (pre-immunoprecipitated) and HA-AGO2_{DAD} co-immunoprecipitated fractions from inoculated rosette leaves and noninoculated inflorescences of TuMV-infected plants were analyzed from duplicate biological samples. Only reads that matched to either the *A. thaliana* or TuMV genomes without mismatches were analyzed (S1 Table). For each individual sample, read counts were scaled with respect to the total number of adaptor-parsed reads (reads per million) for the corresponding flow cell (eight individual samples). In mock-inoculated plants, a small number of reads from the input fractions mapped to TuMV (S1–S4 Tables, and S1 Fig). The source of these reads could be contamination, sequencing error, or portions of the *A. thaliana* genome. Based on the number of reads from mock-inoculated plants mapping to the TuMV genome, the false positive rate (proportion of parsed reads artifactually mapping to TuMV) was estimated to be between 9.8×10^{-6} and 1.0×10^{-4} , which should not have affected subsequent analyses.

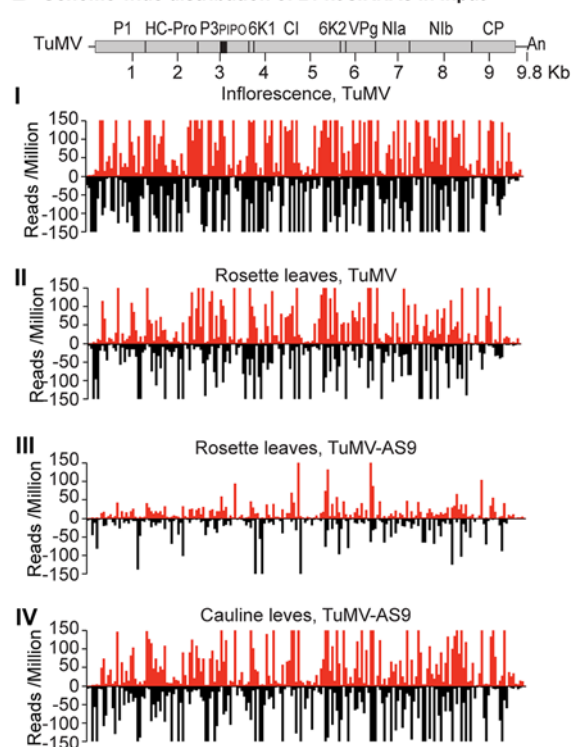
In input fractions from TuMV-infected plants expressing HA-AGO2_{DAD}, the proportion of reads mapping to the *A. thaliana* genome, as opposed to TuMV, varied from 77% (averaged across replicates) to 84% for different tissues (S1A Fig). Sequences mapping to TuMV were mainly 21-nt and 22-nt (S1A Fig). Accordingly, the detailed analyses for HA-AGO2_{DAD} and other proteins (discussed below) were focused on 21-nt (Figs 3–6) and 22-nt sequences (S3–S7) Figs.

Endogenous *A. thaliana* 21-nt small RNAs were enriched in HA-AGO2_{DAD} immunoprecipitates from leaves or inflorescence of mock-inoculated (4.5 to 10 fold) or TuMV-infected samples (2.7 to 6.3 fold) (S2A Fig). Enriched sequences in HA-AGO2_{DAD} immunoprecipitates had predominantly a 5'A nucleotide, as previously reported for AGO2-associated small RNAs [52, 53], or a 5'U nucleotide (S2A Fig). Specific miRNA, miRNA* and trans-acting siRNA (tasiRNA) populations were enriched in HA-AGO2_{DAD} immunoprecipitates from both mock-inoculated (2.3 to 31 fold), and to a lesser extent, TuMV-infected (1.8 to 16 fold) rosette leaves (S8A Fig). MicroRNA read counts for input and immunoprecipitates from this and subsequent analyses are provided in S1 Dataset. MiR390 and miR393* were shown previously to co-immunoprecipitate with AGO2 [35, 52]. In mock-inoculated and TuMV-infected rosette leaves, the number of miR390 reads in HA-AGO2_{DAD} immunoprecipitates was 260 and 65 fold higher, respectively, than in the corresponding input samples. Similarly, miR393* reads were enriched 125 and 60 fold in HA-AGO2_{DAD} immunoprecipitates from mock-inoculated and TuMV-infected rosette leaves, respectively. Therefore, enrichment of *A. thaliana* small RNA populations that are known to be associated with AGO2 occurred as expected.

A Association of TuMV-derived siRNAs with HA-AGO2_{DAD}



B Genome-wide distribution of 21 nt siRNAs in input



C Genome-wide distribution of 21 nt siRNAs in HA-AGO2_{DAD} IP

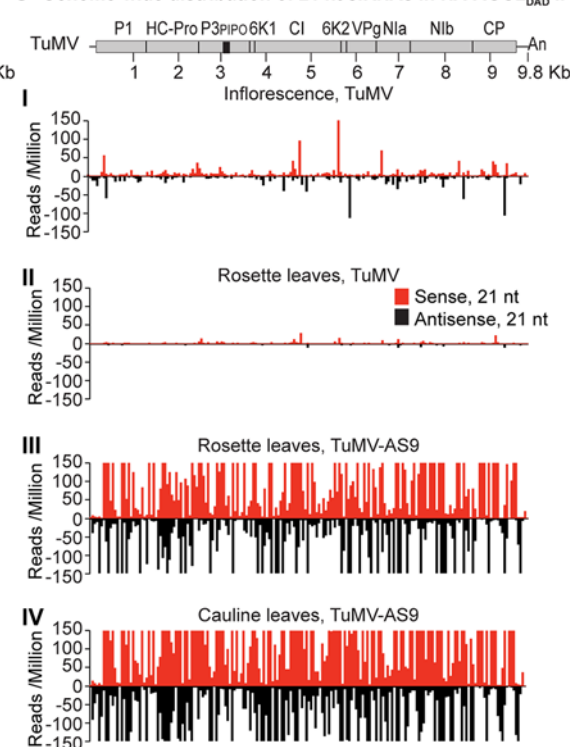


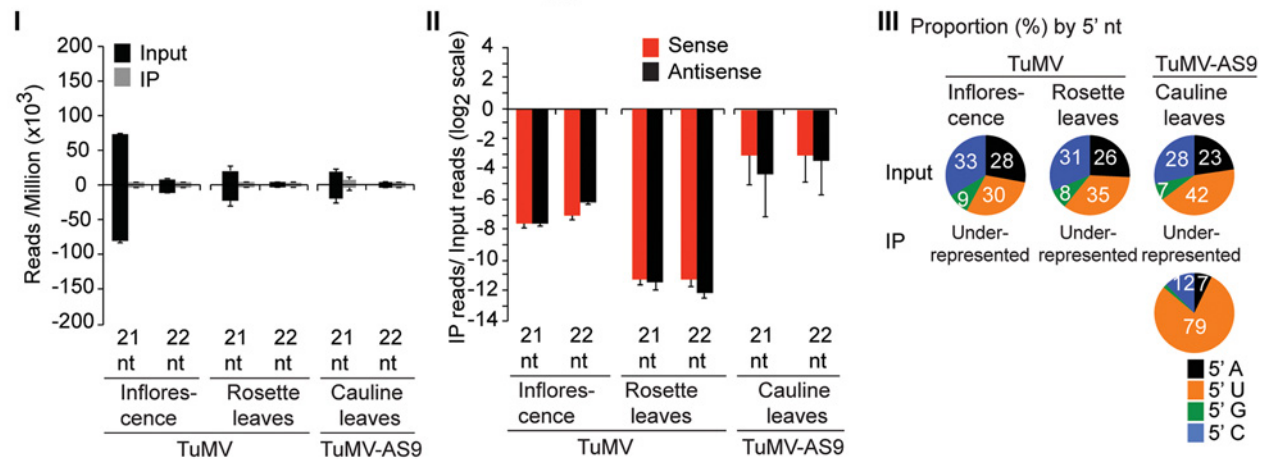
Fig 3. Profile of endogenous and TuMV-derived siRNAs in plants expressing HA-AGO2_{DAD} in an ago2-1 background. Values are average and SE from two biological replicates normalized to reads per million. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively. Inflorescence samples were collected at 10 dpi. (A) Panel I: number of reads by size, class, and polarity, for TuMV-derived siRNAs in input and HA-AGO2_{DAD} IP. Panel II: for 21 and 22 nt TuMV-derived siRNAs, enrichment in HA-AGO2_{DAD} IP. Enrichment is defined as immunoprecipitate (IP) reads/ input reads, expressed on a \log_2 scale. Panel III: proportion (in percentage) of 5' nt in 21 nt and 22 nt TuMV-derived siRNAs by fraction. Numbers were rounded to the nearest integer. (B) and (C) TuMV genome-wide distribution of 21 nt TuMV-derived siRNAs in input (B) and HA-AGO2_{DAD} IP (C). Panel I:

TuMV-infected inflorescence. Panel II: TuMV-inoculated rosette leaves. Panel III: rosette leaves inoculated with TuMV-AS9. Panel IV: cauline leaves systemically infected with TuMV-AS9. Reads were plotted for each 1 nt position. The scale was capped at 150 reads.

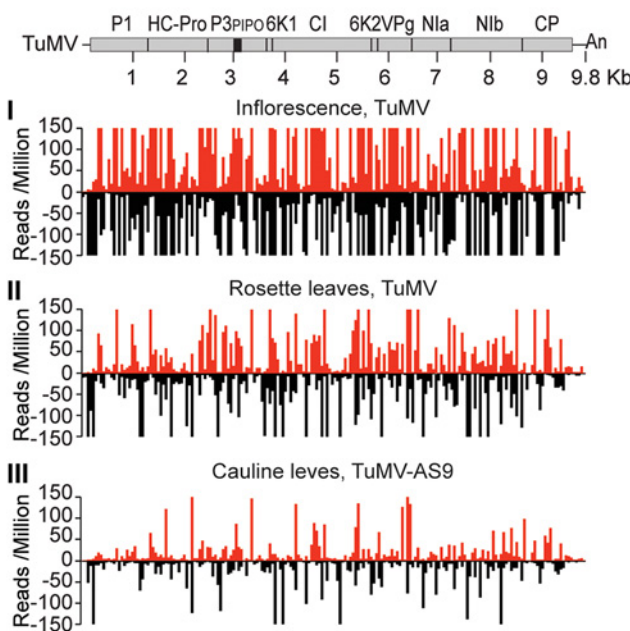
doi:10.1371/journal.ppat.1004755.g003

In TuMV-inoculated rosette leaves, and systemically infected inflorescence, virus-derived siRNAs were abundant, representing 17% and 23%, respectively, of mapped reads in input samples (S1A Fig). Reads mapped to both sense (genomic strand) and antisense strands across the entire TuMV genome. However, both 21- and 22-nt TuMV-derived siRNAs were depleted in HA-AGO2_{DAD} immunoprecipitates (Fig. 3A panels I and II, Fig. 3B and 3C panels I and II,

A Association of TuMV-derived siRNAs with HA-AGO1_{DAH}



B Genome-wide distribution of 21 nt siRNAs in input



C Genome-wide distribution of 21 nt siRNAs in HA-AGO1_{DAH} IP

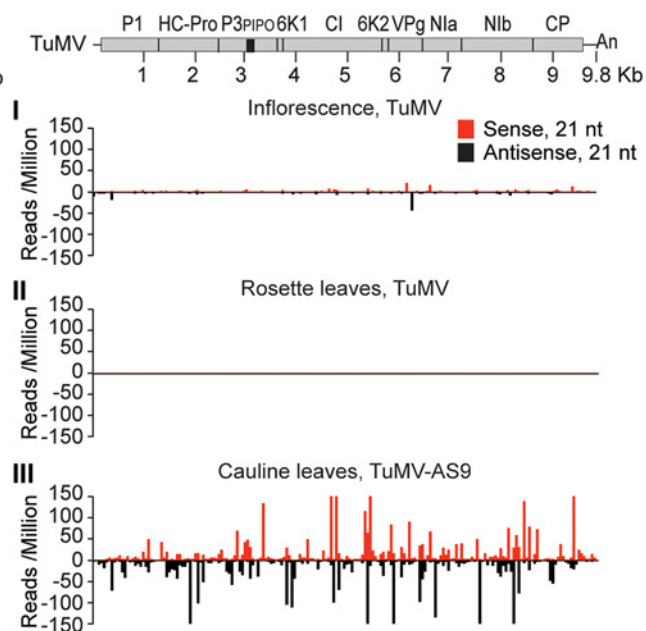
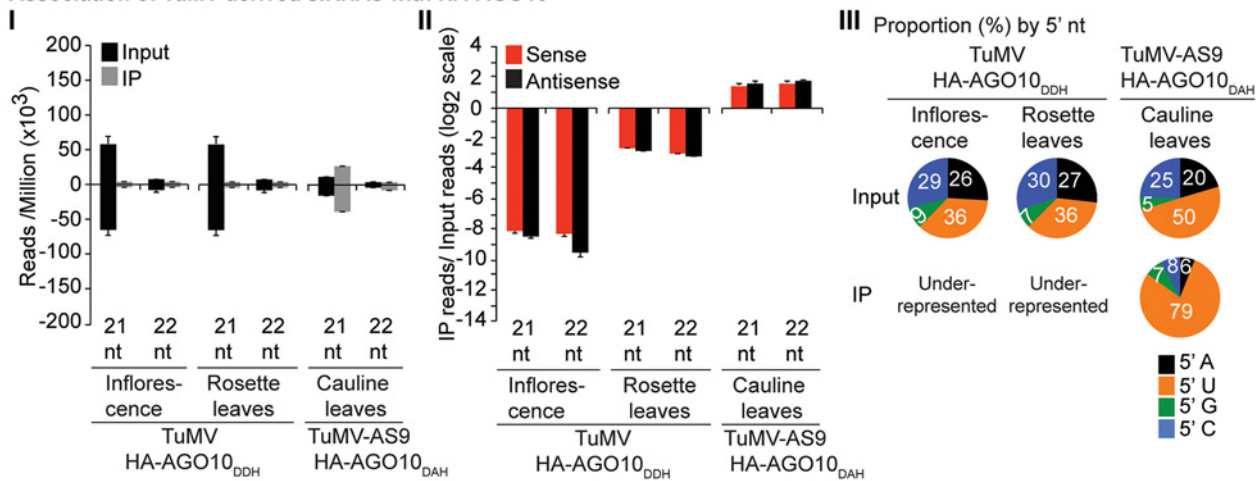


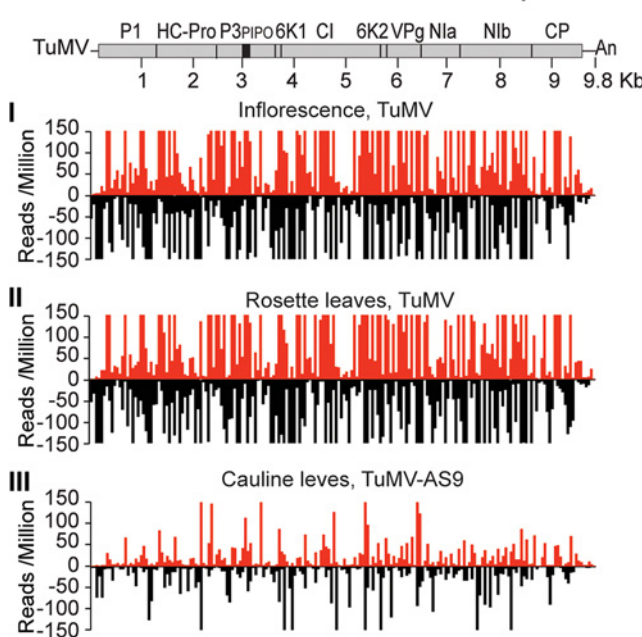
Fig 4. Profile of endogenous and TuMV-derived siRNAs in plants expressing HA-AGO1_{DAH} in an ago2-1 background. Labels are as in Fig. 3. Inflorescence samples were collected at 10 dpi. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively.

doi:10.1371/journal.ppat.1004755.g004

A Association of TuMV-derived siRNAs with HA-AGO10



B Genome-wide distribution of 21 nt siRNAs in input



C Genome-wide distribution of 21 nt siRNAs in HA-AGO10 IP

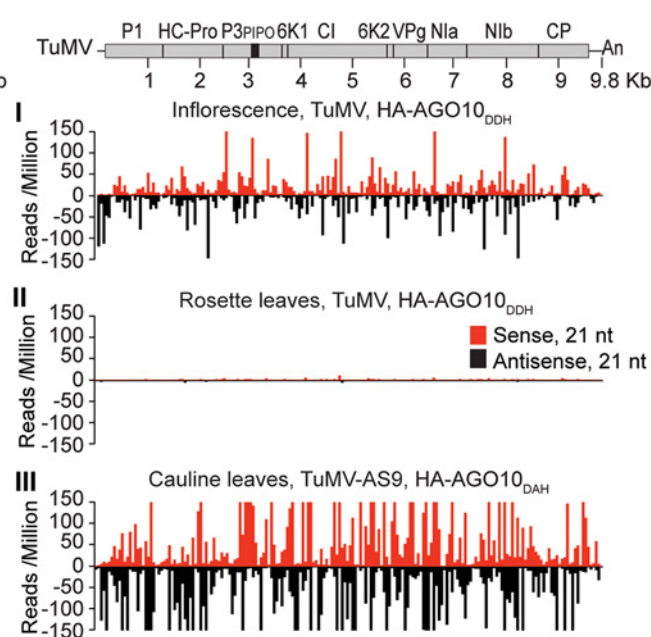


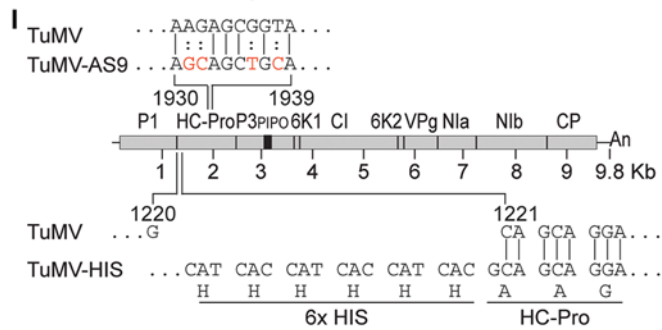
Fig 5. Profile of endogenous and TuMV-derived siRNAs in plants expressing HA-AGO10. Labels are as in Fig 3. Catalytically active HA-AGO10_{DDH} and catalytic mutant HA-AGO10_{DAH} were expressed in a wild-type Col-0 (*AGO2*) or *ago2-1* background, respectively. Inflorescence samples were collected at 10 dpi. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively.

doi:10.1371/journal.ppat.1004755.g005

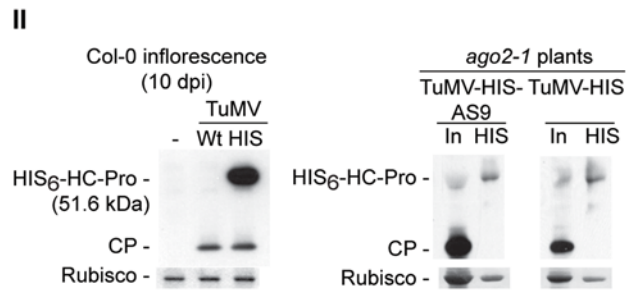
and S3 Fig); only a small number of individual TuMV-derived siRNAs were marginally enriched.

In leaves of TuMV-AS9-infected plants, endogenous *A. thaliana* small RNAs were again enriched (2.7 to 4.6 fold) in HA-AGO2_{DAD} immunoprecipitates, with patterns expected of AGO2-associated small RNAs (S2A Fig). Virus-derived siRNAs represented 7% or 16% of mapped reads in input samples from inoculated rosette leaves or systemically infected cauline leaves, respectively (S1A Fig). However, in striking contrast to TuMV-infected samples, both 21- and 22-nt TuMV-AS9-derived siRNAs were highly enriched relative to TuMV-derived

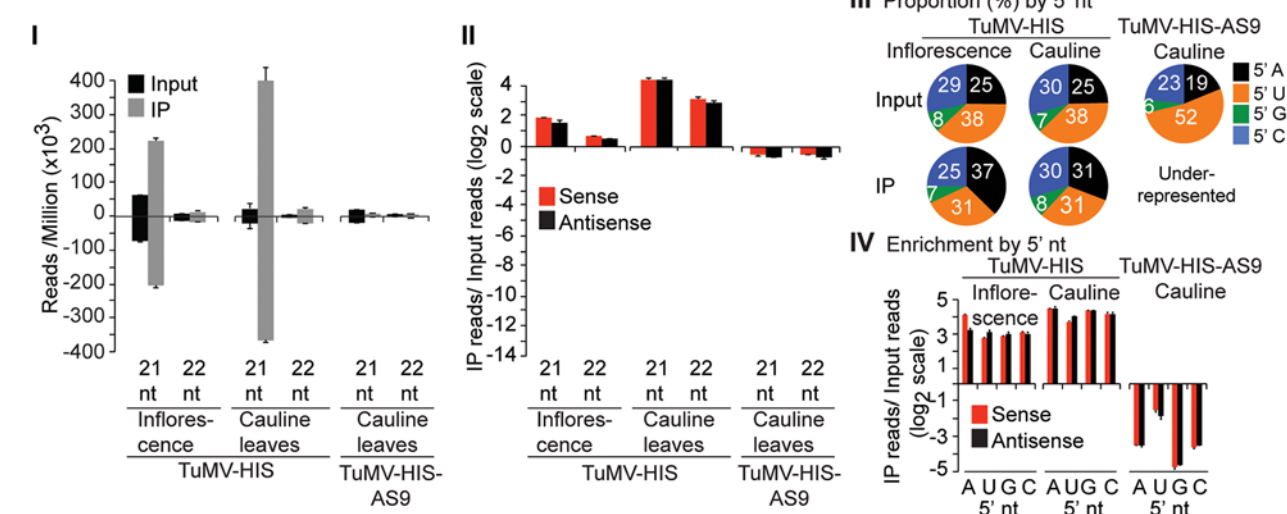
A Infectious clones of parental and modified TuMV



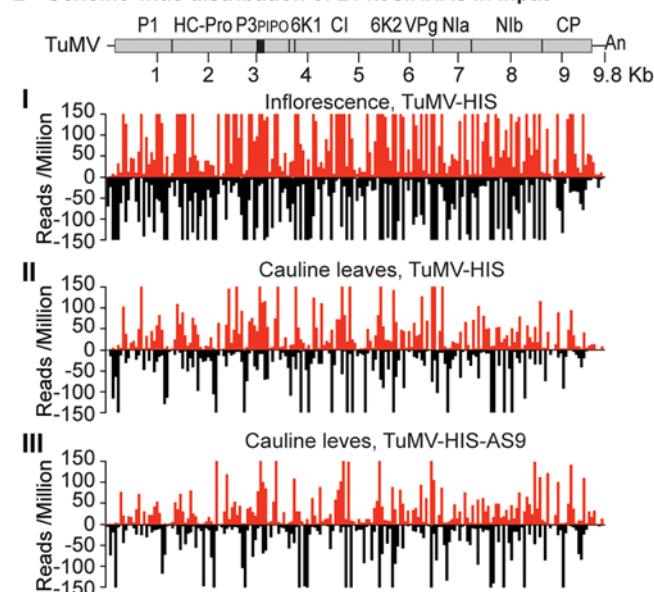
B HC-Pro IP



C Association of TuMV-derived siRNAs with HC-Pro



D Genome-wide distribution of 21 nt siRNAs in input



E Genome-wide distribution of 21 nt siRNAs in HC-Pro IP

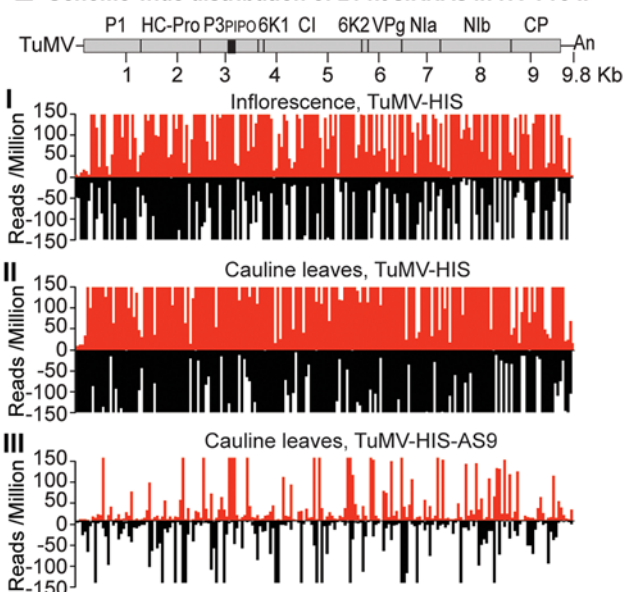


Fig 6. Profile of TuMV-derived siRNAs in plants infected with TuMV-HIS or TuMV-HIS-AS9. (A) Panel I: schematic representation of the TuMV genome and modified clones with an AS9 mutation and a 6xHIS tag (TuMV-HIS). Coordinates correspond to wild-type TuMV. The 6xHIS tag fused in frame to HC-Pro is underlined. Panel II: representative blot CP and HC-Pro accumulation in inflorescence of Col-0 at 10 dpi. (B) CP and HC-Pro accumulation in input and HC-Pro (wild-type and AS9) immunoprecipitation from cauline leaves of *ago2-1* plants. Samples from plants infected with TuMV-HIS or TuMV-HIS-AS9 were collected at 10 and 15 dpi, respectively. 6.25 µg of total protein or 10 µl of immunoprecipitate (IP) were loaded for TuMV-HIS input and IP samples,

respectively. Amounts were doubled for TuMV-HIS-AS9 input and IP. (C) Panel I: number of reads by size, class, and polarity, for TuMV-derived siRNAs in input and wild-type or AS9 HC-Pro IP. Panel II: enrichment in HC-Pro IP as in Fig. 3. Panel III: proportion (in percentage) of 5' nt in 21 nt and 22 nt TuMV-derived siRNAs by fraction. Panel IV: bars show the enrichment of TuMV-derived siRNAs by 5' nt and polarity. (D) and (E) TuMV genome-wide distribution of 21 nt TuMV-derived siRNAs in input (D) and HC-Pro IP (E). Reads were plotted for each 1 nt position. The scale was capped at 150 reads.

doi:10.1371/journal.ppat.1004755.g006

siRNAs in HA-AGO2_{DAD} immunoprecipitates from both inoculated rosette leaves and systemically infected cauline leaves (Fig. 3A panels I and II, Fig. 3B and 3C panels III and IV, and S3 Fig). Among co-immunoprecipitated siRNAs, those containing a 5'A were overrepresented (Fig. 3A panel III). Association of AGO2 with siRNAs derived from TuMV-AS9, but not from TuMV, was verified by small RNA northern blot assays (S9 Fig). These results indicate that programming of AGO2 with TuMV-derived siRNAs is inhibited in the presence of active HC-Pro.

Differential association of AGO1 and AGO10 with viral siRNAs in the presence and absence of functional HC-Pro

A similar experimental design was used to test the association of tagged AGO1 and AGO10 with TuMV and TuMV-AS9-derived siRNAs. To enable infection by suppressor-deficient TuMV-AS9, transgenic *A. thaliana* plants expressing catalytically defective HA-AGO1_{DAH} [31] or HA-AGO10_{DAH} were produced in the TuMV-AS9-permissive *ago2-1* background. Phenotypic defects associated to catalytic mutant HA-AGO1_{DAH} were more severe in an *ago1-25* mutant that in a wild-type (*AGO1*) background [31]. Effects of catalytically defective HA-AGO10_{DAH} on plant phenotype were not known, so transgenic *A. thaliana* plants expressing catalytically active HA-AGO10_{DDH} in a wild-type Col-0 background were also generated. Transgenic lines were inoculated with TuMV or TuMV-AS9 and samples from inoculated rosette leaves and systemically infected cauline leaves or inflorescences were collected from biological replicates. Small RNAs from input samples and immunoprecipitated fractions were sequenced, and reads were mapped and counts were scaled as described above. Tagged versions of AGO1 and AGO10 associated with small RNAs with a 5'U, as expected (S2B and S2C Fig panel II) [36, 52–54], and the proportion of *A. thaliana* and TuMV-derived siRNAs (S1B and S1C Fig) was similar to the observed in plants expressing HA-AGO2 (S1A Fig).

In mock-inoculated samples, endogenous *A. thaliana* 21-nt small RNAs were enriched 5 to 15 fold, and 5 to 7 fold, in HA-AGO1_{DAH} and HA-AGO10_{DAH} immunoprecipitates, respectively. In TuMV- and TuMV-AS9-infected samples, *A. thaliana* 21-nt small RNAs were enriched 5 and 15 fold, respectively, in HA-AGO1_{DAH} immunoprecipitates (S2B Fig panel I). In TuMV-infected samples, *A. thaliana* 21-nt small RNAs were enriched 1.5 and 2.5 fold in HA-AGO10_{DDH} immunoprecipitates from inflorescences and rosette leaves, respectively (S2C Fig panel I). In TuMV-AS9-infected samples, *A. thaliana* 21-nt small RNAs were enriched 7 fold in HA-AGO10_{DAH} immunoprecipitates from cauline leaves (S2C Fig panel I). Sequences with a 5'U were enriched with both AGOs (panel II in S2B and S2C Fig), as expected [36, 52–54]. MiRNAs were enriched in HA-AGO1_{DAH} and HA-AGO10_{DAH} immunoprecipitates from both mock-inoculated (7 to 50 fold) and TuMV-infected (3 to 25 fold) samples, while miRNA* and tasiRNA populations were variable (S8B and S8C Fig). For example, miR166 reads were enriched 30 and 45 fold in HA-AGO1_{DAH} immunoprecipitates from inflorescences of mock-inoculated and TuMV-infected plants, respectively. MiR168 reads were likewise enriched 20 and 12 fold. MiR166 reads were enriched 900 and 60 fold in HA-AGO10_{DAH} immunoprecipitates from mock-inoculated and TuMV-infected plants, respectively, in agreement with previous observations [36].

In rosette and inflorescence tissues from each of the transgenic lines, TuMV infection triggered abundant 21- and 22-nt siRNAs that originated from sense and antisense strands across the entire viral genome (Figs. 4B and 5B). However, as with HA-AGO2_{DAD} immunoprecipitates, TuMV-derived siRNAs were depleted in both HA-AGO1_{DAH} (Fig. 4A–4C panels I and II, and S4 Fig) and HA-AGO10_{DDH} (Fig. 5A–5C panels I and II, and S5 Fig) immunoprecipitates. By contrast, in plants infected with suppressor-deficient TuMV-AS9, virus-derived siRNAs were enriched in HA-AGO10_{DAH} immunoprecipitates (Fig. 5A panels I and II, Fig. 5B and 5C panels III, and S5 Fig), and had predominantly a 5'U nucleotide (Fig. 5A panel III). Individual highly enriched sequences were distributed across the TuMV-AS9 genome (Fig. 5C panel III and S5 Fig), suggesting that AGO10 may target all regions of TuMV-AS9 genome. TuMV-AS9-derived siRNAs were present in HA-AGO1_{DAH} immunoprecipitates at a higher level than in immunoprecipitates from plants infected with parental TuMV, although the overall population of TuMV-AS9-derived siRNAs was depleted relative to the input fraction (Fig. 4A panels I and II, Fig. 4B and 4C panel III, and S4 Fig). Only a few individual sequences were enriched; these sequences had predominantly a 5'U nucleotide (Fig. 4A panel III). Because depletion of TuMV-AS9-derived siRNAs in HA-AGO1_{DAH} immunoprecipitates was 60 to 1,200 fold lower than in TuMV-infected samples, we reasoned that AGO1 does interact with virus-derived siRNAs, but to a lesser extent than both AGO2 and AGO10.

HC-Pro associates with siRNAs derived from the entire TuMV genome

Results described above show that AGO1, AGO2 and AGO10 associate at low levels with parental TuMV-derived siRNAs. In contrast, AGO2 and AGO10, and to a much lesser extent AGO1, associate with siRNAs derived from the suppressor-deficient TuMV-AS9 genome. Only two residues (R238A and V240A) in HC-Pro differ between TuMV and TuMV-AS9 (Fig. 6A panel I) [23, 38]. We hypothesized that i) HC-Pro associates with siRNAs-derived from the entire TuMV genome and sequesters them from AGO proteins, and ii) the AS9 mutation in HC-Pro reduces siRNA-binding activity. HC-Pro is known to have small RNA-binding activity [39, 43, 44, 55], but the extent to which it binds siRNAs in the context of TuMV infection has not been described. To measure the extent to which HC-Pro binds small RNA using the immunoprecipitation assay, we introduced an N-terminal 6xHistidine tag (HIS₆) in the context of the TuMV (TuMV-HIS) and TuMV-AS9 (TuMV-HIS-AS9) genomes (Fig. 6A panel I). The addition of HIS₆ to HC-Pro did not affect viral coat protein accumulation (Fig. 6A panel II), but enabled specific immunoprecipitation of HC-Pro from plants infected with TuMV-HIS and TuMV-HIS-AS9 (Fig. 6B).

Small RNAs from input and immunoprecipitated fractions obtained from plants inoculated with TuMV-HIS and TuMV-HIS-AS9 were sequenced. Because TuMV-HIS-AS9 accumulated more slowly than TuMV-HIS, TuMV-HIS samples were collected earlier than TuMV-HIS-AS9 samples (10 and 15 dpi, respectively), and twice as much input and immunoprecipitate materials for TuMV-HIS-AS9 samples were analyzed. The longer infection time and doubling of materials for TuMV-HIS-AS9 resulted in similar protein levels for HIS-HC-Pro and HIS-HC-Pro-AS9 input and immunoprecipitate fractions (Fig. 6B).

Endogenous *A. thaliana* small RNAs were depleted in suppressor-deficient HC-Pro-AS9 immunoprecipitates. Similarly, 22-, 23- and 24-nt *A. thaliana* endogenous small RNAs were depleted in wild-type HC-Pro immunoprecipitates (S6A Fig). In samples from systemically infected inflorescence or cauline leaves, *A. thaliana* endogenous 21-nt small RNAs were marginally enriched (2 fold) or depleted, respectively, in wild-type HC-Pro immunoprecipitates (S6A Fig). While miRNAs were depleted, miRNA* and tasiRNAs were enriched in HC-Pro immunoprecipitates (S6B–S6C Fig). Specifically, reads corresponding to miR390 and miR390*

were enriched 8 and 64 fold, respectively, in wild-type HC-Pro immunoprecipitates. MiR166 reads were depleted 5 fold, whereas miR166* reads were enriched 16 fold in wild-type HC-Pro immunoprecipitates.

In contrast with results obtained for HA-AGO1_{DAH}, HA-AGO2_{DAD} and HA-AGO10_{DDH} from TuMV-infected plants (compare panel I in Fig. 3C–5C to Fig. 6E), TuMV-derived siRNAs were highly enriched in HIS-HC-Pro immunoprecipitates from cauline leaves and inflorescence (Fig. 6C panels I and II, and Fig. 6D and 6E panels I and II). No 5' nt preference was evident (Fig. 6C panels III and IV). HIS-HC-Pro associated preferentially with 21-nt over 22-nt siRNAs in samples from both cauline leaves and inflorescences (Fig. 6C, 6D–E panels I and II, and S7 Fig). In contrast, TuMV-HIS-AS9-derived siRNAs from across the genome were depleted in the HIS-HC-Pro-AS9 immunoprecipitates from systemically infected cauline leaves; only a few individual sequences were enriched (Fig. 6C panels I and II, 6D and 6E panel III, and S7 Fig). These results indicate that wild-type HC-Pro associates with TuMV-derived siRNAs, and that the AS9 mutation disrupts this association. We concluded that HC-Pro interferes with antiviral silencing, at least in part, by sequestering TuMV-derived siRNAs and preventing their association with antiviral AGO proteins. Suppression activity of HC-Pro is not tissue specific and affects AGO1, AGO2, AGO10 and possibly other AGO proteins.

Discussion

Genetic and co-immunoprecipitation analyses were combined to reveal that i) several AGOs function as anti-TuMV defense modules in *A. thaliana*, ii) viral siRNAs generally fail to load into AGO proteins with antiviral functions during wild-type TuMV infection, and iii) HC-Pro sequesters viral siRNA away from AGOs with antiviral functions.

Functions of AGO-small RNA complexes in anti-TuMV defense

AGO proteins target endogenous transcripts to regulate plant development and innate immunity [2, 56], which may indirectly affect susceptibility to viruses. It is likely, however, that at least some AGO proteins with an antiviral role are programmed with virus-derived siRNA to directly target viral RNA [8, 10, 57, 58]. The genetic analysis described here revealed several AGO proteins that participate in modular fashion during anti-TuMV defense (Fig. 7). AGO2 has the most influential role in protecting inoculated rosette and cauline leaves (Fig. 1), while AGO1 and AGO10 have genetically redundant roles in protecting inflorescence tissues. A larger proportion of *ago1 ago2 ago10* triple mutants than *ago1 ago10* double mutants were systemically infected (Table 3), perhaps suggesting that AGO2 also contributes to restricting virus spread to inflorescences.

The antiviral effects of different AGO proteins in different tissues may depend on a number of factors, including expression patterns, AGO-interacting partners, small RNA binding preferences, or subcellular localization. Microarray data suggest that *AGO10* and *AGO1* are expressed more strongly than *AGO2* in flowers and meristems [59]. However, *AGO1* and *AGO10* transcript levels are also higher than *AGO2* transcript levels in rosette leaves. Therefore, expression levels alone do not explain the effectiveness of individual AGOs in different organs. It is conceivable that modular, tissue-specific functionality is controlled by AGO-interacting or AGO-promoting factors that are tissue-specific. In *ago1 ago10* double mutants, systemic infection of inflorescences could be partially restricted because AGO2 limits virus accumulation in leaves, acts directly in inflorescences, or functions in both of these tissues.

Direct down-regulation of viral RNA requires that AGOs bind virus-derived siRNAs (or endogenous small RNAs complementary to a given viral genome) and then viral RNA, followed by slicing of the viral RNA, repression of translation, and/or recruitment of factors for silencing

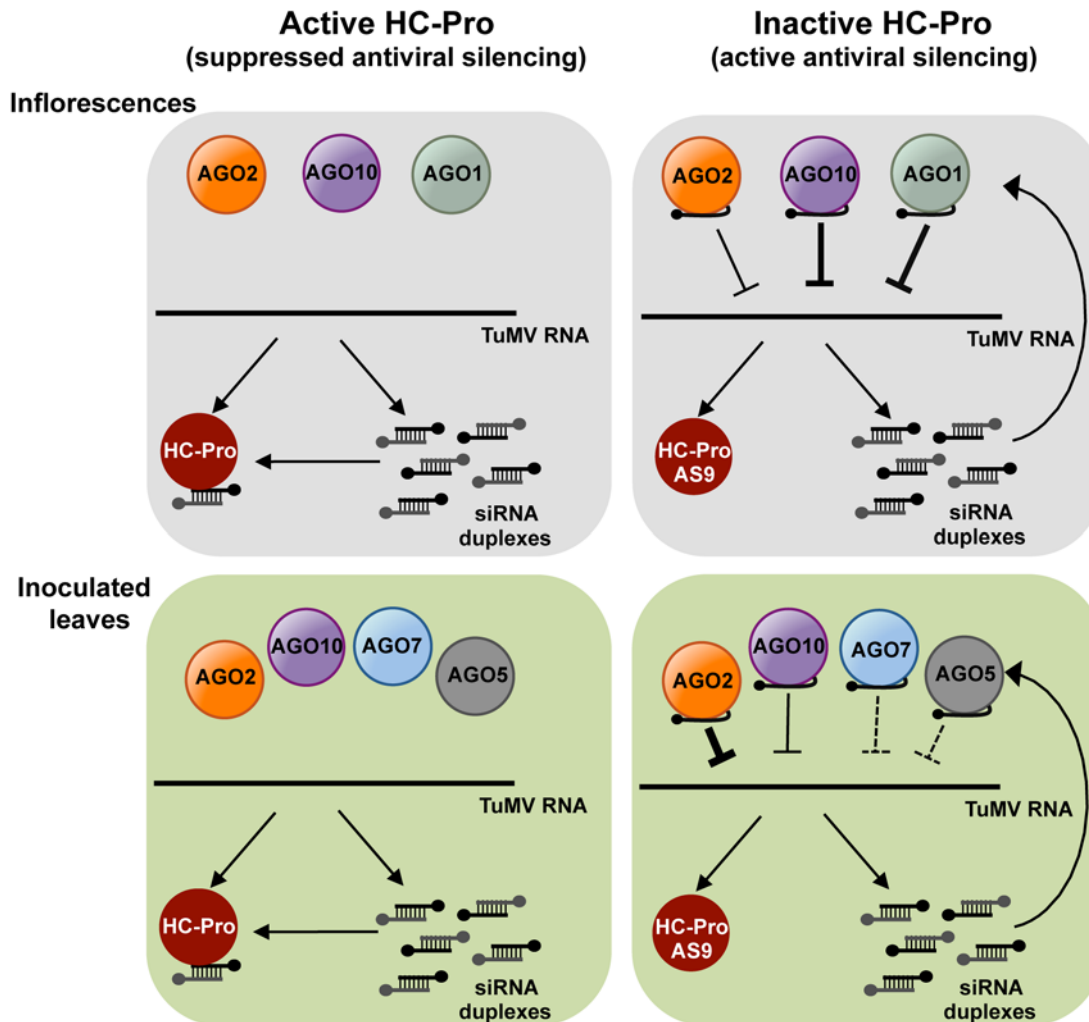


Fig 7. A model for direct action of *A. thaliana* AGO proteins in anti-TuMV defense. AGO-mediated antiviral silencing is suppressed through sequestration of TuMV-derived siRNAs by silencing suppressor HC-Pro (left panels), in both inoculated leaves and inflorescences. In the absence of active HC-Pro (right panels), AGO2, AGO10 and, to a lesser extent AGO1, associate with TuMV-AS9-derived siRNAs to potentially repress TuMV RNAs through slicing or translational repression. AGO2 protects leaves from TuMV infection and movement, with non-additive contributions by AGO10, AGO5 and AGO7. Redundant activities of AGO10 and AGO1 protect inflorescence from TuMV infection, with an additive contribution by AGO2.

doi:10.1371/journal.ppat.1004755.g007

amplification. Results described here show that AGO2, AGO10 and at much lower levels AGO1 associate with TuMV-AS9-derived siRNA in the absence of HC-Pro (Fig. 3C panels III and IV, and Figs. 4C and 5C panel III). AGO2-mediated slicing of viral RNAs could be a significant anti-viral mechanism, as catalytically defective forms of AGO2 lack anti-TuMV activity [31]. Evidence of direct targeting of TuMV RNA by AGO1 and AGO10 is lacking. In other studies, AGO1 was reported to bind small RNAs derived from *Turnip yellow mosaic virus* and CMV strains Fny and NT9 [20], but not CMV strain I17F or *Crucifer-infecting tobamovirus* [60]. The basis for differential interaction of TuMV-derived siRNAs and AGO1, AGO2 and AGO10 is not clear. It is possible that different AGOs have privileged access to viral siRNAs. In this context, AGO1 pools may have limited access to viral siRNAs during TuMV infection.

In inoculated rosette leaves of *ago2* mutant and *dcl2 dcl3 dcl4* triple mutant plants, TuMV-AS9 accumulated to comparable levels (Figs. 1 and 2). In contrast, accumulation of TuMV-AS9 was consistently lower in cauline leaves and inflorescences of all *ago* mutants tested,

including the *ago1 ago2 ago10* triple mutant, compared to the respective tissues in *dcl2 dcl3 dcl4* mutant plants. If it is assumed that all small RNA-mediated antiviral activity is lost in the *dcl* triple mutant, then it is reasonable to conclude that all antiviral silencing in inoculated rosette leaves is mediated by AGO2. The far greater effect of the *dcl* mutations, relative to the *ago* mutations, in systemic tissues, especially inflorescences, argues that the combined effects of AGO1, AGO2 and AGO10 account for only a small proportion of overall anti-TuMV silencing activity. This could indicate that other AGO proteins that were not tested here, or that were not tested in the right genetic combinations, play specific roles in systemic tissues. It could also mean that DCL proteins play a more dominant, direct antiviral role in systemic tissues, as suggested by genetic analyses with CMV [4, 29], BMV [30], PVX [27], *Tobacco rattle virus* [61], TCV [3, 6, 33, 62], *Cauliflower mosaic virus*, *Cabbage leaf curl virus*, and *Oil rape mosaic virus* [63].

Different antiviral AGO proteins may also have distinct effects on amplification of secondary, virus-derived siRNAs, which may be important for production of systemic signals [2, 7, 13, 64]. Full anti-TuMV silencing requires both RDR1 and RDR6 [23], presumably for production of dsRNA from viral RNA. If this occurs like dsRNA formation during tasiRNA biogenesis, then RDR proteins may be recruited to viral RNA after targeting by AGO-small RNA complexes [52, 65–68]. Given the role of AGO1-small RNA complexes in triggering formation of several families of tasiRNA, AGO1 could conceivably play a trigger role for secondary viral siRNA.

The interpretation of *ago1* mutant susceptibility experiments is challenging because of the pleiotropic developmental phenotypes of *ago1* hypomorphic mutants and the large number of genes that are dysregulated when AGO1 is disrupted. In particular, disruption of AGO1-miR403 activity increases *AGO2* mRNA and protein levels [26, 69], which could result in a net increase in virus resistance, even if AGO1 directly targets viral RNA.

Other AGOs might also have indirect roles in anti-TuMV defense, perhaps by affecting expression of defense-related genes [35, 56, 70]. Expression of potyviral HC-Pro [45], infection with TCV [26], and infection with *Pseudomonas syringae* [35] result in increased AGO2 expression; AGO2 regulates expression of *MEMB12* [35] and possibly other genes. AGO2 also associates with virus-activated endogenous siRNAs [56]. The significance of AGO2-dependent gene regulation for virus infection, if any, is not yet clear.

Suppression of antiviral silencing by HC-Pro

Multiple virus-encoded suppressors of RNA silencing target AGO1 [16, 17, 20, 21, 33, 60], and P25 from PVX interact with AGO2, AGO3 and AGO4 [17] although the biological significance of this interaction remains to be elucidated. During TuMV infection, no evidence was obtained to indicate that AGO1, AGO2 or AGO10 were destabilized or otherwise down-regulated. Each AGO accumulated to normal levels.

TuMV-infected plants accumulate large amounts of virus-derived siRNAs that map across the entire genome (Figs. 3B, 4B, 5B, 6B, and S3–S5 Figs) [23], and co-immunoprecipitation and high-throughput sequencing showed that HC-Pro associates with viral siRNAs in leaf and inflorescence tissue (Fig. 6E panels I and II). Viral siRNAs associate with HC-Pro without a 5' nt preference (Fig. 6C panels III and IV). Importantly, HC-Pro was shown to sequester viral siRNAs away from AGO1, AGO2 and AGO10 (Figs. 3C, 4C and 5C panels I and II), leading to the obvious proposal that HC-Pro interferes with antiviral silencing by preventing AGOs from loading with virus-derived siRNAs (Fig. 7). Mutant HC-Pro-AS9 is deficient in associating with viral siRNAs (Fig. 6C–E panels III, and S6 Fig), and concomitantly loses silencing suppression activity.

The basis for sequestration of siRNAs by HC-Pro is not yet clear. HC-Pro may outcompete AGOs for siRNAs. Alternatively, HC-Pro may intercept viral siRNAs prior to AGO loading, perhaps due to subcellular localization properties. Further analyses will be necessary to resolve this issue.

Materials and Methods

DNA plasmids

Recombinant plasmids were made as follows.

pCB-TuMV-HIS and pCB-TuMV-HIS-AS9. To introduce a 6xHIS (HIS₆) tag on HC-Pro, two PCR fragments were amplified from pCB-TUMV [16] using two sets of primers: TuMV764 d(AGGACGGTGCACAGAATATGC) and E101-B2Rev d(CCAGAAGTTGG CTCCTGCTGCGTGATGGTGATGGTGATGACCTGCCTGGTGATAGACACAGCTAGC ACTAAAGTGCAC); and E101-B2For d(GTGCACTTTAGTGCTAGCTGTGTCTATCACC AGGCAGGTCATCACCATCACCATCACGCAGCAGGAGCCAACTTCTGG) and TuMV-GFP-2873 d(CGCCTGATTCTGTTGTGACAC). The two PCR fragments were stitched into a final PCR product using primers TuMV764 and TuMV-GFP-2873. The final PCR product was digested with StuI-AgeI and used to replace the StuI-AgeI fragment in pCB-TuMV, creating pCB-TuMV-HIS. The same insert was used to replace the StuI-AgeI fragment in pCB-TuMV-AS9 [16], to generate pCB-TuMV-HIS-AS9. Both HIS₆-tagged clones have a Nla cleavage site between P1 and the HIS₆-tag on HC-Pro.

pMDC99-pAGO10:3xHA-AGO10_{DDH} and pMDC99-pAGO10:3xHA-AGO10_{DAH}. For in-frame N-terminal 3xHA-tagging of wild-type AGO10_{DDH} in its natural genomic context, a 9072 bp genomic region was TOPO cloned into pENTR (Invitrogen) in two pieces: an upstream region (with primers caccGATTCTATAAAAAATAcattcc and CTCGAGGCG GCCGCCATGGTTTTTGTGTTTGGATTTTC) and the coding and downstream regions (with HA-containing forward primer caccATGGCCTATCCTTATGATGTACCTGATTATG CCTACCCATACGACGTTCCAGACTACGCTTACCCATACGACGTTCCAGACTACG CTCCGATTAGGCAAATGAAAGATAG and reverse primer cctagaattgacgggttagatcg). The first piece was ligated upstream of the second using a *NotI* site in pENTR and a *NcoI* site created by the cloning primers, producing pENTR-pAGO10-3xHA-AGO10_{DDH}. To disrupt the AGO10 PIWI domain catalytic triad, A2384 in the coding sequence of pENTR-pAGO10:3xHA-AGO10_{DDH} was mutated to G by GENEWIZ Inc., causing amino acid substitution D795A to generate pENTR-pAGO10-3xHA-AGO10_{DAH}. Transgenes from pENTR-pAGO10-3xHA-AGO10_{DDH} and pENTR-pAGO10-3xHA-AGO10_{DAH} were LR recombined into binary vector pMDC99 [71], producing pMDC99-pAGO10:3xHA-AGO10_{DDH} and pMDC99-pAGO10:3xHA-AGO10_{DAH}, respectively.

Plant materials

All *Arabidopsis thaliana* plants used in this study (including mutant lines and transgenic lines) descended from the Columbia-0 (Col-0) accession, and were grown under long day (16 h light/ 8 h dark) at 22°C. The following single mutant lines were described before: *ago1-25* and *ago1-27* [25], *ago2-1* [72], *ago3-2* [32], *ago4-2* [73], *ago5-2* [32], *ago6-3* [32], *zip-1* [74], *ago8-1* [32], and *ago9-5* (SALK_126176). T-DNA insertion mutant GABI_818H06 (*ago10-5*) was obtained from The GABI KAT project [75]. Homozygous mutants were confirmed by PCR-based genotyping using a three-primer reaction: one on the left border, one in the flanking DNA, and one in the T-DNA insertion site [76]. Lack of AGO10 expression in homozygous plants was confirmed by RT-PCR using oligos AGO10_qF (GGTATTCAGGGAACAAGCAG) and

AGO10_{qR} (GCTGGAGGAAGTATAGAGACCG). Double and triple *ago* mutants were generated by crossing. *dcl2-1 dcl3-1 dcl4-2* triple mutants have been described [3].

Transgenic *A. thaliana* plants expressing HA-tagged AGO1 or AGO2 catalytic mutants from their native promoters have been described [31]. Transgenic *A. thaliana* plants expressing HA-tagged wild-type or catalytic mutant AGO10 from its native promoter were made by dipping Col-0 plants in *Agrobacterium tumefaciens* GV3101 carrying the pMDC99-pAGO10:3xHA-AGO10_{DDH} or pMDC99-pAGO10:3xHA-AGO10_{DAH} constructs as described [77]. Transgenic plants were grown on MS medium containing hygromycin (50 mg/ml) for 7 days, transferred to soil, and maintained in greenhouse conditions. Catalytic mutant HA-AGO1_{DAH}, HA-AGO2_{DAD} and HA-AGO10_{DAH} and wild-type HA-AGO10_{DDH} transgenes were introduced into *ago2-1* by crossing.

Virus infection assays

A. thaliana plants were inoculated with TuMV-GFP, TuMV-AS9-GFP, wild-type TuMV, TuMV-AS9, TuMV-HIS, or TuMV-HIS-AS9 as described previously [23]. Local and systemic infection by TuMV-GFP or TuMV-AS9-GFP was determined by GFP fluorescence under UV illumination. To measure coat protein (CP) or HIS₆-tagged HC-Pro (HIS-HC-Pro) accumulation, at 15 days post inoculation (dpi), four noninoculated cauline leaves or five inflorescence clusters per plant were randomly collected and pooled into a single sample. Four biological replicates were randomly collected per virus-plant genotype combination. Samples were ground in glycine buffer [78] at a ratio of 0.5 mL per 1g of leaf, or 0.25 mL per five inflorescence clusters. Protein extracts were normalized to 0.5 mg/mL. For western blot assays, 6.25 µg or 1.5 µg of total protein were used for leaf or inflorescence samples, respectively. Immunoblotting and chemiluminescence detection were done as described [23]. TuMV CP was detected using antibody PVAS-134 (1:40,000) and HIS-HC-Pro was detected using anti-HIS antibody 27E8-HRP (Cell Signaling) at a 1:5,000 dilution. Ponceau staining of the large subunit of rubisco was used as a loading control. Unless otherwise indicated, CP and HIS-HC-Pro were detected simultaneously on the same blot. In experiments involving HA-tagged AGOs, HA-AGO, CP and HIS-HC-Pro were detected on the same blot. The top part of the blot, containing proteins larger than 70 kDa was incubated with anti-HA antibodies, to detect HA-AGOs. The part of the blot containing proteins between 70 and 27 kDa was probed for CP and HIS-HC-Pro.

Immunoprecipitation of HA-tagged ARGONAUTES and HIS-tagged HC-Pro

Immunoprecipitation of epitope tagged proteins was performed as described [31] with minor adjustments. Briefly, one gram of leaf or inflorescence tissue was ground in 6 ml of lysis buffer. Lysates were pre-cleared by incubating with protein A agarose (Roche) beads (0.8 mL per 1g of tissue) for 30 min at 4°C, and beads were not treated with P1 nuclease. For immunoblot detection of proteins (CP, HA-AGOs or HIS-HC-Pro), 6.25 µg or 1.5 µg of total protein from leaf or inflorescence samples were used, respectively. From the immunoprecipitated beads 5% of the samples was diluted with 38 µl of 2x protein dissociation buffer, and 5 to 15 µL used for immunoblotting. For small RNA northern blotting, 15 µg were used from the input fractions and 25% of the RNA immunoprecipitate fraction (HA or HIS).

Small RNA library construction for high-throughput sequencing

Small RNA libraries from mock-inoculated or TuMV-infected plants, input or immunoprecipitate (HA or HIS) fractions were generated using sequencing-by synthesis technology (Illumina High Seq 2000) as described [31, 79]. For input fractions, 50 µg of total RNA were fractionated

by electrophoresis. The area from 16 to 26 nt was sliced and used for small RNA purification. 30 ng of small RNAs were used to make the libraries from total fraction. 50% of the immunoprecipitated RNA was used without fractionation to make libraries from immunoprecipitate fractions. For each treatment, small RNA libraries were made independently from two biological replicates. Bar-coded PCR amplification primers were used for multiplexing purposes. Eight individual samples were multiplexed and run in a single flow cell.

Bioinformatic analysis of small RNA libraries

Bioinformatic analysis of endogenous and TuMV-derived siRNAs was as described [23, 31, 80]. After removing 5' and 3' adaptors, sequences were aligned to the *A. thaliana* genome and to the TuMV genome. Only sequences with a perfect match were used for downstream analysis. For each sample, reads were normalized per 1,000,000 total reads (RPM), including all size classes. Enrichment with respect to the immunoprecipitate was calculated as the ratio of reads in the immunoprecipitate to reads in the input, and expressed on a log₂ scale.

Accession numbers

Sequence data from this article can be found in Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>) accession number GSE64911.

Supporting Information

S1 Fig. Proportion of *A. thaliana* endogenous and TuMV-derived small RNAs in mock-inoculated and in TuMV-infected plants. Samples for immunoprecipitation were collected from inflorescence 10 (dpi), rosette leaves (7 dpi), or cauline leaves (15 dpi). Numbers are the relative abundance, in percentage, of reads mapping to *A. thaliana* or to TuMV with respect to the total number of reads with a perfect match to either genome. Proportion of TuMV-derived siRNAs by size class is indicated by numbers (percentage) in color pie charts. Numbers were rounded to the nearest integer. Plants expressing (A) HA-AGO2_{DAD}, (B) HA-AGO1_{DAH} from an *ago2-1* background and were inoculated with wild-type TuMV or TuMV-AS9. (C) HA-AGO10_{DDH} or HA-AGO10_{DAD} were expressed from a AGO2 or *ago2-1* background, respectively. (D) Wild-type Col-0 or single *ago2-1* mutant plants were inoculated with TuMV-HIS or TuMV-HIS-AS9. Color codes are as in (A). (TIF)

S2 Fig. Association of endogenous siRNAs with HA-tagged AGO1, AGO2 and AGO10. Values are average and SE from two biological replicates normalized to reads per million. Inoculated rosette leaf, systemically infected cauline leaves or inflorescence samples were collected at 7, 15 or 10 dpi, respectively. (A) HA-AGO2_{DAD} in an *ago2-1* background. Panel I: enrichment [immunoprecipitate (IP) reads/ input reads, expressed in a log₂ scale] of endogenous (21 to 24 nt) small RNAs in mock-inoculated plants and in plants infected with wild-type TuMV or TuMV-AS9. In the scale was capped at 4 and at -4. Panel II: proportion (in percentage) of 5' nt in 21 nt and 22 nt small RNAs in input and in HA-AGO2_{DAD} immunoprecipitated (IP) fractions. Numbers were rounded to the nearest integer. (B) HA-AGO1_{DAH} in an *ago2-1* background. Labels for panels I and II are as in (A). (C) Catalytically active HA-AGO10_{DDH} and catalytic mutant HA-AGO10_{DAH} were expressed in a wild-type Col-0 (AGO2) or *ago2-1* background, respectively. Labels for panels I and II are as in (A). (TIF)

S3 Fig. TuMV genome-wide distribution and enrichment of 22-nt TuMV-derived siRNAs in plants expressing HA-AGO2_{DAD} in an *ago2-1* background. Values are average and SE

from two biological replicates normalized to reads per million. Scale was capped at 150. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively. Inflorescence samples were collected at 10 dpi. (A) and (B) TuMV genome-wide distribution of 22 nt TuMV-derived siRNAs in input (A) and in HA-AGO2_{DAD} immunoprecipitated (IP) fractions (B). Scale was capped at 150.

(TIF)

S4 Fig. TuMV genome-wide distribution and enrichment of 22-nt TuMV-derived siRNAs in plants expressing HA-AGO1_{DAH} in an *ago2-1* background. Values are average and SE from two replicates normalized to reads per million. Inflorescence samples were collected at 10 dpi. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively. (A) and (B) TuMV genome-wide distribution of 22 nt TuMV-derived siRNAs in input (A) and in HA-AGO1_{DAH} immunoprecipitated fractions (IP) (B). Scale was capped at 150.

(TIF)

S5 Fig. TuMV genome-wide distribution and enrichment of 22-nt TuMV-derived siRNAs in plants expressing HA-AGO10_{DDH} or HA-AGO10_{DAH}. Values are average and SE from two replicates normalized to reads per million. Inflorescence samples were collected at 10 dpi. Inoculated rosette leaf and systemically infected cauline leaf samples were collected at 7 and 15 dpi, respectively. (A) and (B) TuMV genome-wide distribution of 22 nt TuMV-derived siRNAs in input (A) and in HA-AGO10 immunoprecipitated (IP) fractions (B). Scale was capped at 150.

(TIF)

S6 Fig. Association of endogenous siRNAs (21–24-nt) with HC-Pro in plants infected with TuMV-HIS or TuMV-HIS-AS9. Values are average and SE from two biological replicates normalized to reads per million. Inflorescence and cauline leaf samples from plants infected with TuMV-HIS were collected at 10 dpi. Cauline leaf samples from plants infected with TuMV-HIS-AS9 were collected at 15 dpi. (A) Number of reads of endogenous *A. thaliana* siRNAs by size class in input and HC-Pro immunoprecipitated (IP) fractions from inflorescence and cauline leaves. (B) Number of reads for miRNAs, miRNA* and tasiRNAs in input and mock or HC-Pro IP. (C) Enrichment (IP reads/ Input reads, expressed in a log₂ scale) of miRNAs, miRNA* and tasiRNAs (TAS) in mock or HC-Pro IP. Scales was capped at 3 and -3.

(TIF)

S7 Fig. TuMV genome-wide distribution and enrichment of 22-nt TuMV-derived siRNAs in Col-0 or *ago2-1* plants infected with TuMV-HIS or TuMV-HIS-AS9. Values are average and SE from two biological replicates normalized to reads per million. Scale was capped at 500. Inflorescence samples were from Col-0 plants at 10 dpi. Cauline leaf samples were from single *ago2-1* mutant plants infected with TuMV-HIS or TuMV-HIS-AS9 at 10 or 15 dpi, respectively. (A) and (B) TuMV genome-wide distribution of 22 nt TuMV-derived siRNAs in input (A) or immunoprecipitated (IP) fractions of wild-type or AS9 HC-Pro.

(TIF)

S8 Fig. Association of *A. thaliana* miRNAs, miRNA* and tasiRNAs with HA-tagged AGO2_{DAD}, AGO1_{DAH}, AGO10_{DDH} or AGO10_{DAH}. Transgenic HA-AGO1_{DAH} and HA-AGO2_{DAD} were expressed from an *ago2-1* background. Transgenic HA-AGO10_{DDH} and HA-AGO10_{DAH} were expressed from a wild-type Col-0 (AGO2) or an *ago2-1* background, respectively. Plants were mock-inoculated or infected with TuMV or with TuMV-AS9. Rosette leaf and samples were collected at 7 dpi. Cauline leaf and inflorescence samples were collected at 15 and 10 dpi, respectively. Values are average and SE from two biological replicates. The

histograms show average fold enrichment in AGO IP (IP reads/ input reads, expressed in log₂ scale) of miRNAs, miRNA* and tasiRNAs. A) HA-AGO2_{DAD} IP. B) HA-AGO1_{DAH} IP, and C) HA-AGO10_{DDH} or HA-AGO10_{DAH} IP. (TIF)

S9 Fig. Association of HA-AGO1_{DAH} and HA-AGO2_{DAD} with endogenous and virus-derived siRNAs. Blots show accumulation of CP, HA-AGO, and virus-derived small RNAs in immunoprecipitation (IP) fractions of HA-AGO1_{DAH} and HA-AGO2_{DAD} from cauline leaves (1g) at 15 dpi. HA-AGO1_{DAH} and HA-AGO2_{DAD} were expressed from transgenic *ago2-1* plants. Mock-inoculated plants and non-transgenic single *ago2-1* mutants were used as controls. Representative blots showing accumulation of HA-AGOs, CP, TuMV-derived siRNAs (CI) and selected miRNAs in input and HA-AGO immunoprecipitation fractions (IPs). TuMV CP and HA-AGO were detected by immunoblotting in input and IP fractions. TuMV-derived siRNAs were detected with a DIG-labeled probe made by random priming of cDNA corresponding to CI. miR390 and miR168 were used as IP controls, and U6 as loading control. Endogenous siRNAs were detected with DIG-labeled oligonucleotides. Duplicated blots were stripped and re-probed. A) IP of HA-AGO1_{DAH} and HA-AGO2_{DAD} from cauline leaves of plants infected with wt TuMV. Panel I: protein accumulation in input samples. Panel II: protein accumulation in IP fractions. B) IP of HA-AGO1_{DAH} and HA-AGO2_{DAD} from cauline leaves of plants infected with suppressor-deficient TuMV-AS9. Panels I and II are as in (A). (TIF)

S1 Table. Abundance of endogenous *A. thaliana* and TuMV-derived small RNAs of all size classes in input and HA-AGO2 immunoprecipitation fractions. (DOCX)

S2 Table. Abundance of endogenous *A. thaliana* and TuMV-derived small RNAs of all size classes in input and AGO1 immunoprecipitation fractions. (DOCX)

S3 Table. Abundance of endogenous *A. thaliana* and TuMV-derived small RNAs of all size classes in input and AGO10 immunoprecipitation fractions. (DOCX)

S4 Table. Abundance of endogenous *A. thaliana* and TuMV-derived small RNAs of all size classes in input and HC-Pro immunoprecipitation fractions. (DOCX)

S1 Dataset. MicroRNA read counts for input and immunoprecipitates of HA-AGO2, AGO1, AGO10 and HC-Pro. (XLSX)

Acknowledgments

We thank Goretti Nguyen and Robyn Stevens for excellent technical assistance. We thank Pablo Vera, the Salk Institute Genomic Analysis Laboratory, the Arabidopsis Biological Resource Center, and the GABI-KAT project for generating and distributing mutant seed.

Author Contributions

Conceived and designed the experiments: HGR JCC. Performed the experiments: HGR JSH AT AG MTGR MGM NL MTMB. Analyzed the data: HGR NF AC JSH JCC. Contributed

reagents/materials/analysis tools: HGR AT AC NF JSH KBG JCC. Wrote the paper: HGR AC NF JSH JCC.

References

1. Ding SW, Voinnet O. Antiviral immunity directed by small RNAs. *Cell*. 2007; 130(3):413–26. doi: [10.1016/j.cell.2007.07.039](https://doi.org/10.1016/j.cell.2007.07.039) PMID: [17693253](https://pubmed.ncbi.nlm.nih.gov/17693253/)
2. Pumplin N, Voinnet O. RNA silencing suppression by plant pathogens: defence, counter- defence and counter-counter-defence. *Nat Rev Microbiol*. 2013; 11(11):745–60. doi: [10.1038/nrmicro3120](https://doi.org/10.1038/nrmicro3120) PMID: [24129510](https://pubmed.ncbi.nlm.nih.gov/24129510/)
3. Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science*. 2006; 313(5783):68–71. doi: [10.1126/science.1128214](https://doi.org/10.1126/science.1128214) PMID: [16741077](https://pubmed.ncbi.nlm.nih.gov/16741077/)
4. Diaz-Pendon JA, Li F, Li WX, Ding SW. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell*. 2007; 19(6):2053–63. doi: [10.1105/tpc.106.047449](https://doi.org/10.1105/tpc.106.047449) PMID: [17586651](https://pubmed.ncbi.nlm.nih.gov/17586651/)
5. Curtin SJ, Watson JM, Smith NA, Eamens AL, Blanchard CL, Waterhouse PM. The roles of plant dsRNA-binding proteins in RNAi-like pathways. *FEBS Lett*. 2008; 582(18):2753–60. doi: [10.1016/j.febslet.2008.07.004](https://doi.org/10.1016/j.febslet.2008.07.004) PMID: [18625233](https://pubmed.ncbi.nlm.nih.gov/18625233/)
6. Qu F, Ye X, Morris TJ. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc Natl Acad Sci U S A*. 2008; 105(38):14732–7. doi: [10.1073/pnas.0805760105](https://doi.org/10.1073/pnas.0805760105) PMID: [18799732](https://pubmed.ncbi.nlm.nih.gov/18799732/)
7. Bologna NG, Voinnet O. The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annu Rev Plant Biol*. 2014; 65:473–503. doi: [10.1146/annurev-arplant-050213-035728](https://doi.org/10.1146/annurev-arplant-050213-035728) PMID: [24579988](https://pubmed.ncbi.nlm.nih.gov/24579988/)
8. Schuck J, Gursinsky T, Pantaleo V, Burgyan J, Behrens SE. AGO/RISC-mediated antiviral RNA silencing in a plant in vitro system. *Nucleic Acids Res*. 2013; 41(9):5090–103. doi: [10.1093/nar/gkt193](https://doi.org/10.1093/nar/gkt193) PMID: [23535144](https://pubmed.ncbi.nlm.nih.gov/23535144/)
9. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, et al. Widespread translational inhibition by plant miRNAs and siRNAs. *Science*. 2008; 320(5880):1185–90. doi: [10.1126/science.1159151](https://doi.org/10.1126/science.1159151) PMID: [18483398](https://pubmed.ncbi.nlm.nih.gov/18483398/)
10. Ciomperlik JJ, Omarov RT, Scholthof HB. An antiviral RISC isolated from Tobacco rattle virus-infected plants. *Virology*. 2011; 412(1):117–24. doi: [10.1016/j.virol.2010.12.018](https://doi.org/10.1016/j.virol.2010.12.018) PMID: [21272908](https://pubmed.ncbi.nlm.nih.gov/21272908/)
11. Iwakawa HO, Tomari Y. Molecular Insights into microRNA-Mediated Translational Repression in Plants. *Mol Cell*. 2013; 52(4):591–601. Epub 2013/11/26. doi: [10.1016/j.molcel.2013.10.033](https://doi.org/10.1016/j.molcel.2013.10.033) PMID: [24267452](https://pubmed.ncbi.nlm.nih.gov/24267452/)
12. Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature reviews Genetics*. 2011; 12(2):99–110. doi: [10.1038/nrg2936](https://doi.org/10.1038/nrg2936) PMID: [21245828](https://pubmed.ncbi.nlm.nih.gov/21245828/)
13. Szittya G, Burgyan J. RNA interference-mediated intrinsic antiviral immunity in plants. *Current topics in microbiology and immunology*. 2013; 371:153–81. doi: [10.1007/978-3-642-37765-5_6](https://doi.org/10.1007/978-3-642-37765-5_6) PMID: [23686235](https://pubmed.ncbi.nlm.nih.gov/23686235/)
14. Incarbone M, Dunoyer P. RNA silencing and its suppression: novel insights from in planta analyses. *Trends Plant Sci*. 2013; 18(7):382–92. doi: [10.1016/j.tplants.2013.04.001](https://doi.org/10.1016/j.tplants.2013.04.001) PMID: [23684690](https://pubmed.ncbi.nlm.nih.gov/23684690/)
15. Baumberger N, Tsai CH, Lie M, Havecker E, Baulcombe DC. The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. *Curr Biol*. 2007; 17(18):1609–14. doi: [10.1016/j.cub.2007.08.039](https://doi.org/10.1016/j.cub.2007.08.039) PMID: [17869110](https://pubmed.ncbi.nlm.nih.gov/17869110/)
16. Bortolamiol D, Pazhouhandeh M, Marrocco K, Genschik P, Ziegler-Graff V. The Polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. *Curr Biol*. 2007; 17(18):1615–21. doi: [10.1016/j.cub.2007.07.061](https://doi.org/10.1016/j.cub.2007.07.061) PMID: [17869109](https://pubmed.ncbi.nlm.nih.gov/17869109/)
17. Chiu MH, Chen IH, Baulcombe DC, Tsai CH. The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway. *Mol Plant Pathol*. 2010; 11(5):641–9. doi: [10.1111/j.1364-3703.2010.00634.x](https://doi.org/10.1111/j.1364-3703.2010.00634.x) PMID: [20696002](https://pubmed.ncbi.nlm.nih.gov/20696002/)
18. Csorba T, Lozsa R, Hutvagner G, Burgyan J. Polerovirus protein P0 prevents the assembly of small RNA-containing RISC complexes and leads to degradation of ARGONAUTE1. *Plant J*. 2010; 62(3):463–72. doi: [10.1111/j.1365-3113.2010.04163.x](https://doi.org/10.1111/j.1365-3113.2010.04163.x) PMID: [20128884](https://pubmed.ncbi.nlm.nih.gov/20128884/)
19. Derrien B, Baumberger N, Schepetilnikov M, Viotti C, De Cillia J, Ziegler-Graff V, et al. Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. *Proc Natl Acad Sci U S A*. 2012; 109(39):15942–6. doi: [10.1073/pnas.1209487109](https://doi.org/10.1073/pnas.1209487109) PMID: [23019378](https://pubmed.ncbi.nlm.nih.gov/23019378/)

20. Zhang X, Yuan YR, Pei Y, Lin SS, Tuschi T, Patel DJ, et al. Cucumber mosaic virus- encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. *Genes Dev.* 2006; 20(23):3255–68. [10.1101/gad.1495506](https://doi.org/10.1101/gad.1495506) PMID: [17158744](https://pubmed.ncbi.nlm.nih.gov/17158744/)
21. Giner A, Lakatos L, Garcia-Chapa M, Lopez-Moya JJ, Burguan J. Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs. *PLoS Pathog.* 2010; 6(7):e1000996. [10.1371/journal.ppat.1000996](https://doi.org/10.1371/journal.ppat.1000996) PMID: [20657820](https://pubmed.ncbi.nlm.nih.gov/20657820/)
22. Nakahara KS, Masuta C. Interaction between viral RNA silencing suppressors and host factors in plant immunity. *Curr Opin Plant Biol.* 2014; 20:88–95. [10.1016/j.pbi.2014.05.004](https://doi.org/10.1016/j.pbi.2014.05.004) PMID: [24875766](https://pubmed.ncbi.nlm.nih.gov/24875766/)
23. Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Bremel KJ, et al. Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *Plant Cell.* 2010; 22(2):481–96. [10.1105/tpc.109.073056](https://doi.org/10.1105/tpc.109.073056) PMID: [20190077](https://pubmed.ncbi.nlm.nih.gov/20190077/)
24. Vaucheret H. Plant ARGONAUTES. *Trends Plant Sci.* 2008; 13(7):350–8. Epub 2008/05/30. S1360–1385 (08)00138–6[pil] [10.1016/j.tplants.2008.04.007](https://doi.org/10.1016/j.tplants.2008.04.007) PMID: [18508405](https://pubmed.ncbi.nlm.nih.gov/18508405/)
25. Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, Feuerbach F, et al. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell.* 2002; 14(3):629–39. PMID: [11910010](https://pubmed.ncbi.nlm.nih.gov/11910010/)
26. Harvey JJ, Lewsey MG, Patel K, Westwood J, Heimstadt S, Carr JP, et al. An antiviral defense role of AGO2 in plants. *PLoS One.* 2011; 6(1):e14639. [10.1371/journal.pone.0014639](https://doi.org/10.1371/journal.pone.0014639) PMID: [21305057](https://pubmed.ncbi.nlm.nih.gov/21305057/)
27. Jaubert MJ, Bhattacharjee S, Mello AF, Perry KL, Moffett P. AGO2 mediates RNA silencing anti-viral defenses against Potato virus X in Arabidopsis. *Plant physiology.* 2011. Epub 2011/05/18. [10.1104/pp.111.178012](https://doi.org/10.1104/pp.111.178012)
28. Zhang X, Singh J, Li D, Qu F. Temperature-dependent survival of Turnip crinkle virus- infected arabidopsis plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. *Journal of virology.* 2012; 86(12):6847–54. Epub 2012/04/13. [10.1128/JVI.00497-12](https://doi.org/10.1128/JVI.00497-12) PMID: [22496240](https://pubmed.ncbi.nlm.nih.gov/22496240/)
29. Wang XB, Jovel J, Udornporn P, Wang Y, Wu Q, Li WX, et al. The 21-Nucleotide, but Not 22-Nucleotide, Viral Secondary Small Interfering RNAs Direct Potent Antiviral Defense by Two Cooperative Argonautes in Arabidopsis thaliana. *The Plant cell.* 2011; 23(4):1625–38. Epub 2011/04/07. [10.1105/tpc.110.082305](https://doi.org/10.1105/tpc.110.082305) PMID: [21467580](https://pubmed.ncbi.nlm.nih.gov/21467580/)
30. Dzianott A, Sztuba-Solinska J, Bujarski JJ. Mutations in the antiviral RNAi defense pathway modify Brome mosaic virus RNA recombinant profiles. *Molecular plant-microbe interactions: MPMI.* 2012; 25(1):97–106. Epub 2011/09/23. [10.1094/MPMI-05-11-0137](https://doi.org/10.1094/MPMI-05-11-0137) PMID: [21936664](https://pubmed.ncbi.nlm.nih.gov/21936664/)
31. Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, et al. Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *The Plant cell.* 2012; 24(9):3613–29. Epub 2012/10/02. [10.1105/tpc.112.099945](https://doi.org/10.1105/tpc.112.099945) PMID: [23023169](https://pubmed.ncbi.nlm.nih.gov/23023169/)
32. Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y. The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. *Plant Cell Physiol.* 2008; 49(4):493–500. [10.1093/pcp/pcn043](https://doi.org/10.1093/pcp/pcn043) PMID: [18344228](https://pubmed.ncbi.nlm.nih.gov/18344228/)
33. Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, et al. Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. *Genes Dev.* 2010; 24(9):904–15. [10.1101/gad.1908710](https://doi.org/10.1101/gad.1908710) PMID: [20439431](https://pubmed.ncbi.nlm.nih.gov/20439431/)
34. Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, et al. A role for small RNAs in DNA double-strand break repair. *Cell.* 2012; 149(1):101–12. Epub 2012/03/27. [10.1016/j.cell.2012.03.002](https://doi.org/10.1016/j.cell.2012.03.002) PMID: [22445173](https://pubmed.ncbi.nlm.nih.gov/22445173/)
35. Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang HD, et al. Arabidopsis Argonaute 2 regulates innate immunity via miRNA393 (*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Mol Cell.* 2011; 42(3):356–66. [10.1016/j.molcel.2011.04.010](https://doi.org/10.1016/j.molcel.2011.04.010) PMID: [21549312](https://pubmed.ncbi.nlm.nih.gov/21549312/)
36. Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, et al. Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell.* 2011; 145(2):242–56. [10.1016/j.cell.2011.03.024](https://doi.org/10.1016/j.cell.2011.03.024) PMID: [21496644](https://pubmed.ncbi.nlm.nih.gov/21496644/)
37. Mallory AC, Hinze A, Tucker MR, Bouche N, Gascoli V, Elmayan T, et al. Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. *PLoS Genet.* 2009; 5(9):e1000646. [10.1371/journal.pgen.1000646](https://doi.org/10.1371/journal.pgen.1000646) PMID: [19763164](https://pubmed.ncbi.nlm.nih.gov/19763164/)
38. Kasschau KD, Cronin S, Carrington JC. Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology.* 1997; 228(2):251–62. [10.1006/viro.1996.8368](https://doi.org/10.1006/viro.1996.8368) PMID: [9123832](https://pubmed.ncbi.nlm.nih.gov/9123832/)
39. Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, Liu YP, et al. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J.* 2006; 25(12):2768–80. [10.1038/sj.emboj.7601164](https://doi.org/10.1038/sj.emboj.7601164) PMID: [16724105](https://pubmed.ncbi.nlm.nih.gov/16724105/)

40. Mallory AC, Reinhart BJ, Bartel D, Vance VB, Bowman LH. A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro- RNAs in tobacco. *Proc Natl Acad Sci U S A*. 2002; 99(23):15228–33. [10.1073/pnas.232434999](https://doi.org/10.1073/pnas.232434999) PMID: [12403829](https://pubmed.ncbi.nlm.nih.gov/12403829/)
41. Chapman EJ, Prokhnevsky AI, Gopinath K, Dolja VV, Carrington JC. Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes Dev*. 2004; 18(10):1179–86. [10.1101/gad.1201204](https://doi.org/10.1101/gad.1201204) PMID: [15131083](https://pubmed.ncbi.nlm.nih.gov/15131083/)
42. Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, et al. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA uncton. *Dev Cell*. 2003; 4(2):205–17. PMID: [12586064](https://pubmed.ncbi.nlm.nih.gov/12586064/)
43. Schott G, Mari-Ordóñez A, Himber C, Alioua A, Voinnet O, Dunoyer P. Differential effects of viral silencing suppressors on siRNA and miRNA loading support the existence of two distinct cellular pools of ARGONAUTE1. *The EMBO journal*. 2012; 31(11):2553–65. Epub 2012/04/26. [10.1038/emboj.2012.92](https://doi.org/10.1038/emboj.2012.92) PMID: [22531783](https://pubmed.ncbi.nlm.nih.gov/22531783/)
44. Shibolet Y, Haronsky E, Leibman D, Arazi T, Wassenegger M, Whitham SA, et al. The conserved FRNK box in HC-Pro, a plant viral suppressor of gene silencing, is required for small RNA binding and mediates symptom development. *J Virol*. 2007; 81(23):13135–48. [10.1128/JVI.01031-07](https://doi.org/10.1128/JVI.01031-07) PMID: [17898058](https://pubmed.ncbi.nlm.nih.gov/17898058/)
45. Endres MW, Gregory BD, Gao Z, Foreman AW, Mlotshwa S, Ge X, et al. Two plant viral suppressors of silencing require the ethylene-inducible host transcription factor RAV2 to block RNA silencing. *PLoS Pathog*. 2010; 6(1):e1000729. [10.1371/journal.ppat.1000729](https://doi.org/10.1371/journal.ppat.1000729) PMID: [20084269](https://pubmed.ncbi.nlm.nih.gov/20084269/)
46. Ala-Poikela M, Goytia E, Haikonen T, Rajamäki ML, Valkonen JP. Helper component proteinase of the genus Potyvirus is an interaction partner of translation initiation factors eIF(iso)4E and eIF4E and contains a 4E binding motif. *J Virol*. 2011; 85(13):6784–94. [10.1128/JVI.00485-11](https://doi.org/10.1128/JVI.00485-11) PMID: [21525344](https://pubmed.ncbi.nlm.nih.gov/21525344/)
47. Anandalakshmi R, Marathe R, Ge X, Herr JM Jr., Mau C, Mallory A, et al. A calmodulin- related protein that suppresses posttranscriptional gene silencing in plants. *Science*. 2000; 290(5489):142–4. Epub 2000/10/06. PMID: [11021800](https://pubmed.ncbi.nlm.nih.gov/11021800/)
48. Iki T, Yoshikawa M, Nishikiori M, Jaudal MC, Matsumoto-Yokoyama E, Mitsuhashi I, et al. In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Mol Cell*. 2010; 39(2):282–91. [10.1016/j.molcel.2010.05.014](https://doi.org/10.1016/j.molcel.2010.05.014) PMID: [20605502](https://pubmed.ncbi.nlm.nih.gov/20605502/)
49. Ballut L, Drucker M, Pugniere M, Cambon F, Blanc S, Roquet F, et al. HC-Pro, a multifunctional protein encoded by a plant RNA virus, targets the 20S proteasome and affects its enzymic activities. *J Gen Virol*. 2005; 86(Pt 9):2595–603. [10.1099/vir.0.81107-0](https://doi.org/10.1099/vir.0.81107-0) PMID: [16099919](https://pubmed.ncbi.nlm.nih.gov/16099919/)
50. Soitamo AJ, Jada B, Lehto K. HC-Pro silencing suppressor significantly alters the gene expression profile in tobacco leaves and flowers. *BMC Plant Biol*. 2011; 11:68. [10.1186/1471-2229-11-68](https://doi.org/10.1186/1471-2229-11-68) PMID: [21507209](https://pubmed.ncbi.nlm.nih.gov/21507209/)
51. Lellis AD, Kasschau KD, Whitham SA, Carrington JC. Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. *Curr Biol*. 2002; 12(12):1046–51. Epub 2002/07/19. PMID: [12123581](https://pubmed.ncbi.nlm.nih.gov/12123581/)
52. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, et al. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans- acting siRNA formation. *Cell*. 2008; 133(1):128–41. Epub 2008/03/18. [10.1016/j.cell.2008.02.033](https://doi.org/10.1016/j.cell.2008.02.033) PMID: [18342362](https://pubmed.ncbi.nlm.nih.gov/18342362/)
53. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell*. 2008; 133(1):116–27. [10.1016/j.cell.2008.02.034](https://doi.org/10.1016/j.cell.2008.02.034) PMID: [18342361](https://pubmed.ncbi.nlm.nih.gov/18342361/)
54. Wang H, Zhang X, Liu J, Kiba T, Woo J, Ojo T, et al. Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J*. 2011; 67(2):292–304. Epub 2011/04/05. [10.1111/j.1365-3113.2011.04594.x](https://doi.org/10.1111/j.1365-3113.2011.04594.x) PMID: [21457371](https://pubmed.ncbi.nlm.nih.gov/21457371/)
55. Merai Z, Kerenyi Z, Kertesz S, Magna M, Lakatos L, Silhavy D. Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. *J Virol*. 2006; 80(12):5747–56. [10.1128/JVI.01963-05](https://doi.org/10.1128/JVI.01963-05) PMID: [16731914](https://pubmed.ncbi.nlm.nih.gov/16731914/)
56. Cao M, Du P, Wang X, Yu YQ, Qiu YH, Li W, et al. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. *Proc Natl Acad Sci U S A*. 2014; 111(40):14613–8. [10.1073/pnas.1407131111](https://doi.org/10.1073/pnas.1407131111) PMID: [25201959](https://pubmed.ncbi.nlm.nih.gov/25201959/)
57. Omarov RT, Ciomperlik JJ, Scholthof HB. RNAi-associated ssRNA-specific ribonucleases in Tombusvirus P19 mutant-infected plants and evidence for a discrete siRNA-containing effector complex. *Proc Natl Acad Sci U S A*. 2007; 104(5):1714–9. [10.1073/pnas.0608117104](https://doi.org/10.1073/pnas.0608117104) PMID: [17244709](https://pubmed.ncbi.nlm.nih.gov/17244709/)
58. Pantaleo V, Szittyá G, Burgyn J. Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. *J Virol*. 2007; 81(8):3797–806. [10.1128/JVI.02383-06](https://doi.org/10.1128/JVI.02383-06) PMID: [17267504](https://pubmed.ncbi.nlm.nih.gov/17267504/)

59. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, et al. A gene expression map of *Arabidopsis thaliana* development. *Nat Genet.* 2005; 37(5):501–6. [10.1038/ng1543](https://doi.org/10.1038/ng1543) PMID: [15806101](https://pubmed.ncbi.nlm.nih.gov/15806101/)
60. Baumberger N, Baulcombe DC. *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci U S A.* 2005; 102(33):11928–33. [10.1073/pnas.0505461102](https://doi.org/10.1073/pnas.0505461102) PMID: [16081530](https://pubmed.ncbi.nlm.nih.gov/16081530/)
61. Donaire L, Barajas D, Martinez-Garcia B, Martinez-Priego L, Pagan I, Llave C. Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. *J Virol.* 2008; 82(11):5167–77. [10.1128/JVI.00272-08](https://doi.org/10.1128/JVI.00272-08) PMID: [18353962](https://pubmed.ncbi.nlm.nih.gov/18353962/)
62. Cao M, Ye X, Willie K, Lin J, Zhang X, Redinbaugh MG, et al. The capsid protein of Turnip crinkle virus overcomes two separate defense barriers to facilitate systemic movement of the virus in *Arabidopsis*. *J Virol.* 2010; 84(15):7793–802. [10.1128/JVI.02643-09](https://doi.org/10.1128/JVI.02643-09) PMID: [20504923](https://pubmed.ncbi.nlm.nih.gov/20504923/)
63. Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park HS, et al. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* 2006; 34(21):6233–46. [10.1093/nar/gkl886](https://doi.org/10.1093/nar/gkl886) PMID: [17090584](https://pubmed.ncbi.nlm.nih.gov/17090584/)
64. Wang N, Zhang D, Wang Z, Xun H, Ma J, Wang H, et al. Mutation of the RDR1 gene caused genome-wide changes in gene expression, regional variation in small RNA clusters and localized alteration in DNA methylation in rice. *BMC Plant Biol.* 2014; 14:177. [10.1186/1471-2229-14-177](https://doi.org/10.1186/1471-2229-14-177) PMID: [24980094](https://pubmed.ncbi.nlm.nih.gov/24980094/)
65. Montgomery TA, Yoo SJ, Fahlgren N, Gilbert SD, Howell MD, Sullivan CM, et al. AGO1-miR173 complex initiates phased siRNA formation in plants. *Proc Natl Acad Sci U S A.* 2008; 105(51):20055–62. Epub 2008/12/11. [0810241105](https://doi.org/10.1073/pnas.0810241105) [pii] [10.1073/pnas.0810241105](https://doi.org/10.1073/pnas.0810241105) PMID: [19066226](https://pubmed.ncbi.nlm.nih.gov/19066226/)
66. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, et al. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in *Arabidopsis*. *Nature structural & molecular biology.* 2010; 17(8):997–1003. Epub 2010/06/22. [10.1038/nsmb.1866](https://doi.org/10.1038/nsmb.1866)
67. Rajeswaran R, Aregger M, Zvereva AS, Borah BK, Gubaeva EG, Pooggin MM. Sequencing of RDR6-dependent double-stranded RNAs reveals novel features of plant siRNA biogenesis. *Nucleic Acids Res.* 2012; 40(13):6241–54. [10.1093/nar/gks242](https://doi.org/10.1093/nar/gks242) PMID: [22434877](https://pubmed.ncbi.nlm.nih.gov/22434877/)
68. Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH. 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc Natl Acad Sci U S A.* 2010; 107(34):15269–74. [10.1073/pnas.1001738107](https://doi.org/10.1073/pnas.1001738107) PMID: [20643946](https://pubmed.ncbi.nlm.nih.gov/20643946/)
69. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell.* 2005; 121(2):207–21. [10.1016/j.cell.2005.04.004](https://doi.org/10.1016/j.cell.2005.04.004) PMID: [15851028](https://pubmed.ncbi.nlm.nih.gov/15851028/)
70. Bhattacharjee S, Zamora A, Azhar MT, Sacco MA, Lambert LH, Moffett P. Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. *The Plant journal: for cell and molecular biology.* 2009; 58(6):940–51. Epub 2009/02/18. [10.1111/j.1365-3113.2009.03832.x](https://doi.org/10.1111/j.1365-3113.2009.03832.x)
71. Curtis MD, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in plants. *Plant Physiol.* 2003; 133(2):462–9. [10.1104/pp.103.027979](https://doi.org/10.1104/pp.103.027979) PMID: [14555774](https://pubmed.ncbi.nlm.nih.gov/14555774/)
72. Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. SERRATE: a new player on the plant micro-RNA scene. *EMBO Rep.* 2006; 7(10):1052–8. [10.1038/sj.embor.7400806](https://doi.org/10.1038/sj.embor.7400806) PMID: [16977334](https://pubmed.ncbi.nlm.nih.gov/16977334/)
73. Agorio A, Vera P. ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant Cell.* 2007; 19(11):3778–90. [10.1105/tpc.107.054494](https://doi.org/10.1105/tpc.107.054494) PMID: [17993621](https://pubmed.ncbi.nlm.nih.gov/17993621/)
74. Hunter C, Sun H, Poethig RS. The *Arabidopsis* heterochronic gene *ZIPPY* is an ARGONAUTE family member. *Curr Biol.* 2003; 13(19):1734–9. PMID: [14521841](https://pubmed.ncbi.nlm.nih.gov/14521841/)
75. Kleinboelting N, Huep G, Kloetgen A, Viehoveer P, Weisshaar B. GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res.* 2012; 40(Database issue):D1211–5. [10.1093/nar/gkr1047](https://doi.org/10.1093/nar/gkr1047) PMID: [22080561](https://pubmed.ncbi.nlm.nih.gov/22080561/)
76. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, et al. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science.* 2003; 301(5633):653–7. [10.1126/science.1086391](https://doi.org/10.1126/science.1086391) PMID: [12893945](https://pubmed.ncbi.nlm.nih.gov/12893945/)
77. Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 1998; 16(6):735–43. PMID: [10069079](https://pubmed.ncbi.nlm.nih.gov/10069079/)
78. Varallyay E, Valoczi A, Agyi A, Burgyan J, Havelda Z. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *EMBO J.* 2010; 29(20):3507–19. [10.1038/emboj.2010.215](https://doi.org/10.1038/emboj.2010.215) PMID: [20823831](https://pubmed.ncbi.nlm.nih.gov/20823831/)
79. Gilbert K, Faber N, Kasschau K, Chapman EJ, Carrington JC, Carbonell A. Preparation of Multiplexed Small RNA Libraries From Plants. *Bioprotocol.* 2014; 4(21).
80. Fahlgren N, Sullivan CM, Kasschau KD, Chapman EJ, Cumbie JS, Montgomery TA, et al. Computational and analytical framework for small RNA profiling by high-throughput sequencing. *RNA.* 2009; 15(5):992–1002. [10.1261/ma.1473809](https://doi.org/10.1261/ma.1473809) PMID: [19307293](https://pubmed.ncbi.nlm.nih.gov/19307293/)

Appendix B

Raspberry Pi powered imaging for plant phenotyping

B.1 Preface

This appendix is a manuscript in “Protocol Note” format that has been tentatively accepted for publication in Applications in Plant Sciences, pending acceptance of minor revisions. The initially submitted version is available at <http://www.biorxiv.org/content/early/2017/09/01/183822>

This paper provides step-by-step instructions for constructing and using three imaging system, along with a protocol for software configuration steps common to all three methods. I developed the first system for the work described in chapter 3. The second and third system were developed by the Gehan lab, and Malia Gehan made both main-text figures. Several of the scripts and configuration files that accompany the paper were adapted from my online documentation, which is described in further detail in Appendix C. As co-first author, I made major contributions to the drafting and revision of the manuscript.

Raspberry Pi Powered Imaging for Plant Phenotyping¹

Jose C. Tovar^{2*}, J. Steen Hoyer^{2,3*}, Andy Lin, Allison Tielking, Steven T. Callen², S. Elizabeth Castillo², Michael Miller, Monica Tessman, Noah Fahlgren², James C. Carrington, Dmitri A. Nusinow², and Malia A. Gehan^{2,4}

² Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, Missouri 63132

³ Computational and systems biology program, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130

* These authors contributed equally

Email addresses: JCT: jtovar@danforthcenter.org

JSH: j.s.hoyer@wustl.edu

AL: azlin231@gmail.com

AT: atielkin@stanford.edu

STC: scallen@danforthcenter.org

SEC: ecastillo@danforthcenter.org

MM: mbmiller808@gmail.com

MT: monica.tessman@gmail.com

NF: nfahlgren@danforthcenter.org

JCC: jcarrington@danforthcenter.org

DAN: dnusinow@danforthcenter.org

MAG: mgehan@danforthcenter.org

¹ Manuscript received _____; revision accepted _____.

⁴ Author for correspondence: mgehan@danforthcenter.org

Acknowledgements: This work was funded by the Danforth Plant Science Center and the National Science Foundation: IOS-1456796, IOS-1202682, EPSCoR IIA-1355406, IIA-1430427, IIA-1430428, MCB-1330562, DBI-1156581. We thank Neff Power (St. Louis, MO) for help in determining the framing to build the multi-image octagon and camera stand. We thank G. Wang (Computomics Corporation), J. Tate, C. Lizárraga, L. Chavez, and N. Shakoor for helpful discussions and the editors of this series for the opportunity to contribute this paper.

ABSTRACT

- *Premise of the study:* Image-based phenomics is a powerful approach to capture and quantify plant diversity. However, commercial platforms that make consistent image acquisition easy are often cost-prohibitive. To make high-throughput phenotyping methods more accessible, low-cost microcomputers and cameras can be used to acquire plant image data.
- *Methods and Results:* We used low-cost Raspberry Pi computers and cameras to manage and capture plant image data. Detailed here are three different applications of Raspberry Pi controlled imaging platforms for seed and shoot imaging. Images obtained from each platform were suitable for extracting quantifiable plant traits (shape, area, height, color) *en masse* using open-source image processing software such as PlantCV.
- *Conclusion:* This protocol describes three low-cost platforms for image acquisition that are useful for quantifying plant diversity. When coupled with open-source image processing tools, these imaging platforms provide viable low-cost solutions for incorporating high-throughput phenomics into a wide range of research programs.

Key words: imaging; low-cost phenotyping; morphology; Raspberry Pi

INTRODUCTION

Image-based high-throughput phenotyping has been heralded as a solution for measuring diverse traits across the tree of plant life (Araus and Cairns, 2014; Goggin et al., 2015). In general, there are five steps in image-based plant phenotyping: 1) image and metadata acquisition; 2) data transfer; 3) image segmentation (separation of target object and background); 4) trait extraction (object description); and 5) group-level data analysis. Image segmentation, trait extraction, and data analysis are the most time-consuming steps of the phenotyping process, but protocols that increase the speed and consistency of image and metadata acquisition greatly speed up downstream analysis steps. Commercial high-throughput phenotyping platforms are powerful tools to collect consistent image data and metadata and are even more effective when designed for targeted biological questions (Topp et al., 2013; Chen et al., 2014; Honsdorf et al., 2014; Yang et al., 2014; Al-Tamimi et al., 2016; Pauli et al., 2016; Feldman et al., 2017; Zhang et al., 2017). However, commercial phenotyping platforms are cost-prohibitive to many laboratories and institutions. There is also no such thing as a ‘one-size fits all’ phenotyping system; different biological questions often require different hardware configurations. Therefore, low-cost technologies that can be used and repurposed for a variety of phenotyping applications are of great value to the plant community.

Raspberry Pi computers are small (credit card sized or smaller), low-cost, and were originally designed for educational purposes (Upton and Halfacree, 2014). Several generations of Raspberry Pi single-board computers have been released, and most models now feature built-in

modules for wireless and bluetooth connectivity (Monk, 2016). The Raspberry Pi Foundation also releases open source software and accessories such as camera modules (5 and 8 megapixel). Additional sensors or controllers can be connected via USB ports and general-purpose input/output pins. A strong online community of educators and hobbyists provide support (including project ideas and documentation), and a growing population of researchers use Raspberry Pi computers for a wide range of applications including phenotyping. We and others (e.g. Huang et al., 2016; Mutka et al., 2016; Minervini et al., 2017) have utilized Raspberry Pi computers in a number of configurations to streamline collection of image data and metadata. Here, we document three different methods for using Raspberry Pi computers for plant phenotyping (Figure 1). These protocols are a valuable resource because while there are many phenotyping papers that outline phenotyping systems in detail (Granier et al., 2006; Iyer-Pascuzzi et al., 2010; Jahnke et al., 2016; Shafiekhani et al., 2017), there are few protocols that provide step-by-step instructions for building them (Bodner et al., 2017; Minervini et al., 2017). We provide examples illustrating automation of photo capture with open source tools (based on the Python programming language and standard Linux utilities). Further, to demonstrate that these data are of high quality and suitable for quantitative trait extraction, we segmented example image data (plant isolated from background) using the open-source open-development phenotyping software PlantCV (Fahlgren et al., 2015).

METHODS AND RESULTS

Raspberry Pi Initialization: This work describes three protocols (Appendices 2-4) that utilize Raspberry Pi computers for low-cost image-based phenotyping and gives examples of the data they produce. Raspberry Pi computers can be reconfigured for different phenotyping projects and can be easily purchased from online retailers. The first application is time-lapse plant imaging (Appendix 2); the second protocol describes setup and use of an adjustable camera stand for top-view photography (Appendix 3); and the third project describes construction and use of an octagonal box for acquiring plant images from several angles simultaneously (Appendix 4). For all three phenotyping protocols, the same protocol to initialize Raspberry Pi computers is used and is provided in Appendix 1. The initialization protocol in Appendix 1 parallels the Raspberry Pi Foundation's online documentation and provides additional information on setting up passwordless secure shell (SSH) login to a remote host for data transfer and/or to control multiple Raspberry Pis. Passwordless SSH allows one to pull data from the data collection computer to a remote server without having to manually enter login information each time. Reliable data transfer is an important consideration in plant phenotyping projects because, while it is possible to process image data directly on a Raspberry Pi computer, most users will prefer to process large image datasets on a bioinformatics cluster. Remote data transfer is especially important for time-lapse imaging setups, such as the configuration described in Appendix 2, because data can be generated at high frequency over the course of long experiments, and thus can easily exceed available disk space on the micro secure digital (SD) cards that serve as local hard-drives. Once one Raspberry Pi has been properly configured and tested, the fully configured

operating system can be backed up, yielding a disk image that can be copied (“cloned”) onto as many additional SD cards as are needed for a given phenotyping project (Appendix 1).

Raspberry Pi Time-lapse Imaging: Time-lapse imaging is a valuable tool for documenting plant development and can reveal differences that would not be apparent from endpoint analysis.

Raspberry Pi computers and camera modules work effectively as phenotyping systems in controlled-environment growth chambers; and low cost of Raspberry Pi computers allows this approach to scale well. Growth chambers differ from (agro)ecological settings but are an essential tool for precise control and reproducible experimentation (Poorter et al., 2016).

Time-lapse imaging with multiple cameras allows for simultaneous imaging of many plants and can capture higher temporal resolution than conveyor belt and mobile-camera systems. Appendix 2 provides an example protocol for setting up the hardware and software necessary to capture plant images in a growth chamber. The main top-view imaging setup described is aimed at imaging flats or pots of plants in a growth chamber. We include instructions for adjusting the camera-plant focal distance (yielding higher plant spatial resolution) and describe how to adjust the temporal resolution of imaging. The focal distance can be optimized to the target plant, trait, and degree of precision required; large plant-camera distances allow a larger field of view, at the cost of lower resolution. For traits like plant area, where segmentation of individual plant organs is not critical, adjusting the focal length might not be necessary. Projected leaf area in top-down photos correlates well with fresh and dry weight, especially for relatively flat plants such as *Arabidopsis thaliana* (Leister et al., 1999). A stable and level imaging configuration is important for consistent imaging across long experiments and to compare data from multiple Raspberry

Pi/Camera rigs. Although there is more than one way to suspend Raspberry Pi/Camera rigs in a flat and stable top-view configuration, AC power socket adapters were attached to the back of cases with silicone adhesive (Appendix 2). Raspberry Pi boards and cameras were then encased and screwed into the incandescent bulb sockets built into the growth chamber (Figure 1). Users with access to a 3D printer may prefer to print cases, so we have provided a link to instructions for printing a suitable case (with adjustable ball-joint Raspberry Pi camera module mount) in Appendix 2. This type of 3D printed case also works well for side-view imaging of plants grown on plates (Huang et al., 2016; Mutka et al., 2016). For this top-down imaging example, twelve Raspberry Pi/Camera rigs were powered through two USB power supplies drawing power (via extension cord and surge protector) from an auxiliary power outlet built into the growth chamber. Although we use twelve Raspberry Pi/Camera rigs in this example, the setup can be scaled up or down, with a per-unit cost of approximately \$100. A single Pi/Camera rig is enough for a new user to get started, and laboratories can efficiently scale up imaging as they develop experience and refine their goals. Time-lapse imaging was scheduled at five-minute intervals using the software utility cron. A predictable file naming scheme that includes image metadata (field of view number, timestamp, and a common identifier) was employed to confirm that all photo timepoints were captured and transferred as scheduled. Images were pulled from each Raspberry Pi to a remote server twice per hour (using a standard utility called rsync) by a server-side cron process using the configuration files described in Appendix 2.

Optimizing imaging conditions for maximum consistency can simplify downstream image processing. To aid in image normalization during processing, color standards and size markers

can be included in images. Placing rubberized blue mesh (e.g. Con-Tact Brand, Pomona, California, USA) around the base of plants can sometimes simplify segmentation (i.e. distinguishing plant foreground pixels from soil background pixels), though this was not necessary for the *A. thaliana* example described here. Care should be taken to ensure that large changes in the scene (including gradual occlusion of blue mesh by leaves) do not dramatically alter automatic exposure and color balance settings over the course of an experiment. If automatic exposure becomes an issue, camera settings can be manually set (see Appendix 4). In this example, cameras and flats were set up to yield a similar vantage point (a 4 x 5 grid of pots) in each field of view, such that very similar computational pipelines can be used to process images from all twelve cameras. An example image has been processed with PlantCV (Fahlgren et al., 2015) in Figure 2, and a script showing and describing each step in the analysis is provided at <https://github.com/danforthcenter/apps-phenotyping>. Further image processing tutorials and tips can be found at <http://plantcv.readthedocs.io/en/latest/>.

Raspberry Pi Camera Stand: An adjustable camera stand is a versatile piece of laboratory equipment for consistent imaging. Appendix 3 is a protocol for pairing a low-cost home-built camera stand with a Raspberry Pi computer for data capture and management. Altogether, the camera stand system costs approximately \$750. The camera stand (79 cm width x 82.5 cm height) was built from aluminum framing (80/20, Columbia City, Indiana, USA) to hold a Nikon Coolpix L830 camera via a standard mount (Figure 1). For this application, we prefer to use a single-lens reflex (SLR) digital camera (rather than a Raspberry Pi camera module) for adjustable focus and to improve resolution. The camera was affixed to a movable bar, so the

distance between camera and object can be adjusted up to 63 cm. A Python script that utilizes gphoto2 (Figuière and Niedermann, 2017) for data capture and rsync for data transfer to a remote host is included in the protocol (Appendix 3). When the ‘camerastand.py’ script is run, the user is prompted to enter the filename for the image. The script verifies that the camera is connected to the Raspberry Pi, acquires the image with the SLR camera, retrieves the image from the camera, renames the image file to the user-provided filename, saves a copy in a local Raspberry Pi directory, and transfers this copy to the desired directory on a remote host. As image filenames are commonly used as the primary identifier for downstream image processing, it is advised to use a filename that identifies the species, accession, treatment, and replicate, as appropriate. The Python script provided appends a timestamp to the filename automatically. We regularly use this Raspberry Pi Camera Stand to image seeds, plant organs (e.g. inflorescences), and short-statured plants. For seed images, a white background with a demarcated black rectangular area ensures that separated seeds are in frame, which speeds up the imaging process. Color cards (white, black, and gray; DGK Color Tools, New York, New York, USA) and a size marker to normalize area are also included in images to aid in downstream processing and analysis steps. It is advised to use the same background, and, if possible, the same distance between object and camera for all images in an experimental set. However, including a size marker in images can be used to normalize data extracted from images if the vantage point does change. *Chenopodium quinoa* (quinoa) seed images are shown as example data from the camera stand (Figure 2). To show that images collected from the camera stand are suitable for image analysis, seed images acquired with the camera stand were processed using PlantCV (Fahlgren et al., 2015) to quantify individual seed size, shape, color, and count; these types of measurements

are valuable for quantifying variation within a population. The step-by-step image processing instructions are provided at <https://github.com/danforthcenter/apps-phenotyping>. This overall process (Appendix 3) provides a considerable cost savings relative to paying for seed imaging services or buying a commercial seed imaging station.

Raspberry Pi Multi-Image Octagon: Different plant architecture types require different imaging configurations for capture. For example, top-down photographs can capture most of the information about the architecture of rosette plants (as described above), but plants with orthotropic growth such as rice or quinoa are better captured with a combination of both side-view and top-view images. Therefore, platforms for simultaneously imaging plants from multiple angles are valuable. In Appendix 4, a protocol is described to set up an octagon-shaped chamber for imaging at different angles. The complete octagon-shaped imaging system costs approximately \$1500. A ‘master’ Raspberry Pi computer with a Raspberry Pi camera module is used to collect image data and also to trigger three other Raspberry Pi computers and cameras. Data is transferred from the four Raspberry Pi computers to a remote host using rsync. The octagon chamber (122 cm height and 53.5 cm of each octagonal side) was constructed from aluminum framing and 3mm white polvinyl chloride (PVC) panels (80/20, Columbia City, Indiana, USA; Figure 1). The top of this structure is left open but is covered with a translucent white plastic tarp to diffuse light when acquiring images. A latched door was built into the octagon chamber to facilitate loading of plants. Four wheels were attached at the bottom of the chamber for mobility. The four Raspberry Pis with Raspberry Pi camera modules (one top-view and three side-views approximately 45° angle apart) in cases were affixed to the octagon

chamber using heavy-duty velcro. To maintain a consistent distance between the Raspberry Pi cameras and a plant within the Raspberry Pi multi-image octagon, a pot was affixed to the center of the octagon chamber, with color cards affixed to the outside of the stationary pot (white, black, and gray; DGK Color Tools, New York, New York, USA) so that a potted plant could be quickly placed in the pot during imaging.

To facilitate data acquisition and transfer on all four Raspberry Pis, scripts are written so the user only needs to interact with a single ‘master’ Raspberry Pi (here the master Raspberry Pi is named ‘octagon’). From a laptop computer one would connect to the ‘master’ pi via SSH, then run the ‘sshScript.sh’ on that Pi. The ‘sshScript.sh’ script triggers the image capture and data transfer sequence in all four Raspberry Pis and appends the date to a user-input barcode. When the ‘sshScript.sh’ script is run, a prompt asks the user for a barcode sequence. The barcode can be inputted manually, or, if a barcode scanner (e.g. Socket 7Qi) is available, a barcode can be used to input the filename information. Again, it is advised to use a plant barcode that identifies the species, accession, treatment, and replicate, as appropriate. Once a barcode name has been inputted, another prompt asks if the user would like to continue with image capture. This pause in the ‘sshScript.sh’ script gives the user the opportunity to place the plant in the octagon before image capture is triggered. The sshScript.sh runs the script piPicture.py on all four Raspberry Pis. The ‘piPicture.py’ script captures an image and appends the user inputted filename with the Raspberry Pi camera id and the date. The image is then saved to a local directory on the Raspberry Pi. The ‘syncPi.sh’ script is then run by ‘sshScript.sh’ to transfer the images from the four Raspberry Pis to a remote host. The final script (shutdown_all_pi) is optionally run when

image acquisition is over, allowing the user to shut down all four Raspberry Pis simultaneously. Examples of quinoa plant images captured with the Raspberry Pi multi-image octagon are analyzed with PlantCV (Fahlgren et al., 2015) to show that the data can be area and shape can be extracted (Figure 2). Step-by-step analysis scripts are provided at <https://github.com/danforthcenter/apps-phenotyping>.

Protocol Feasibility: The protocols provided in the appendices that follow provide step-by-step instructions for using Raspberry Pi computers for plant phenotyping in three different configurations. The majority of components for all three protocols are readily available for purchase online. Low-cost computers and components are especially important since some experiments might test harsh environmental conditions and need to be replaced long-term. Each of the platforms were built and programmed in large part by high-school students, undergraduates, or graduate students and do not require a large investment of time to build or set-up. Since Raspberry Pi computers are widely used by educators, hobbyists, and researchers there is a strong online community that can be called upon for troubleshooting or to extend the functionality of a project. The best way to start troubleshooting is to use an online search engine to see if others have solved similar issues. If an error message has been triggered, start by using the error message as search terms. If a satisfactory answer is not found through an online search, posting on a community support forum like Stack Overflow is a good next step (<https://raspberrypi.stackexchange.com/>; <https://stackoverflow.com/questions/tagged/raspberry-pi>). When posting on online community forums is helpful to be specific. For example, if an error message is triggered it is vital to include

the exact text of the error message, to describe the events that triggered that error message, and what the target end goal is. Automation increases the consistency of image and metadata capture, which streamlines image segmentation (Figure 2) and is thus preferable to manual image capture. Furthermore, the low cost of each system and the flexibility to reconfigure Raspberry Pi computers for multiple purposes makes automated plant phenotyping accessible to most researchers.

CONCLUSION

The low-cost imaging platforms presented here provide an opportunity for labs to introduce phenotyping equipment into their research toolkit, and thus increase the efficiency, reproducibility, and thoroughness of their measurements. These protocols make high-throughput phenotyping accessible to researchers unable to make a large investment in commercial phenotyping equipment. Paired with open-source open-development high-throughput plant phenotyping software like PlantCV (Fahlgren et al., 2015), image data collected from these phenotyping systems can be used to quantify plant traits for populations of plants that are amenable to genetic mapping. These Raspberry Pi powered tools are also useful for education and training. In particular, we have used time-lapse imaging to introduce students and teachers to the Linux environment, image processing, and data analysis in a classroom setting (<http://github.com/danforthcenter/outreach/>). As costs continue to drop and hardware continues to improve, there is enormous potential for the plant science community to capitalize on creative applications, well-documented designs, and shared datasets and code.

LITERATURE CITED

Al-Tamimi, N., C. Brien, H. Oakey, B. Berger, S. Saade, Y.S. Ho, S.M. Schmöckel, et al. 2016.

Salinity tolerance loci revealed in rice using high-throughput non-invasive phenotyping.

Nature Communications 7: 13342.

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science* 19: 52–61.

Bodner, G., M. Alsalem, A. Nakhforoosh, T. Arnold, and D. Leitner. 2017. RGB and spectral root imaging for plant phenotyping and physiological research: experimental setup and imaging protocols. *Journal of Visualized Experiments: JoVE* 126: e56251.

Chen, D., K. Neumann, S. Friedel, B. Kilian, M. Chen, T. Altmann, and C. Klukas. 2014.

Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *The Plant Cell* 26: 4636–4655.

Fahlgren, N., M. Feldman, M.A. Gehan, M.S. Wilson, C. Shyu, D.W. Bryant, S.T. Hill, et al.

2015. A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in *Setaria*. *Molecular Plant* 8: 1520–1535.

Feldman, M.J., R.E. Paul, D. Banan, J.F. Barrett, J. Sebastian, M.-C. Yee, H. Jiang, et al. 2017.

Time dependent genetic analysis links field and controlled environment phenotypes in the model C4 grass *Setaria*. *PLOS Genetics* 13: e1006841.

- Figuière, H., and H.U. Niedermann. 2017. Version 2.5.14 (2017-06-05). gPhoto - opensource digital camera access and remote control. GitHub, San Francisco, California, USA. Website <https://github.com/gphoto> [accessed 22 August 2017].
- Goggin, F.L., A. Lorence, and C.N. Topp. 2015. Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops. *Current Opinion in Insect Science* 9: 69–76.
- Granier, C., L. Aguirrezabal, K. Chenu, S.J. Cookson, M. Dauzat, P. Hamard, J.-J. Thioux, et al. 2006. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *The New Phytologist* 169: 623–635.
- Honsdorf, N., T.J. March, B. Berger, M. Tester, and K. Pillen. 2014. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLOS ONE* 9: e97047.
- Huang, H., C.Y. Yoo, R. Bindbeutel, J. Goldsworthy, A. Tielking, S. Alvarez, M.J. Naldrett, et al. 2016. PCH1 integrates circadian and light-signaling pathways to control photoperiod-responsive growth in *Arabidopsis*. *eLife* 5: e13292.
- Iyer-Pascuzzi, A.S., O. Symonova, Y. Mileyko, Y. Hao, H. Belcher, J. Harer, J.S. Weitz, and P.N. Benfey. 2010. Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant Physiology* 152: 1148–1157.
- Jahnke, S., J. Roussel, T. Hombach, J. Kochs, A. Fischbach, G. Huber, and H. Scharr. 2016.

phenoSeeder - A robot system for automated handling and phenotyping of individual seeds.

Plant Physiology 172: 1358–1370.

Jones, D. 2017. Version 1.13 (2017-02-25). Picamera. GitHub, San Francisco, California, USA.

Website <https://github.com/waveform80/picamera> [accessed 01 September 2017].

Leister, D., C. Varotto, P. Pesaresi, A. Niwergall, and F. Salamini. 1999. Large-scale evaluation of plant growth in *Arabidopsis thaliana* by non-invasive image analysis. *Plant Physiology and Biochemistry* 37: 671–678.

Minervini, M., M.V. Giuffrida, P. Perata, and S.A. Tsaftaris. 2017. Phenotiki: An open software and hardware platform for affordable and easy image-based phenotyping of rosette-shaped plants. *The Plant Journal: for Cell and Molecular Biology* 90: 204-216.

Monk, S. 2016. Raspberry Pi cookbook: software and hardware problems and solutions. O'Reilly Media, Sebastopol, California, USA.

Mutka, A.M., S.J. Fentress, J.W. Sher, J.C. Berry, C. Pretz, D.A. Nusinow, and R. Bart. 2016. Quantitative, image-based phenotyping methods provide insight into spatial and temporal dimensions of plant disease. *Plant Physiology* 172: 650–660.

Pauli, D., P. Andrade-Sanchez, A.E. Carmo-Silva, E. Gazave, A.N. French, J. Heun, D.J.

Hunsaker, et al. 2016. Field-based high-throughput plant phenotyping reveals the temporal patterns of quantitative trait loci associated with stress-responsive traits in cotton. *G3* 6: 865–879.

- Poorter, H., F. Fiorani, R. Pieruschka, T. Wojciechowski, W.H. van der Putten, M. Kleyer, U. Schurr, and J. Postma. 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *The New Phytologist* 212: 838–855.
- Shafiekhani, A., S. Kadam, F.B. Fritschi, and G.N. DeSouza. 2017. Vinobot and Vinoculer: two robotic platforms for high-throughput field phenotyping. *Sensors* 17: 214.
- Upton, E., and G. Halfacree. 2014. Raspberry Pi user guide. John Wiley & Sons, Hoboken, New Jersey, USA.
- Yang, W., Z. Guo, C. Huang, L. Duan, G. Chen, N. Jiang, W. Fang, et al. 2014. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications* 5: 5087.
- Zhang, X., C. Huang, D. Wu, F. Qiao, W. Li, L. Duan, K. Wang, et al. 2017. High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiology* 173: 1554–1564.

APPENDICES

Appendix 1. Initializing a Raspberry Pi for phenotyping projects. The camera stand, growth-chamber imaging stations, and multi-image octagon phenotyping platforms that are described in detail in Appendices 2-4 use Raspberry Pis to trigger image acquisition, append metadata to filenames, and move data to remote host machines. The following are the required parts and steps to initialize a single Raspberry Pi. The initialization protocol is based on the installation guidelines from the Raspberry Pi Foundation, which are under a Creative Commons license (<https://www.raspberrypi.org/documentation/>).

Parts List:

Item
Raspberry Pi single-board microcomputer
Micro USB power supply
Mini Secure Digital (SD) card, we recommend 16GB
HDMI Monitor, HDMI cable, keyboard, and mouse

General Raspberry Pi Initialization Protocol:

1. Install ‘Raspbian Stretch with Desktop’ (here version 4.9 is used, but the latest version is recommended) onto the SD card by following the installation guide at <https://www.raspberrypi.org/downloads/raspbian/>

2. Insert mini SD card into Raspberry Pi and plug in monitor, keyboard, and mouse to Raspberry Pi.
3. Plug in Micro USB power supply and connect to power. The Raspberry Pi will boot to the desktop interface, which is also known as the graphical user interface (GUI). If the Raspberry Pi does not boot to the desktop interface, you can type `sudo raspi-config` and go to the third option ‘Enable Boot to Desktop/Scratch’ to change this. Alternately, if the Raspberry Pi boots to the command line you can get to the GUI by typing ‘startx’ and hitting the Enter key.
4. Once at the desktop, open Raspberry Pi Configuration under Applications Menu > Preferences. Alternatively, you can get to the configurations menu by typing “`sudo raspi-config`” in the Terminal program.
 - a. In the System tab, set hostname (see Appendices 2, 3, or 4 for specific hostnames to use; alternatively, a static IP address can be set up for the Raspberry PI).
 - b. In the Interfaces tab, set SSH and Camera to enabled.
 - c. In the Localization tab:
 - i. Set Locale to appropriate Language and Country, and leave Character Set as UTF-8 (default option).
 - ii. Set Timezone to an appropriate Area and Location. Universal Coordinated Time (UTC) can be advantageous for long-running time-lapse experiments.
 - iii. Set Keyboard to appropriate Country and Variant.

iv. Set WiFi Country.

5. Configure WiFi using the network icon on the top right of the desktop.

Alternatively, use an Ethernet cable connection.

6. Optionally, make a local copy of the scripts that accompany this paper. In

Terminal, change directory to the Desktop by typing “cd Desktop”. Then type “git clone <https://github.com/danforthcenter/apps-phenotyping.git>”. If prompted with “The authenticity of host 'remote-host' can't be established (...) Are you sure you want to continue connecting?” enter “yes”. This will download the project scripts and examples for all three phenotyping platforms (Appendices 2-4). Some of these scripts may need be adjusted after they have been copied on the Raspberry Pi, as described below. The Git version control system can be used to track the history of changes of these files.

Raspberry Pi SD card cloning protocol:

Once you have gone through the initialization protocol for one Raspberry Pi, the disk image of the SD card from that Raspberry Pi can be cloned if you need additional Raspberry Pis for your project. Any project specific scripts that need further adjustments on individual Raspberry Pis can then be completed (see Appendices 2 to 4). Cloning an SD card will generate a file of the exact size of the SD card (e.g. 16 GB), and it is therefore essential ensure that new SD cards to be “flashed” with the original disk image are at least as large as the initialized SD card (e.g. 16 GB or larger).

To clone an SD card on a Windows computer:

1. Download and install Win32 Disk Imager from
<https://sourceforge.net/projects/win32diskimager/>
2. Before opening the Win32 Disk Imager, insert the SD card (in an SD card reader if needed) from the initialized Raspberry Pi into your computer.
3. Open Win32 Disk Imager.
4. Click on the blue folder icon. A file explorer window will appear.
5. Select the directory to store the SD card image, and provide a filename for the image.
6. Click Open to confirm your selection. The file explorer window will close.
7. Under Device, select the appropriate drive letter for the SD card.
8. Click the Read button.
9. Once the image is created, a 'Read Successful' message will appear. Click OK.
10. Eject the SD card, and close Win32 Disk Imager.
11. Insert the new SD card where the image will be cloned. Make sure this SD card has as much or more storage capacity as the SD card from the initialized Raspberry Pi that was imaged.
12. Reopen Win32 Disk Imager.
13. Click on the blue folder icon, and select the image that was just created.
14. Under Device, select the appropriate drive letter for the SD card where the image will be cloned.
15. Click the Write button.

16. Click Yes.
17. Once the image is created, a 'Write Successful' message will appear. Click OK.
18. Eject the SD card, and insert it into the Raspberry Pi. The Raspberry Pi is now initialized.

To clone an SD card on a Mac computer:

1. Download and install ApplePi-Baker from
<https://www.tweaking4all.com/software/macosx-software/macosx-apple-pi-baker/>
2. Insert the SD card (in an SD card reader if needed) from the initialized Raspberry Pi into your computer.
3. Under Pi-Crust: Select SD-Card or USB drive, select the initialized Raspberry Pi SD card.
4. Click on Create Backup.
5. Click OK.
6. Under Save As, provide a filename for the SD card image.
7. Under Where, select directory to store the SD card image, and click Save.
8. Once the image is created, a 'Your ApplePi is Frozen!' message will appear.
Click OK.
9. Eject the SD card.
10. Insert the new SD card where the image will be cloned. Make sure this SD card has as much or more storage capacity as the SD card from the initialized Raspberry Pi that was imaged.

11. Under Pi-Crust: Select SD-Card or USB drive, select the SD card where the image will be cloned.
12. Click on Restore Backup.
13. Browse and select the image that was just created.
14. Click OK.
15. Once the SD card is cloned, a 'Your ApplePi is ready!' message will appear.
Click OK.
16. Eject the SD card, and insert it into the Raspberry Pi. The Raspberry Pi is now initialized.

To clone an SD card on a Linux computer:

This protocol is adapted from The PiHut

(<https://thepihut.com/blogs/raspberry-pi-tutorials/17789160-backing-up-and-restoring-your-raspberry-pis-sd-card>) and Raspberry Pi Stack Exchange

(<https://raspberrypi.stackexchange.com/questions/311/how-do-i-backup-my-raspberry-pi>)

.

1. First, use the command 'df -h' to see a list of existing devices.
2. Insert the SD card (in an SD card reader if needed) from the initialized Raspberry Pi into your computer.
3. Use the command 'df -h' again. The SD card will be the new item on the list (e.g. /dev/sdbp1 or /dev/sdb1). The last part of the name (e.g. p1 or 1) is the partition number.

4. Use the command ‘`sudo dd if=/dev/SDCardName of=/path/to/SDCardImage.img`’ to create the SD card image (e.g. `sudo dd if=/dev/sdb of=~/InitializedPi.img`). Make sure to remove the partition name to image the entire SD card (e.g. use `/dev/sdb` instead of `/dev/sdb1`).
5. There is no progress indicator, so wait until the command prompt reappears.
6. Unmount the SD card by typing: `sudo umount /dev/SDCardName`.
7. Remove the SD card.
8. Insert the new SD card where the image will be cloned. Make sure this SD card has as much or more storage capacity as the SD card from the initialized Raspberry Pi that was imaged.
9. Use the command ‘`df -h`’ again to discover the new SD card name, or names if there is more than 1 partition.
10. Unmount every partition using the command ‘`sudo umount /dev/SDCardName`’ (e.g. `sudo umount /dev/sdb1`).
11. Copy the initialized Raspberry Pi SD card image using the command ‘`sudo dd if=/path/to/SDCardImage.img of=/dev/SDCardName`’.
12. There is no progress indicator, so wait until the command prompt reappears.
13. Unmount the SD card, and insert it into the Raspberry Pi. The Raspberry Pi is now initialized.

General instructions for installing and testing a Raspberry Pi Camera:

1. Make sure the Raspberry Pi is not connected to power.

2. Pull up the top part of the connector located between the HDMI and ethernet ports, until loose.
3. Insert the Raspberry Pi camera flex cable into the connector, with the silver rectangular plates at the end of the cable facing the HDMI port.
4. While holding the cable in place, push down the top part of the connector to prevent the flex cable from moving.
5. Remove the small piece of blue plastic covering the camera lens, if present.
6. Turn the Raspberry Pi on, and test the camera by opening a Terminal window, then entering “`raspistill -o image-name-here.jpg`” to take a picture.

General instructions for using a SSH keys for passwordless connection to a remote

host: These instructions are to allow a Raspberry Pi computer to access a remote host (e.g. another Raspberry Pi computer, bioinformatics cluster, or other computer), without having to enter login information. If a user would like to pull data from Raspberry Pi to a remote host, rather than pushing data from a Raspberry Pi to a remote host, similar instructions would be followed on the remote computer.

1. In the Terminal window, enter “`ssh-keygen`” to create a public SSH key for passwordless access to a remote host.
2. Press the Enter key to use the default location when asked to “Enter file in which to save the key”.
3. Press Enter two more times, to use the default passphrase setting (no passphrase), if desired.

4. Optionally, enter “ls ~/.ssh” to verify the SSH key was generated. The files “id_rsa” and “id_rsa.pub” should be listed.
5. Use the command “ssh-copy-id -i ~/.ssh/id_rsa.pub user@remote-host” in the Terminal window to copy the public SSH key to the remote host, where “user@remote-host” should be replaced by the name of the remote host where the images will be stored (e.g. ssh-copy-id -i ~/.ssh/id_rsa.pub jdoe@serverx).
6. If prompted with “The authenticity of host 'remote-host' can't be established (...) Are you sure you want to continue connecting?” enter “yes”.
7. Enter the user’s password for the remote server, if prompted.
8. Verify the SSH key was successfully copied to the remote host. SSH to the remote host from the Raspberry Pi, using the command “ssh user@remote-host” in the Terminal window (e.g. ssh jdoe@serverx). No password should be required if the key was copied successfully.

Additional Notes:

- For official distributions of the Raspbian Raspberry Pi operating system, the default username is “pi”, and default password is “raspberry”. The following protocols assume these default settings are unchanged. To change passwords, open Raspberry Pi Configuration under Applications Menu > Preferences. In the System tab, click on ‘Change Password...’. Enter current password, new password, and confirm new password. If using the command line type passwd, then follow the command prompts to change password.

- If an Ethernet cable connection is used, a Power over Ethernet adapter (such as UCTronics LS-POE-B0525) and a Power over Ethernet-capable Ethernet switch (such as Ubiquiti ES-48-750W) can be used, eliminating the need for Raspberry Pi power cables. If connecting multiple computers to a single power supply, ensure that all computers can draw adequate power. For details see <https://www.raspberrypi.org/documentation/hardware/raspberrypi/power/README.md>
- There can be a learning curve associated with command-line tools and Linux-based operating systems. As noted above, material on Linux and Raspberry Pi configuration is available online, including many technical mailing lists and user forums. Learning to use the Linux “man” utility to read manual page documentation can be helpful for quickly looking up command-line flag options and occasionally for understanding differences between the precise versions of software available on different host machines.

Appendix 2. Raspberry Pi Top-View or Side-View Time-Lapse Imaging.

The protocol that follows describes how to set up one or more Raspberry Pi/Camera rigs for time-lapse photography, and is based on the tutorial at

<https://www.raspberrypi.org/documentation/usage/camera/raspicam/timelapse.md>

Here, we focus on a 12-camera configuration that has worked well in a reach-in growth chamber.

We describe one low-cost method for stably fixing Raspberry Pi/Camera rigs to the top of the chamber. Zip ties may work well for attaching Pi/Camera rigs in some growth chambers

(Minervini et al., 2017), and we have also used heavy-duty velcro, so that Pi/Camera rigs can be removed and used for other purposes, such as side-view imaging of seedlings on petri plates. We

have separately provided a protocol for imaging plants grown on vertical petri plates with a

NoIR camera module and an IR light-emitting diode (LED) panel for backlighting (as in Huang et al. 2016). This method was used for quantifying hypocotyl growth, and can be used with other tissues such as roots. See

<http://maker.danforthcenter.org/tutorial/raspberry%20pi/led/raspberry%20pi%20camera/RPi-LE>

D-Illumination-and-Imaging

Parts list:

Item
Initialized Raspberry Pi (Appendix 1)
Raspberry Pi case, such as SmartPi LEGO-compatible case for Raspberry Pi model B and camera module (Smartcase LLC, Philadelphia PA, USA)

Raspberry Pi camera (either RGB or NoIR camera depending on application; see additional notes below)
AC light bulb socket adapters and silicone adhesive sealant, e.g. IS-808
HDMI Monitor, HDMI cable, keyboard, and mouse (for optional lens focus process)

Optional/Alternative Parts:

Item
Heavy-duty velcro, as an alternative method of mounting Pi/Camera rigs. We recommend against using small pieces of low-cost consumer-grade velcro.
3D printed petri plate stand, for imaging plates vertically. STL file for 3D printing is available here: https://www.thingiverse.com/thing:418614/#files

Adjustment of Raspberry Pi Camera Focus (Optional):

As noted above, the Raspberry Pi camera is fixed focus, with a focal length such that objects 1 m away or further will be in focus. Therefore, if you want to alter the focus you have to alter the Raspberry Pi camera module itself with pliers. We recommend watching the excellent youtube tutorial by George Wang:

<https://www.youtube.com/watch?v=u6VhRVH3Z6Y>. For the main configuration described here, the lens-plant distance was 55.2 cm.

1. Install the Raspberry Pi camera on a Raspberry Pi computer that is plugged into a Monitor, Mouse and Keyboard as described above. Plug the power in last.
2. To change the focus, use a sharp object and carefully remove the glue around the lens. Every few incisions, carefully grip the side of the lens-ring with pliers and check if it will turn counter-clockwise. (Removing the glue is optional, as noted in the video linked above.)
3. When the camera lens-ring does turn, use the following command in the Terminal of the Raspberry Pi to check whether a target is adequately focused at your desired imaging distance: “`raspistill -o image.jpg`” We suggest using a ruler and white business card as targets. Continue adjusting the camera at the desired focal distance.

Mounting of Raspberry Pi/Camera rigs for top-view time-lapse imaging:

1. Connect a camera module to each Raspberry Pi, as described in Appendix 1.
2. Attach the AC light bulb socket adapters to the Raspberry Pi cases with silicone adhesive, and install an initialized Raspberry Pi computer (see Appendix 1) and camera in each case.
3. Plug each of the Raspberry Pis into the monitor, keyboard, mouse, and USB power. Following Appendix 1, change the hostnames of the Raspberry Pis. For example, the sample configuration files we provide assume that the twelve

Raspberry Pi have hostnames ‘timepi01’, ‘timepi02, ... ‘timepi12’. Recall that protocol 1 describes how to copy these configuration files, e.g. to /home/pi/Desktop.

4. Make a folder for images on each of the Raspberry Pis. In our example we made a folder ‘/home/pi/images’. To do this on the Terminal type: ‘mkdir /home/pi/images’. If you wish to use a different path image folder, change line 14 in the example file ‘pull-images-from-raspi.crontab’ (within the appendix.2.time-lapse subdirectory). You will also need to change the ‘photograph-all-5min.crontab’ and ‘photograph-all-5min-vhflipped.crontab’ lines 12, 15, 18 ,21, and 24 so that the images are saved to the correct location during acquisition (both of these scripts are located in Desktop/apps-phenotyping/appendix.2.time-lapse)
5. Physically position each Pi/Camera rig within an experimental growth space (e.g. by screwing adapters into sockets or joining velcro strips together). Take photos (with the raspistill command) to confirm that each camera covers a suitable field of view. See additional notes below on taking photos remotely and optionally flipping photo orientation.

Starting and ending a single imaging experiment:

6. The ‘photograph-all-5min.crontab’ and ‘photograph-all-5min-vhflipped.crontab’ files are cron tables and contain the commands that trigger regular image capture. Both scripts are currently written to capture data every 5 minutes between the hours of 8:30 and 17:30 (8:30am to 5:30pm standard time). If that frequency is

too high, the first number or comma separated list of numbers on lines 12, 15, 18, 21, and 24 have to be altered to reflect that change. If the hours of imaging are different, then the second number or range of numbers on lines 12, 15, 18, 21, and 24 has to be altered.

7. The ‘`photograph-all-5min.crontab`’ and ‘`photograph-all-5min-vhflipped.crontab`’ files also control wifi, turning off outside of the imaging window/photoperiod (see comments in files). Wifi is set to turn on ten minutes before the start of imaging and turn off 10 minutes after imaging ends. If the minute or hour of imaging is different from our experimental setup then the first two numbers on both lines 37 and 43 have to be altered.
8. Once both ‘`photograph-all-5min.crontab`’ and ‘`photograph-all-5min-vhflipped.crontab`’ scripts are satisfactory, install the cron jobs on each Raspberry Pi. The ‘`install-twelve-crontabs.sh`’ script (run from a remote machine, and depending on reasonable wifi connectivity) does this for all twelve Raspberry Pis, but first the user has to determine if the images need to be flipped or not. The difference between the ‘`photograph-all-5min.crontab`’ and ‘`photograph-all-5min-vhflipped.crontab`’ scripts is that the ‘`photograph-all-5min-vhflipped.crontab`’ imaging command flips the images in both the vertical and horizontal directions. Flipping the images might be necessary if there are differences in the orientation of the cameras, and thus images, between the Raspberry Pis. If a Raspberry Pi’s images are in ‘wrong’

orientation, open the ‘install-twelve-crontabs.sh’ file and follow the directions for commenting and uncommenting.

9. To pull data from the Raspberry Pi computers to a remote host, line 15 of the ‘pull-images-from-raspi.crontab’ must be changed to the path that you would like the images to go to on the remote host. The remote host must have passwordless SSH set up so that it can connect to each Raspberry Pi without a password. This is very much like the ‘general instructions to generate a SSH key for passwordless SSH to a remote host’ in Appendix 1, but in reverse. Briefly, on the remote host you would generate an ssh key (command “ssh-keygen”), then copy that to the Raspberry Pi (e.g. “ssh-copy-id -i ~/.ssh/id_rsa.pub pi@timepi01”).
10. Once the RASPIDIR and SERVERDIR paths are changed in ‘pull-images-from-raspi.crontab’, put the ‘pull-images-from-raspi.crontab’ file on the remote host, then install it on the remote host on the command line by typing: ‘crontab pull-images-from-raspi.crontab’. Warning: this will overwrite any preexisting user-specific cron tables.
11. Upon conclusion of an experiment, suspend photography on each Raspberry Pi by “removing” the active crontab (crontab -r). Once the experiment is done, you can safely shutdown the Raspberry Pis using the ‘shut-down-all.sh’ script, if desired.
12. If a cron table job is set up on the Raspberry Pi it will take images as long as the Raspberry Pi has power, disk space, and a functioning camera module.

Additional Notes:

- Here we set up twelve Raspberry Pis for time-lapse imaging, but you may want to set up more or fewer. If fewer than twelve Raspberry Pis are used then simply comment out the excessive commands with a '#' in each of the Appendix 2 scripts. If more than twelve Raspberry Pis are used, then follow the commented code to add more Raspberry Pis with unique hostnames.
- Because of the low cost of the hardware, this approach scales well. If you intend to use a large number of Raspberry Pis (tens, hundreds) you will likely want to investigate management and monitoring tools such as Ansible, Ganglia, and/or Puppet.
- Consider including at least one size marker in each field of view. For example, a white Tough-Spot (Research Products International, Mount Prospect, Illinois, USA) will remain affixed if wet.
- Also consider including color standards or white balance cards (white, gray and black; DGK Color Tools Optek Premium Reference White Balance Cards)
- As noted above, third generation Raspberry Pis have built in wifi and bluetooth modules. Older Raspberry Pi models can be used with USB wifi modules (e.g. Newark 07W8938). Connectivity to our local wireless network from within reach-in growth chambers is generally good, and has been more than sufficient for our monitoring and image transfer purposes. Testing wireless connectivity before setting up Pi/Camera rigs is strongly recommended. Wireless transfer is unlikely to work within a walk-in growth room therefore transferring of data by ethernet is preferable. With a large number of Pi/Camera rigs it is also preferable to transfer

data via ethernet to avoid wireless signal interference. Alternatively, if real-time monitoring is not required and SD card disk space is not a constraint, one or more Raspberry Pis can be left to run autonomously until the conclusion of an experiment, at which point either SD cards can be removed or computers can be moved to another location with better connectivity.

- If wireless connectivity is good, one can run test photo capture commands via a remote connection (and then copy the resulting image files for viewing, e.g. with `rsync`). This removes the need to physically connect a monitor and keyboard to check orientation when mounting each Pi/Camera rig. Minervini et al. (Minervini et al., 2017) have provided instructions for installing and configuring an interface for taking and viewing photos through a web browser.
- The Raspberry Pi NoIR camera can be paired with an infrared (IR) light source to image under low visible light or no visible light conditions. We use a 730 nm cutoff filter (Lee #87) over the NoIR camera lens to block visible light when using an 880 nm LED array to backlight (see link above). The cutoff filter helps prevent changes in contrast during imaging, which makes image processing easier.

Appendix 3. Raspberry Pi Camera Stand. The following are the hardware and software needed to set-up a Raspberry Pi camera stand.

Parts list:

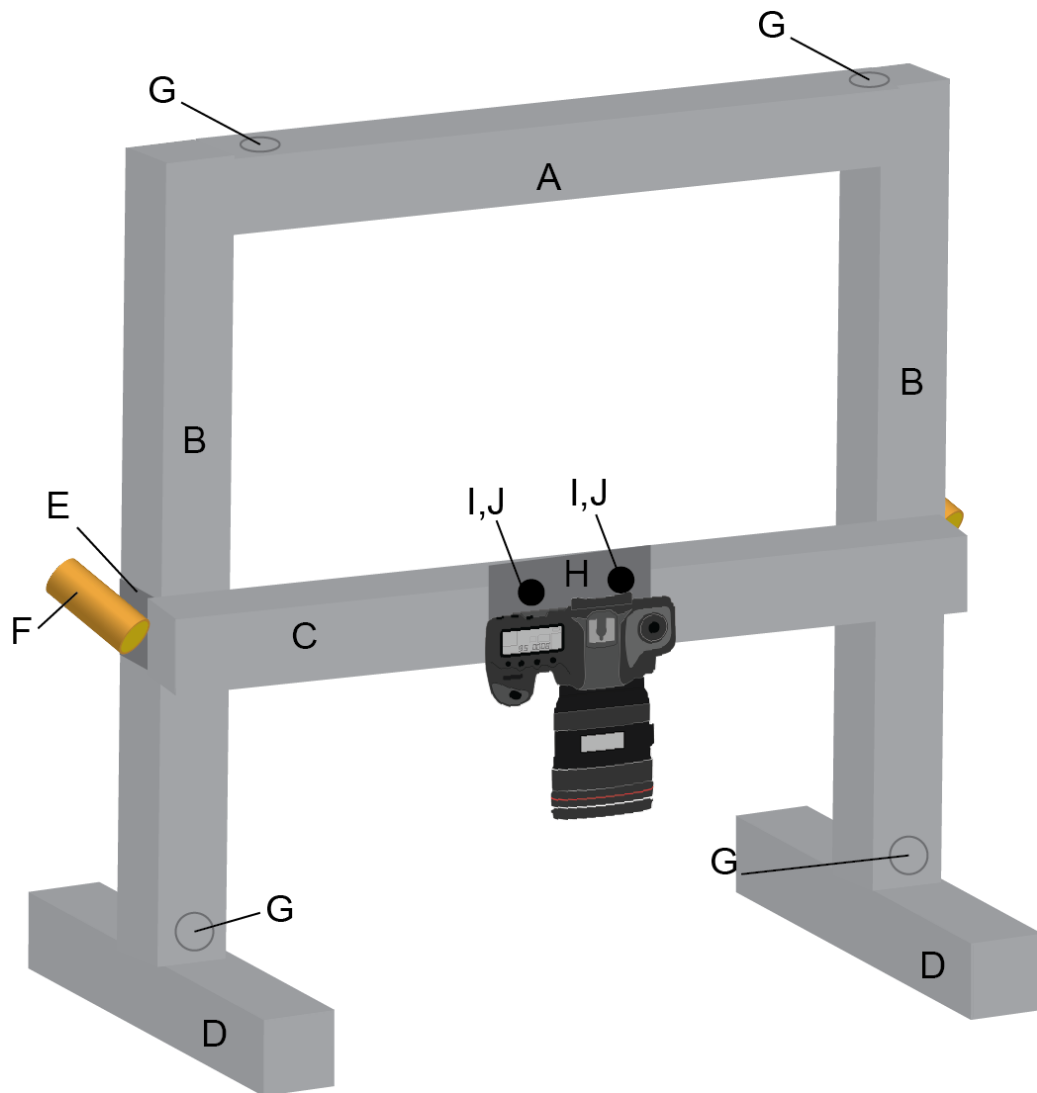
Quantity	Description
1	Initialized Raspberry Pi Computer (see Appendix 1)
1	Raspberry Pi case
1	HDMI cable, mouse, keyboard, and monitor.
1	Nikon Coolpix L830 or other gphoto2 (Figuière and Niedermann, 2017) supported camera
1	Nikon Coolpix power cord
1	Nikon Coolpix L830 USB cable

Aluminum 80/20 Inc. frame parts:

ID	Quantity	80/20 Part Number	Description
A	1	1515-UL	75 cm width x 3.81 cm height x 3.81 cm length T-slotted bar, with 7040 counterbore in A left and A right
B	2	1515-UL	75 cm width x 3.81 cm height x 3.81 cm

			length T-slotted bar, with 7040 counterbore in D left
C	1	1515-UL	88.9 cm width x 3.81 cm height x 3.81 cm length T-slotted bar
D	2	1515-UL	36 cm x 3.81 cm height x 3.81 cm length T-slotted bar
E	2	6525	Double flange short standard linear bearing with brake holes
F	2	6800	15 S gray “L” handle linear bearing brake kit
G	4	3360	15 S 5/16-18 standard anchor fastener assembly
H	1	65-2453	10.16 cm width x 10.16 cm height x 0.3175 cm thick aluminum plate. Three holes will be needed, two to bolt the plate to the crossbar and one hole below the bar to mount the camera with a nut.
I	2	3203	15 series 5/16-18 standard slide in T-nut

J	2	3117	15/16-18 x 0.875" button head socket cap screw
---	---	------	--



Additional set-up of Raspberry Pi for camera stand:

1. In the Applications Menu > Preferences, set hostname to camerastand.

2. Install gphoto2 and libgphoto2 (here stable version 2.5.14 is used, (Figuère and Niedermann, 2017)) by following the installation guide at <https://github.com/gonzalo/gphoto2-updater>.
3. Connect the camera to the Raspberry Pi with the Nikon Coolpix L830 USB cable.
4. Plug in and turn the camera on.
5. Open a Terminal window, enter “gphoto2 --auto-detect” to detect the camera.
6. Optionally, in the Terminal window, enter “gphoto2 --summary” to verify gphoto2 has correctly identified the Nikon Coolpix L830.
7. In the Terminal window, enter “cd Desktop” to change directory to the desktop.
8. Create a folder to store images on the Raspberry Pi desktop, and change the picPath directory in line 25 of camerastand.py to this folder (i.e. /home/pi/Desktop/folder1/)
9. Replace user@remote-host:remote-directory in line 32 of camerastand.py to the camera stand operator’s username, the remote host name, and the directory in the remote host where the images will be stored (e.g. jdoe@serverx:/home/jdoe/camerastand_images). Make sure that an SSH keygen has been generated (see Appendix 1) that will allow the Raspberry Pi to push data to the remote host.

Raspberry Pi camera stand operation protocol:

1. Turn the Raspberry Pi and camera on.
2. Open a Terminal window.

3. Change directory to Desktop (type “cd Desktop”).
4. In the Terminal, use the command “python
`/home/pi/Desktop/apps-phenotyping/appendix.3.camerastand/camerastand.py`
filename” then press enter to acquire and transfer an image, where “filename”
should be replaced by an appropriate filename for the current picture (e.g. python
`/home/pi/Desktop/apps-phenotyping/appendix.3.camerastand/camerastand.py`
speciesx_plant1_treatment1_rep1).

Additional Notes:

- The camera stand allows camera height to be adjusted. We recommend including a size marker in the images to normalize object area during image analysis. We often use a 1.27 cm diameter Tough-Spot (Research Products International, Mount Prospect, Illinois, USA).
- For seed image background, we draw the corners of a box on a white piece of paper or cardboard. We then place color cards (white, gray and black; DGK Color Tools Optek Premium Reference White Balance Card), and the size marker (Tough-Spot; Research Products International, Mount Prospect, Illinois, USA) just outside the box. This ensures that objects to be imaged (e.g. seeds) are within the field of view.
- Images are saved on the Raspberry Pi SD card, as well as in the remote host, in the directories indicated in lines 25 and 32 of camerastand.py, respectively. Alternatively, the rsync command can be changed so that data is deleted from the Raspberry Pi once data transfer has been confirmed. To change the rsync

command so that the image is deleted from the Raspberry Pi once it has been transferred to the remote host, change line 32 in the ‘camerastand.py’ script to

```
“sp.call(["rsync", "-uhrP", picPath, "user@remote-host:remote-directory",  
"--remove-source-files"])"
```

Appendix 4. Raspberry Pi Multi-Image Octagon.

Parts list:

Quantity	Description
1	Laptop or another computer, such as a Raspberry Pi with a monitor, keyboard, and mouse
4	Initialized Raspberry Pi computers (see Appendix 1)
4	Raspberry Pi cases (e.g. SmartiPi brand LEGO-compatible cases for Raspberry Pi model B case and camera module; SmartiCase LLC, Philadelphia PA, USA)
4	Raspberry Pi cameras (RGB)
4	Heavy-duty velcro
1	Power strip
1	White translucent tarp for light diffusion
1	HDMI Monitor, HDMI cable, keyboard, and mouse (for initialization process)

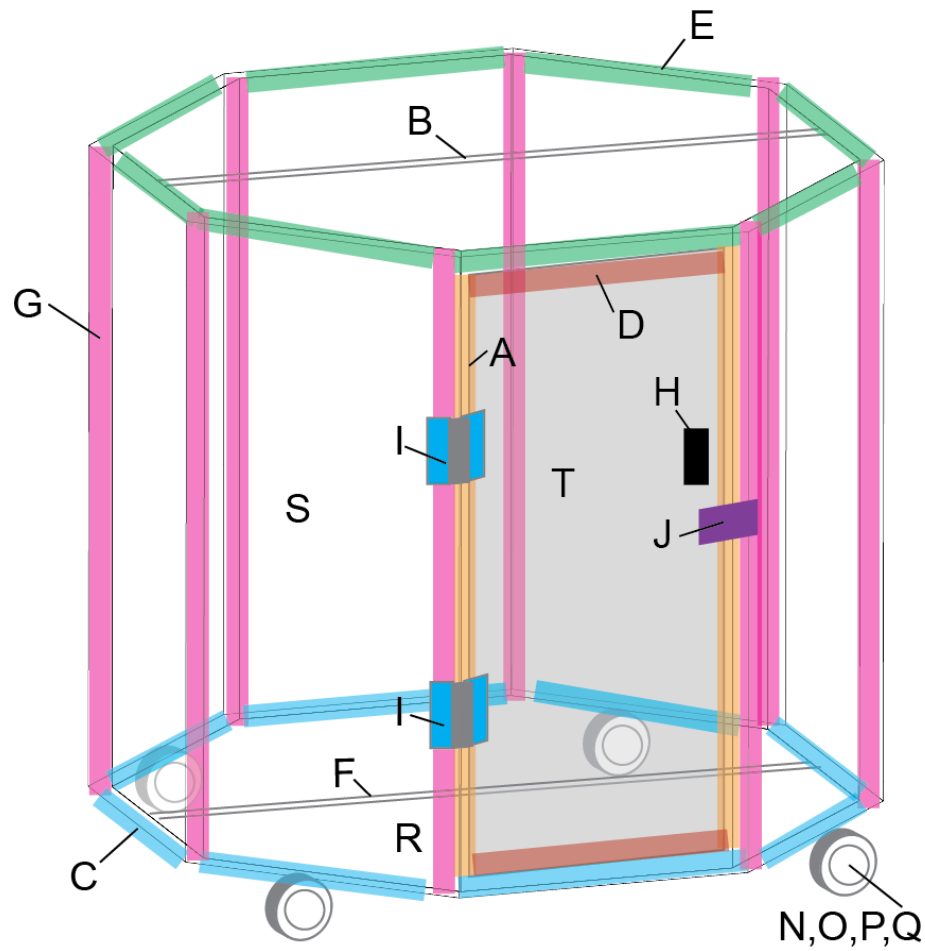
Aluminum 80/20 Inc. frame parts and paneling for Octagon:

ID	Quantity	80/20 Part Number	Description
----	----------	-------------------	-------------

A	2	40-4002	113.919 cm width x 4 cm height x 4 cm length T-slotted bar, with 7040 counterbore in B left and 7040 counterbore in B right.
B	2	40-4002	113.284 cm width x 4 cm height x 4 cm length T-slotted bar. Bi-slot adjacent T-slotted extrusion.
C	8	40-4003	47.1856 cm width x 4 cm height x 4 cm length T-slotted bar, with 7044 counterbore in A left and 7044 counterbore in A right.
D	2	40-4003	30.3911 cm width x 4 cm height x 4 cm length T-slotted bar, with 7044 counterbore in C left; and 7044 counterbore in C right.
E	8	40-4004	47.1856 cm width x 4 cm height x 4 cm length T-slotted bar, with 7044 counterbore in C left and 7044 counterbore in C right.
F	1	40-4080-UL	121.92 cm width x 4 cm height x 8 cm

			length T-slotted bar, with 7044 counterbore in E left, 7044 counterbore in R left, 7044 counterbore in R right and 7044 counterbore in E right.
G	8	40-4094	121.92 cm width x 4 cm height, 40 series T-slotted bar with 45 ° outside radius
H	1	40-2061	Medium plastic door handle, black
I	2	40-2085	40 S aluminum hinge
J	1	65-2053	Deadbolt latch with top latch
K	44	40-3897	Anchor fastener assembly with M8 bolt and standard T-nut
L	8	75-3525	M8 x 1.2 cm black button head socket cap screw (BHSCS) with slide-in economy T-nut
M	2	75-3634	M8 x 1.8 cm black socket head cap screw (SHCS) with slide-in economy T-nut
N	4	40-2426	40 S flange mount caster base plate
O	12	13-8520	M8 x 2 cm SHCS blue

P	8	40-3915	15 S M8 roll-in T-nut with ball spring
Q	4	65-2323	12.7 cm flange mount swivel caster with brake
R	2	2616	129.921 cm width x 64.9605 cm height x 0.3 cm length white PVC panel
S	7	65-2616	116.119 cm width x 49.3856 cm height x 0.3 cm length white PVC panel
T	1	65-2616	107.483 cm width x 32.5911 cm height x 0.3 cm length white PVC panel



Optional parts:

Quantity	Description
1	Barcode scanner (e.g. Socket Mobile 7Qi)

Additional set-up for 4 Raspberry Pis for multi-image octagon:

1. Install Raspberry Pi camera on Raspberry Pi (see Appendix 1).

2. Plug each of the Raspberry Pis into the monitor, keyboard, mouse, and USB power. Following Appendix 1, set hostnames of the four Raspberry Pis to: 1) octagon; 2) sideview1; 3) sideview2; and 4) topviewpi
3. For the Raspberry Pi with the hostname 'octagon', set up passwordless SSH (Appendix 1) so that the 'octagon' Raspberry Pi can trigger scripts on other Raspberry Pis commands. Briefly:
 - a. open a Terminal window, and use the command “ssh-copy-id -i ~/.ssh/id_rsa.pub user@remote-host” for the three other Raspberry Pis (e.g ssh-copy-id -i ~/.ssh/id_rsa.pub pi@sideview1; ssh-copy-id -i ~/.ssh/id_rsa.pub pi@sideview2; ssh-copy-id -i ~/.ssh/id_rsa.pub pi@topviewpi)
 - b. If asked “The authenticity of host 'remote-host' can't be established. Are you sure you want to continue connecting?” enter “yes”.
 - c. Enter the Raspberry Pi password when prompted (the password is “raspberry” if not altered from default).
4. Open a Terminal window, and enter “cd Desktop” to change directory to the desktop. Create folders on each of the Raspberry Pis to temporarily store images on the Raspberry Pi desktop. To facilitate identification of image source, each folder can be given the respective Raspberry Pi's hostname (octagon, sideview1, sideview2, topviewpi).
5. For each of the Raspberry Pis, open the piPicture.py script located at Desktop>apps-phenotyping>appendix.4.octagon.multi-image>piPicture.py.

Change the picPath directory in line 37 of piPicture.py to the folder created in step 4 (e.g. /home/pi/Desktop/octagon).

6. Similarly, for each of the Raspberry Pis, open the syncScript.sh script located at Desktop>apps-phenotyping>appendix.4.octagon.multi-image>syncScript.sh.
Change the rsync local directory in line 1 of syncScript.sh to the folder created in step 4 (e.g. /home/pi/Desktop/octagon).
7. Change the rsync remote directory in line 1 of syncScript.sh to the directory in the remote host where the images will be stored (e.g. jdoe@serverx:/home/jdoe/octagon_images). Make sure that the specified directory exists on the remote host. Remember that passwordless SSH (Appendix 1) must be set up to allow the Raspberry Pi to push data to the remote host.
8. Change '<hostname>' in line 7 of syncScript.sh to the respective Raspberry Pi hostname.
9. Lines 44 to 58 of piPicture.py script set camera parameters using the Picamera package (Jones, 2017). These may need to be adjusted depending on the lighting in the octagon chamber, please refer to the Picamera documentation (<http://picamera.readthedocs.io/en/release-1.10/recipes1.html>) for tips on adjusting parameters. It is important to keep in mind that the Picamera package allows the user to change the camera resolution. If the resolution is set inappropriately for the camera module that is being used (too small, for example) image quality can be reduced.

10. Mount the four Raspberry Pis to the octagon using heavy duty velcro and plug them into a power strip.

Raspberry Pi multi-image octagon operation protocol:

1. Turn all Raspberry Pis on. If desired, put the tarp over the top of the octagon for light diffusion.
2. From your computer (laptop is most convenient), SSH into the octagon Raspberry Pi. Type “ssh pi@octagon” in a Terminal window, and enter the password “raspberry”.
3. To begin imaging, in the Terminal type “bash
/home/pi/Desktop/apps-phenotyping/appendix.4.octagon.multi-image/sshScript.sh
”.
4. When prompted “Please scan barcode or type quit to quit”, type plant id (e.g. speciesx_plant1_treatment1_rep1), or scan a plant id in with a barcode scanner.
5. Place the potted plant into a mounted pot within the octagon (see additional notes)
6. Press enter to acquire images, then wait until prompt “Please scan barcode or type quit to quit” appears again.
7. Repeat steps 6 and 7 to acquire another image, or enter “quit” to quit acquiring images.
8. To shut down sideview1, sideview2, and topviewpi Raspberry Pis, in the Terminal type “bash
/home/pi/Desktop/apps-phenotyping/appendix.4.octagon.multi-image/shutdown_a

ll_pi.sh”. To shut down the octagon Raspberry Pi, in the terminal window type “sudo halt”.

Additional Notes:

- Affixing a pot to the center of the octagon chamber, with color cards affixed to the outside of the stationary pot (white, black and gray; DGK Color Tools, New York, New York, USA) allows a potted plant to be quickly placed in the same relative position to other images.

FIGURE LEGENDS

Figure 1. Low-cost Raspberry Pi phenotyping platforms. A) Raspberry Pi time-lapse imaging in a growth chamber. B) Raspberry Pi camera stand. C) Raspberry Pi multi-image octagon.

Figure 2. Examples of data collected from Raspberry Pi phenotyping platforms that have plant/seed tissue segmented using open-source open-development software PlantCV (Fahlgren et al., 2015). A) PlantCV-segmented image of a flat of *Arabidopsis* acquired from Raspberry Pi time-lapse imaging protocol in a growth chamber. B) PlantCV-segmented image of quinoa seeds acquired from Raspberry Pi camera stand. C) Example side- and top-view images of quinoa plants acquired from Raspberry Pi multi-image octagon. Plant convex hull, width, and length, have been identified with PlantCV and are denoted in red.

FIGURES AND LEGENDS

Figure 1. Low-cost Raspberry Pi phenotyping platforms. A) Raspberry Pi time-lapse imaging in a growth chamber. B) Raspberry Pi camera stand. C) Raspberry Pi multi-image octagon.

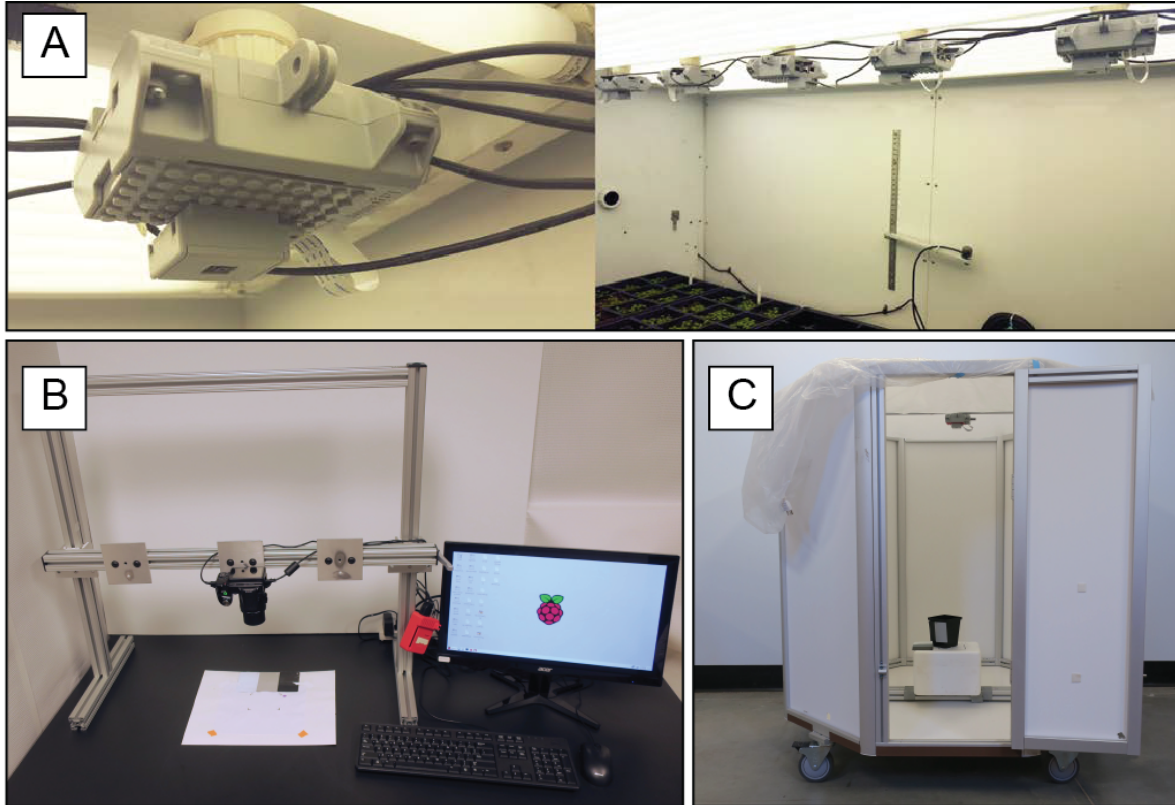
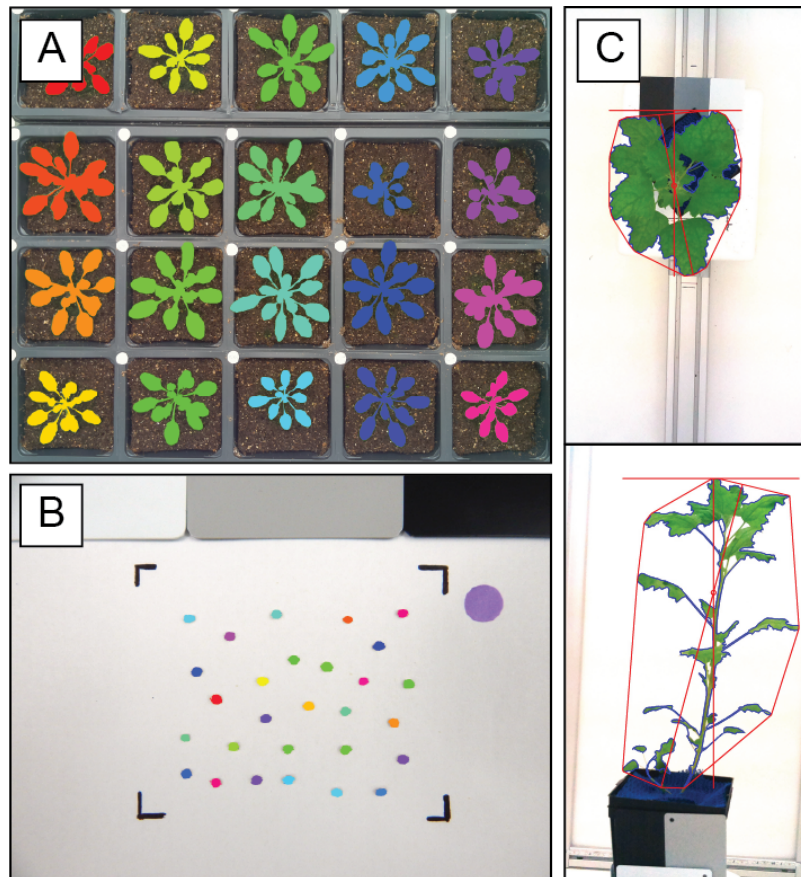


Figure 2. Examples of data collected from Raspberry Pi phenotyping platforms that have plant/seed tissue segmented using open-source open-development software PlantCV (Fahlgren et al., 2015). A) PlantCV-segmented image of a flat of Arabidopsis acquired from Raspberry Pi time-lapse imaging protocol in a growth chamber. B) PlantCV-segmented image of quinoa seeds acquired from Raspberry Pi camera stand. C) Example side- and top-view images of quinoa plants acquired from Raspberry Pi multi-image octagon. Plant convex hull, width, and length, have been identified with PlantCV and are denoted in red.



Appendix C

Online documentation: Configuration for imaging plants with RaspiCams using raspistill, cron, and rsync

The following text accompanies the configuration files I used for time-lapse imaging, as described in chapter 3. They provide additional details beyond the preceding appendix. The files have been archived at Zenodo.org (DOI 10.5281/zenodo.594707) and are also available at <https://github.com/jshoyer/raspi-topdown-plant-imaging-12x>

C.1 Purpose

The scripts and cron tables here are presented as an example of a minimal imaging configuration and as research documentation. It is very unlikely that you will be able to use them directly without substantial modification. In particular, the hostnames for the twelve RasPis (e.g. `ch129-pos01`) are hardcoded into the files. We suggest reading the cron table files and shell scripts. If they seem

useful, modify them to serve your purpose (adjust hardcoded hostnames etc.) and then use them to initiate photo capture with some moderate number of Pi/Camera rigs and (optionally) automate transfer of the files to a remote server.

We use standard GNU/Linux utilities to keep things simple and easy to modify, in keeping with the Raspberry Pi spirit. We encourage plant biologists to experiment with `cron` and `rsync`—they are useful in a wide variety of situations! RaspiCam imaging is a great setting for developing and practicing command line skills.

Copyright © 2017 Donald Danforth Plant Science Center. See LICENSE-MIT.

C.1.1 Links

Useful pages under RaspberryPi.org/documentation/:

- [configuration/camera.md](#)
- [usage/camera/raspicam/time-lapse.md](#)
- [hardware/camera/README.md](#)
- [raspbian/applications/camera.md](#)

We recommend examining examples of similar approaches:

- <http://phenotiki.com> – includes tools for configuration via a web browser.
- Ansible-based configuration for ~10-fold larger imaging setups:

1. <https://github.com/calizarr/PhenoPiSight>

2. Simplified version for topdown RasPiCam imaging: <https://github.com/maliagehan/gehan-bramble>

- Gigavision (code, docs) – works with both RasPiCams and DSLRs. Requires ansible, OpenCV 3, and several python packages, including Flask.

These files here are used for acquiring data. For processing and analysis of the resulting image files you may want to use PlantCV or one of the many other available software packages listed at plant-image-analysis.org

See TAIR page for some information on *A. thaliana*, including a time-lapse video of rosette growth.

C.2 cron tables

C.2.1 Image capture schedule

We install a cron table on each RasPi. There are two variants of this cron table: `photograph-every-5min.crontab` and `photograph-with-flips-every-5min.crontab`. These are nearly identical but the second one does horizontal and vertical flips (`-vf -hf` flags) because five of our Pi/Camera rigs are oriented 180° opposite the others. Files get names like ‘2017-02-10_0300_ch129-pos01.jpg’ where ‘0300’ is the photo capture time, 0300 UTC, which is 10 PM local time, the last capture of the day. ‘ch129-pos01’ is the hostname (chamber 129 position #1, above field-of-view #1). We generally image plants under short-day conditions (8 hours of light, from 8:30 AM to 4:30 PM), with capture throughout the light phase. We also include configuration files (cron tables) long days (16-hour, with light from 6 AM to 10 PM local time) and medium (12-hour) days in subdirectories. See also notes below.

C.2.2 Lights and wifi

Both cron table variants turn off wireless connectivity (`ifdown wlan0`) for most of the night, to avoid exposing plants to blue LED light from the USB WiPi dongles we use. (This will not be an issue if you use the built-in wifi module on a model 3 Raspberry Pi board.) We additionally turn off the green and red indicator LEDs on each RasPi board by editing the `/boot/config.txt` file on each RasPi:

```
dtparam=act_led_trigger=none
dtparam=act_led_activelow=off
```

```
dtparam=pwr_led_trigger=none
dtparam=pwr_led_activelow=off
```

With these parameters set, light from the red LED effectively becomes an indicator that a RasPi has crashed (but is still drawing sufficient power to function), as opposed to simply suffering from network connectivity problems.

We similarly use a pair of minimal cron tables to similarly avoid exposing plants to blue light in between imaging experiments (1, 2).

C.2.3 Image transfer

File transfer is scheduled with a separate cron table installed on a server: `pull-images-from-raspis.crontab`. Syncing photos to a cluster makes them easier to process/monitor, and lets us collect more photos than will fit on a single SD card.

C.2.4 Helper scripts

Two shell scripts are provided to streamline starting and ending image capture on all twelve RasPis at once.

- The first script is used at the start of an experiment (or after adjusting a cron table). This file copies the appropriate cron table onto each of the twelve RasPis (if necessary) and then installs it. The script is run without arguments: `./install-twelve-crontabs`
- A second script (`install-twelve-mini-crontabs`) is used at the end of an experiment (see below) to install the minimal cron tables described above and thereby suspend image capture.

The next section describes use of a third helper script for taking and viewing a single snapshot.

C.3 Procedures

C.3.1 Starting imaging

1. Take snapshots to convince yourself that plants are positioned appropriately. This is an excellent time to photograph color standard cards, to enable later assessment of sensor drift.

To take a snapshot with the RasPi in position #1, run the script like so:

```
./take-one-picture-and-pull-it-with-rsync /path/on/cluster/ 1
```

- You could run this command (and the next one) from your local computer, but things will be easier if you run them on a remote server.

2. Install the correct cron table on each RasPi (as mentioned above) to start regular image capture: `./install-twelve-crontabs`
3. Double check the server cron table. Is the correct (hardcoded) destination path on the server specified?
4. Install the server cron table to pull photos: `crontab pull-images-from-raspis.crontab`

C.3.2 Ending imaging

1. Stop image capture; reinstall cron tables that just monitor lights and cycle wifi on and off:
`./install-twelve-mini-crontabs`
2. If desired, photograph color standard cards, as you remove plants or shortly thereafter, as above: `./take-one-picture-and-pull-it-with-rsync /path/on/cluster/ 1`

C.3.3 Pitfalls

1. Watch out for color drift and consider including standards in your field of view. By default, `raspistill` automatically picks exposure and color balance settings based on a five second video preview. This has been sufficient for our purposes and provides a starting point for testing other settings, but it means that the white balance and capture conditions can vary over the course of an experiment. In particular, the blue rubber mesh often placed over soil for image-based phenotyping experiments (see e.g. Junker et al. 2014, Figures 3 and 4) can cause color balance "overcompensation", resulting in an orange tinge. This tinge steadily recedes over the course of an experiment (as plant leaves cover the mesh), which further complicates image processing.

- We embed raw Bayer data into JPEG file exif metadata (`raspistill -r` flag) to enable post-processing, but only for the first and last capture of the day.
 - Lots of room for improvement here!
2. The clock built into our growth chamber control board does not automatically recalibrate itself by synchronizing with a server, and so the clock steadily drifts forward, at a rate of ~3 seconds per day (~1 minute every three weeks). Unless the clock is manually corrected, the light schedule will eventually shift far enough that the first photo of the day will be captured before "sunrise".
- The most reliable way to deal with this issue is to manually calibrate the chamber clock shortly before the start of every new experiment. Adjusting the clock in our growth chamber requires shutting it down, which in turn necessitates turning off each RasPi. (Our twelve RasPis use a GFCI-protected auxiliary power outlet built into the growth chamber, via an extension cord threaded through a port built into the exterior of the chamber.) We use a shell script to shut down our twelve RasPis, and they turn back on automatically after power is restored.
3. We have used our local timezone in the past, but now recommend using Universal Coordinated Time (UTC) to avoid potential for confusion and/or loss of data caused by the start and end of daylight saving time. If you are not using UTC (controlled via `raspi-config` internationalization settings), the start and end of daylight saving time may trigger an automatic clock shift on each RasPi, which can result in the photo capture schedule being offset by one hour relative to the light cycle.
- We generally image from 9:30 AM to 5:30 PM local time during DST (CDT is UTC -0500) and 8:30 PM to 4:30 PM for the rest of the year (CST is UTC -0600). These are both equivalent to 1430 to 2230 UTC.

- Some growth chamber controllers automatically shift the light cycle at the start and end of daylight saving time. This shift is arguably bad, because re-entrainment of plant circadian clocks to the new light schedule can alter growth. Shifting the start of zeitgeber (ZT) time also makes the experiment more difficult to describe.
 - Switching the timezone on a RasPi takes effect without requiring a reboot, but this will **not** alter cron scheduling until you `sudo service cron restart`
4. Make sure your RasPis are drawing sufficient power! The camera boards draw extra power during photo capture, which can cause one or more RasPis sharing an inadequate power supply to crash. See <https://www.raspberrypi.org/documentation/hardware/raspberrypi/power/README.md>

C.3.4 Transfer schedule

Edge cases (especially the start and end of the day) create some potential for error when editing cron tables. The server cron table is perhaps the most complicated, because of the way I manually staggered file transfers (to try to reduce wifi signal interference across the collection of RasPis). Fortunately, most errors in the server cron table will merely cause error messages/overnight sync delays—errors should not cause loss of data.

The following table of UTC times may help you interpret the server cron table for 16-hour days. First photo capture each day is at 1105 UTC, and the last photo capture each day is at 0300 UTC.

	First pull	Last pull
1	11:06	03:06
2	11:08	03:08
3	11:11	03:11
4	11:13	03:13
5	11:16	03:16
8	11:18	03:16
9	11:21	03:21
10	11:23	03:23
11	11:26	03:26
12	11:28	03:28
6	11:31	03:01
7	11:33	03:03

C.3.5 Monitoring

If desired, one can add an email address (MAILTO variable) at the top of the server cron table. This contact address will then receive an email every time an rsync transfer fails. This measure is noisy: a failed transfer is usually caused by transient wifi interference, and merely delays transfer of the relevant files until the next cycle. Multiple failed transfers can indicate that a RasPi has crashed, especially when initial connection was the step that failed. (Interrupted transfers are a lagging indicator, because rsync processes persist for quite a while before they "give up.")

We additionally use a Ganglia dashboard for monitoring. See <http://ganglia.info>

C.4 Plans

Researchers at the Danforth Center will likely continue using these scripts for imaging experiments. We plan to share any improvements we make, but it is also possible that we will supplant this code with something else entirely. To reiterate: we make these files public primarily as a learning aid and as documentation for related research papers.

Questions, feedback, and contributions are welcome via [GitHub](#), [Bitbucket](#), or [GitLab.com](#).

Appendix D

PlantCV v2: Image analysis software for high-throughput plant phenotyping

D.1 Preface

In the course of the work described in chapter 3, I used and contributed to the PlantCV software package developed at the Danforth Center. The manuscript that follows describes the second major release of this package; the preprint version is available at <https://peerj.com/preprints/3225>. The initial submission received a “minor revisions required” editorial decision at PeerJ and has been resubmitted. The first version of PlantCV was described by Fahlgren, Feldman, Gehan et al. [234], and one section from that paper is particularly relevant:

We developed the open-source and open-development PlantCV image analysis platform to emphasize the following features: flexible user-defined analysis workflows; parallelizable image processing for fast throughput; and a scripting language implementation that lowers the barrier to community contributions that extend functionality. It was important to move away from commercial software for greater control and understanding of the image processing and trait extraction algorithms used to process the data, as well as the freedom to expand analyses at will. While some users may

prefer graphical user interfaces for software, script-based programs are easier to develop and the precise workflows are detailed directly in the scripts themselves, enabling reproducible research.

While PlantCV does not have a graphical user interface *per se*, it has extensive facilities for inspecting output (i.e. image files written to disk), both for building geometric intuition and for testing whether image analysis pipelines are working as desired. (“debugging”). As described in the paper that follows, the second release enables integration with the Jupyter Notebook framework [395], both for accelerated prototyping (via rapid feedback for the user) and for in-line integration of prose, code, and output “chunks”.

Surprisingly, many papers on plant image analysis are not accompanied by any source code; this includes papers that purport to describe a general-purpose extensible tool [396]. Personally, I have found that this has made these papers much more difficult to understand than necessary. Including a license that fits the Open Source Definition with such code would seem to be a simple thing, but is neglected surprisingly often [396], which can create serious intellectual property issues, particularly for companies seeking to use public code.

My main contributions to the computational work described here were in testing, writing documentation, and sharing data. Some of the new functions I have written or adapted have not yet been integrated into the “main line” of PlantCV development. Full details can be seen in the history of the source tree (maintained at <https://github.com/danforthcenter/plantcv>) and on relevant Issue and Pull Request discussion pages. I provided the photo used for Figures 2 and 3, but did not directly make any of the figures. I also participated in outlining and revising the manuscript.

PlantCV v2: Image analysis software for high-throughput plant phenotyping

Malia A Gehan^{*,1}, Noah Fahlgren^{*,1}, Arash Abbasi¹, Jeffrey C Berry¹, Steven T Callen^{1,2}, Leonardo Chavez¹, Andrew N Doust³, Max J Feldman¹, Kerrigan B Gilbert¹, John G Hodge³, J Steen Hoyer^{1,4}, Andy Lin^{1,5}, Suxing Liu^{6,7}, César Lizárraga^{1,8}, Argelia Lorence⁹, Michael Miller^{1,10}, Eric Platon¹¹, Monica Tessman^{1,2}, Tony Sax¹²

¹ Donald Danforth Plant Science Center, St. Louis, Missouri, United States of America

² Current Address: Monsanto Company, St. Louis, Missouri, United States of America

³ Department of Plant Biology, Ecology, and Evolution, Oklahoma State University, Stillwater, Oklahoma, United States of America

⁴ Computational and Systems Biology Program, Washington University in St. Louis, St. Louis, Missouri, United States of America

⁵ Current Address: Unidev, St. Louis, Missouri, United States of America

⁶ Arkansas Biosciences Institute, Arkansas State University, Jonesboro, Arkansas, United States of America

⁷ Current Address: Department of Plant Biology, University of Georgia, Athens, Georgia, United States of America

⁸ Current Address: CiBO Technologies, Cambridge, Massachusetts, United States of America

⁹ Arkansas Biosciences Institute, Department of Chemistry and Physics, Arkansas State University, Jonesboro, Arkansas, United States of America

¹⁰ Current Address: Department of Agronomy and Horticulture, Center for Plant Science Innovation, Beadle Center for Biotechnology, University of Nebraska - Lincoln, Lincoln, Nebraska, United States of America

¹¹ Cosmos X, Tokyo, Japan

¹² Missouri University of Science and Technology, Rolla, Missouri, United States of America

* These authors contributed equally to this work

Corresponding Author:

Malia Gehan; Noah Fahlgren

975 North Warson Road, St. Louis, MO 63132, USA

Email address: mgehan@danforthcenter.org; nfahlgren@danforthcenter.org

ABSTRACT

Systems for collecting image data in conjunction with computer vision techniques are a powerful tool for increasing the temporal resolution at which plant phenotypes can be measured non-destructively. Computational tools that are flexible and extendable are needed to address the diversity of plant phenotyping problems. We previously described the Plant Computer Vision (PlantCV) software package, which is an image processing toolkit for plant phenotyping analysis. The goal of the PlantCV project is to develop a set of modular, reusable, and repurposable tools for plant image analysis that are open-source and community-developed. Here we present the details and rationale for major developments in the second major release of PlantCV. In addition to overall improvements in the organization of the PlantCV project, new functionality includes a set of new image processing and normalization tools, support for analyzing images that include multiple plants, leaf segmentation, landmark identification tools for morphometrics, and modules for machine learning.

INTRODUCTION

All approaches for improving crops eventually require measurement of traits (phenotyping) (Fahlgren, Gehan & Baxter, 2015). However, manual plant measurements are time-consuming and often require destruction of plant materials in the process, which prevents measurement of traits for a single plant through time. Consequently, plant phenotyping is widely recognized as a major bottleneck in crop improvement (Furbank & Tester, 2011). Targeted plant phenotypes can range from measurement of gene expression, to flowering time, to grain yield; therefore, the software and hardware tools used are often diverse. Here, we focus on the software tools required to nondestructively measure plant traits through images. This is a challenging area of research because the visual definition of phenotypes vary depending on the target species. For example, identification of petals can be used to measure flowering time, but petal color can vary by species. Therefore, software tools needed to process high-throughput image data need to be flexible and amenable to community input.

The term ‘high-throughput’ is relative to the difficulty to collect the measurement. The scale that might be considered high-throughput for root phenotyping might not be the same for shoot phenotyping, which can be technically easier to collect depending on the trait and species. Here we define high-throughput as thousands or hundreds of thousands of images per dataset. PlantCV is an open-source, open-development suite of analysis tools capable of analyzing high-throughput image-based phenotyping data (Fahlgren et al., 2015). Version 1.0 of PlantCV (PlantCV v1.0) was released in 2015 alongside the introduction of the Bellwether Phenotyping Facility at the Donald Danforth Plant Science Center (Fahlgren et al., 2015). PlantCV v1.0 was envisioned as a base suite of tools that the community could build upon, which lead to several

design decisions aimed at encouraging participation. First, GitHub was used as a platform to organize the community by integrating version control, code distribution, documentation, issue tracking, and communication between users and contributors (Perez-Riverol et al., 2016). Second, PlantCV was written in Python, a high-level language widely used for both teaching and bioinformatics (Mangalam, 2002; Dudley & Butte, 2009), to facilitate contribution from both biologists and computer scientists. Additionally, the use of Python allows extension of PlantCV with the many tools available from the Python scientific computing community (Oliphant, 2007; Millman & Aivazis, 2011). Third, a focus on modular development fosters code reuse and makes it easier to integrate PlantCV with new or existing systems. Finally, the use of a permissive, open-source license (MIT) allows PlantCV to be used, reused, or repurposed with limited restrictions, for both academic and proprietary applications. The focus of the paper associated with the original release of PlantCV v1.0 (Fahlgren et al., 2015) was not the structure and function of PlantCV for image analysis, but rather an example of the type of biological question that can be answered with high-throughput phenotyping hardware and software platforms. Since the release of PlantCV v1.0 major improvements have been made to increase the flexibility, usability, and functionality of PlantCV, while maintaining all of the functionality in v1.0. Here we document the structure of PlantCV v2 along with examples that demonstrate new functionality.

MATERIALS & METHODS

The latest version or a specific release of PlantCV can be cloned from GitHub. The release for this paper is v2.1. Scripts, notebooks, SQL schema, and simple input data associated with the figures and results presented in this paper are available on GitHub at

<https://github.com/danforthcenter/plantcv-v2-paper>. Project-specific GitHub repositories are kept separate from the PlantCV software repository because their purpose is to make project-specific analyses available for reproducibility, while the main PlantCV software repository contains general purpose image analysis modules, utilities, and documentation.

Images of *Arabidopsis thaliana* were captured with a Raspberry Pi computer and camera in a Conviron growth chamber. Additional details about the imaging set-up are provided in a companion paper (Tovar et al., 2017). Images of *Setaria viridis* (A10) and *Setaria italica* (B100) are from publicly available datasets that are available at <http://plantcv.danforthcenter.org/pages/data.html> (Fahlgren et al., 2015; Feldman et al., 2017). Images of wheat (*Triticum aestivum* L.) infected with wheat stem rust (*Puccinia graminis* f. sp. *tritici*) were acquired with a flatbed scanner.

Image analysis was done in PlantCV using Python v2.7.5, OpenCV v2.4.5 (Bradski, 2000), NumPy v1.12.1 (van der Walt, Colbert & Varoquaux, 2011), Matplotlib v2.0.2 (Hunter, 2007), SciPy v0.19.0 (Jones, Oliphant & Peterson), Pandas v0.20.1 (McKinney & Others, 2010), scikit-image v0.13.0 (van der Walt et al., 2014), and Jupyter Notebook v4.2.1 (Kluyver et al., 2016). Statistical analysis and data visualization was done using R v3.3 (R Core Team, 2017) and RStudio v1.0 (RStudio Team, 2016). Graphs were produced using Matplotlib v2.0.2 (Hunter, 2007) and ggplot2 v2.2.1 (Wickham, 2009).

RESULTS AND DISCUSSION

The following are details on improvements to the structure, usability, and functionality of PlantCV since the v1.0 release. Further documentation for using PlantCV can be found at the project website (<http://plantcv.danforthcenter.org/>).

Organization of the PlantCV project

PlantCV is a collection of modular Python functions, which are reusable units of Python code with defined inputs and outputs (Fig. 1A). PlantCV functions can be assembled into simple sequential or branching/merging pipelines. A pipeline can be as long or as short as it needs to be, allowing for maximum flexibility for users using different imaging systems and analyzing features of seed, shoot, root, or other plant systems. Suggestions on how to approach image analysis with PlantCV, in addition to specific tutorials, are available through online documentation (http://plantcv.readthedocs.io/en/latest/analysis_approach/). Each function has a debugging option to allow users to view and evaluate the output of a single step and adjust parameters as necessary. A PlantCV pipeline is written by the user as a Python script. Once a satisfactory pipeline script is developed, the PlantCV parallelization script ('plantcv-pipeline.py') can be used to deploy the pipeline across a large set of image data (Fig. 1A). The parallelization script also functions to manage data by consolidating measurements and metadata into an SQLite database (Fig. 1B). In terms of speed, the user is only limited by the complexity of the pipeline and the number of available processors.

The modular structure of the PlantCV package makes it easier for members of the community to become contributors. Contributors to PlantCV submit bug reports, develop new functions and

unit tests, or extend existing functionality or documentation. Core PlantCV developers do not filter additions of new functions in terms of perceived impact or number of users but do check that new functions follow the PlantCV contribution guide (see the sections on contributing in the online documentation). PlantCV contributors are asked to follow the PEP8 Python style guide (<https://www.python.org/dev/peps/pep-0008/>). Additions or revisions to the PlantCV code or documentation are submitted for review using pull requests via GitHub. The pull request mechanism is essential to protect against merge conflicts, which are sections of code that have been edited by multiple users in potentially incompatible ways.

In PlantCV v2, several service integrations were added to automate common tasks during pull requests and updates to the code repository. A continuous integration framework using the Travis CI service (<https://travis-ci.org/>) was added so that software builds and unit tests can be run automatically upon pull requests and other software updates. Continuous integration provides a safeguard against code updates that break existing functionality by providing a report that shows which tests passed or failed for each build (Wilson et al., 2014). The effectiveness of continuous integration depends on having thorough unit test coverage of the PlantCV code base. Unit test coverage of the PlantCV Python package is monitored through the Coveralls service (<https://coveralls.io/>), which provides a report on which parts of the code are covered by existing unit tests. In addition to the code, the PlantCV documentation was enhanced to use a continuous documentation framework using the Read the Docs service (<https://readthedocs.org/>), which allows documentation to be updated automatically and versioned in parallel with updates to PlantCV. The documentation was updated to cover all functions in the PlantCV library, tutorials on building pipelines and using specialized tools (e.g. multi-plant analysis and machine learning

tools), a frequently asked questions section, and several guides such as installation, Jupyter notebooks, and instructions for contributors.

Improved usability

PlantCV v1.0 required pipeline development to be done using the command line, where debug mode is used to write intermediate image files to disk for each step. In command-line mode, an entire pipeline script must be executed, even if only a single step is being evaluated. To improve the pipeline and function development process in PlantCV v2, the debugging system was updated to allow for seamless integration with the Jupyter Notebook system (<http://jupyter.org/>) (Kluyver et al., 2016). Jupyter compatibility allows users to immediately visualize output and to iteratively rerun single steps in a multi-step PlantCV pipeline, which makes parameters like thresholds or regions of interest much easier to adjust. Once a pipeline is developed in Jupyter, it can then be converted into a Python script that is compatible with PlantCV parallelization (see online documentation for detailed instructions on conversion; <http://plantcv.readthedocs.io/en/latest/jupyter/>). Because of the web-based interface and useful export options, Jupyter notebooks are also a convenient method of sharing pipelines with collaborators, or in publications, and teaching others to use PlantCV.

PlantCV was initially created to analyze data generated by the Bellwether Phenotyping Facility at the Donald Danforth Plant Science Center. Several updates to PlantCV v2 addressed the need to increase the flexibility of PlantCV to analyze data from other plant phenotyping systems. The PlantCV SQLite database schema was simplified so that new tables do not need to be added for every new camera system (Fig. 1B). The full database schema is available on GitHub (see

Materials and Methods) and in PlantCV documentation. New utilities were added to PlantCV v2 that allow data to be quickly and efficiently exported from the SQLite database into text files that are compatible with R (R Core Team, 2017) for further statistical analysis and data visualization.

Because standards for data collection and management for plant phenotyping data are still being developed (Pauli et al., 2016), image metadata is often stored in a variety of formats on different systems. A common approach is to include metadata within image filenames, but because there is a lack of file naming standards, it can be difficult to robustly capture this data automatically. In PlantCV v2, a new metadata processing system was added to allow for flexibility in file naming both within and between experiments and systems. The PlantCV metadata processing system is part of the parallelization tool and works by using a user-provided template to process filenames. User-provided templates are built using a restricted vocabulary so that metadata can be collected in a standardized way. The vocabulary used can be easily updated to accommodate future community standards.

Performance

In PlantCV v1.0, image analysis parallelization was achieved using a Perl-based multi-threading system that was not thread-safe, which occasionally resulted in issues with data output that had to be manually corrected. Additionally, the use of the Python package Matplotlib (Hunter, 2007) in PlantCV v1.0 limited the number of usable processors to 10-12. For PlantCV v2, the parallelization framework was completely rewritten in Python using a multiprocessing framework, and the use of Matplotlib was updated to mitigate the issues and processor constraints in v1.0. The output of image files mainly used to assess image segmentation quality

is now optional, which should generally increase computing performance. Furthermore, to decentralize the computational resources needed for parallel processing and prepare for future integration with high-throughput computing resources that use file-in-file-out operations, results from PlantCV pipeline scripts (one per image) are now written out to temporary files that are aggregated by the parallelization tool after all image processing is complete.

New Functionality

PlantCV v2 has added new functions for image white balancing, auto-thresholding, size marker normalization, multi-plant detection, combined image processing, watershed segmentation, landmarking, and a trainable naive Bayes classifier for image segmentation (machine learning). The following are short descriptions and sample applications of new PlantCV functions.

White balancing

If images are captured in a greenhouse, growth chamber, or other situation where light intensity is variable, image segmentation based on global thresholding of image intensity values can become variable. To help mitigate image inconsistencies that might impair the ability to use a single global threshold and thus a single pipeline over a set of images, a white balance function was developed. If a white color standard is visible within the image, the user can specify a region of interest. If a specific area is not selected then the whole image is used. Each channel of the image is scaled relative to the reference maximum.

Auto-thresholding functions

An alternative approach to using a fixed, global threshold for image segmentation is to use an auto-thresholding technique that either automatically selects an optimal global threshold value or introduces a variable threshold for different regions in an image. Triangle, Otsu, mean, and Gaussian auto-thresholding functions were added to PlantCV to further improve object detection when image light sources are variable. The ‘triangle_auto_threshold’ function implements the method developed by Zack et al. 1977 (Zack, Rogers & Latp, 1977). The triangle threshold method uses the histogram of pixel intensities to differentiate the target object (plant) from background by generating a line from the peak pixel intensity (Duarte, 2015) to the last pixel value and then finding the point (i.e., the threshold value) on the histogram that maximizes distance to that line. In addition to producing the thresholded image in debug mode, the ‘triangle_auto_threshold’ function outputs the calculated threshold value and the histogram of pixel intensities that was used to calculate the threshold. In cases where the auto-threshold value does not adequately separate the target object from background, the threshold can be adjusted by modifying the stepwise input. Modifying the stepwise input shifts the distance calculation along the x-axis, which subsequently calculates a new threshold value to use.

The Otsu, mean, and Gaussian threshold functions in PlantCV are implemented using the OpenCV library (Bradski, 2000). Otsu’s binarization [‘otsu_auto_threshold;’ (Otsu, 1979)] is best implemented when a grayscale image histogram has two peaks since the Otsu method selects a threshold value that minimizes the weighted within-class variance. In other words, the Otsu method identifies the value between two peaks where the variances of both classes are minimized. Mean and Gaussian thresholding are executed by indicating the desired threshold

type in the function ‘adaptive_threshold.’ The mean and Gaussian methods will produce a variable local threshold where the threshold value of a pixel location depends on the intensities of neighboring pixels. For mean adaptive thresholding, the threshold of a pixel location is calculated by the mean of surrounding pixel values; for Gaussian adaptive thresholding, the threshold value of a pixel is the weighted sum of neighborhood values using a Gaussian window (Gonzalez & Woods, 2002; Kaehler & Bradski, 2016).

Gaussian blur

In addition to the ‘median_blur’ function included in PlantCV v1.0, we have added a Gaussian blur smoothing function to reduce image noise and detail. Both the median and Gaussian blur methods are implemented using the OpenCV library (Bradski, 2000) and are typically used to smooth a grayscale image or a binary image that has been previously thresholded. Image blurring, while reducing detail, can help remove or reduce signal from background noise (e.g. edges in imaging cabinets), generally with minimal impact on larger structures of interest.

Utilizing a rectangular neighborhood around a center pixel, ‘median_blur’ replaces each pixel in the neighborhood with the median value. Alternatively, ‘gaussian_blur’ determines the value of the central pixel by multiplying its and neighboring pixel values by a normalized kernel and then averaging these weighted values (i.e., image convolution) (Kaehler & Bradski, 2016). The extent of image blurring can be modified by increasing (for greater blur) or decreasing the kernel size (which takes only odd numbers; commonly, 3x3) or by changing the standard deviation in the X and/or Y directions.

Size marker normalization

Images that are not collected from a consistent vantage point require one or more size markers as references for absolute or relative scale. The size marker function allows users to either detect a size marker within a user-defined region of interest or to select a specific region of interest to use as the size marker. The pixel area of the marker is returned as a value that can be used to normalize measurements to the same scale. For this module to function correctly we assume that the size marker stays in frame, is unobstructed, and is relatively consistent in position throughout a dataset, though some movement is allowed as long as the marker remains within the defined marker region of interest.

Multi-plant detection

There is growing interest among the PlantCV user community to process images with multiple plants grown in flats or trays, but PlantCV v1.0 was built to process images containing single plants. The major challenge with analyzing multiple plants in an image is successfully identifying individual whole plants as distinct objects. Leaves or other plant parts can sometimes be detected as distinct contours from the rest of the plant and need to be grouped with other contours from the same plant to correctly form a single plant/target object. While creating multiple regions of interest (ROI) to demarcate each area containing an individual plant/target is an option, we developed two modules, 'cluster_contours' and 'cluster_contours_split_img,' that allow contours to be clustered and then parsed into multiple images without having to manually create multiple ROIs (Fig. 2).

The 'cluster_contours' function takes as input: an image, the contours that need to be clustered, a number of rows, and a number of columns. Total image size is detected, and the rows and columns create a grid to serve as approximate ROIs to cluster the contours. The number of rows and columns approximate the desired size of the grid cells. There does not need to be an object in each of the grid cells. Several functions were also added to aid the clustering function. The 'rotate_img' and 'shift_img' functions allow the image to be adjusted so objects are better aligned to a grid pattern.

After objects are clustered, the 'cluster_contour_split_img' function splits images into the individual grid cells and outputs each as a new image so that there is a single clustered object per image. If there is no clustered object in a grid cell, no image is outputted. With the 'cluster_contour_split_img' function, a text file with genotype names can be included to add them to image names. The 'cluster_contour_split_img' function also checks that there are the same number of names as objects. If there is a conflict in the number of names and objects, a warning is printed and a correction is attempted. Alternatively, if the file option is not used, all of the object groups are labeled by position. Once images are split, they can be processed like single plant images using additional PlantCV tools. See the online documentation for an example multi-plant imaging pipeline (http://plantcv.readthedocs.io/en/latest/multi-plant_tutorial/).

The current method for multi-plant identification in PlantCV is flexible but relies on a grid arrangement of plants, which is common for controlled-environment-grown plants. Future releases of PlantCV may incorporate additional strategies for detection and identification of

plants, such as arrangement-independent *K*-means clustering approaches (Minervini, Abdelsamea & Tsafaris, 2014).

Combined image processing

The Bellwether Phenotyping Facility has both RGB visible light (VIS) and near-infrared (NIR) cameras, and images are captured ~1 minute apart (Fahlgren et al., 2015). Compared to VIS images, NIR images are grayscale with much less contrast between object and background. It can be difficult to segment plant material from NIR images directly, even with edge detection steps. Therefore, several functions were added to allow the plant binary mask that results from VIS image processing pipelines to be resized and used as a mask for NIR images. Combining VIS and NIR camera pipelines also has the added benefit of decreasing the number of steps necessary to process images from both camera types, thus increasing image processing throughput. The ‘get_nir’ function identifies the path of the NIR image that matches VIS image. The ‘get_nir’ function requires that the image naming scheme is consistent and that the matching image is in the same image directory. The ‘resize’ function then resizes the VIS plant mask in both the x and y directions to match the size of the NIR image. Resizing values are determined by measuring the same reference object in an example image taken from both VIS and NIR cameras (for example the width of the pot or pot carrier in each image). The ‘crop_position_mask’ function is then used to adjust the placement of the VIS mask over the NIR image and to crop/adjust the VIS mask so it is the same size as the NIR image. It is assumed that the pot position changes consistently between VIS and NIR image datasets. An example VIS/NIR dual pipeline to follow can be accessed online (http://plantcv.readthedocs.io/en/latest/vis_nir_tutorial/).

Object count estimation with watershed segmentation

While segmentation and analysis of whole plants in images provides useful information about plant size and growth, a more detailed understanding of plant growth and development can be obtained by measuring individual plant organs. However, fully automated segmentation of individual organs such as leaves remains a challenge, due to issues such as occlusion (Scharr et al., 2016). Multiple methods for leaf segmentation have been proposed (Scharr et al., 2016), and in PlantCV v2 we have implemented a watershed segmentation approach. The ‘watershed_segmentation’ function can be used to estimate the number of leaves for plants where leaves are distinctly separate from other plant structures (e.g. *A. thaliana* leaves are separated by thin petioles; Fig. 3). The inputs required are an image, an object mask, and a minimum distance to separate object peaks. The function uses the input mask to calculate a Euclidean distance map (Liberti et al., 2014). Marker peaks calculated from the distance map that meet the minimum distance setting are used in a watershed segmentation algorithm (van der Walt et al., 2014) to segment and count the objects. Segmented objects are visualized in different colors, and the number of segmented objects is reported (Fig. 3). An example of how the watershed segmentation method was used to assess the effect of water deficit stress on the number of leaves of *A. thaliana* plants can be found in Acosta-Gamboa et al. 2017 (Acosta-Gamboa et al., 2017).

Landmarking functions for morphometrics

To extend PlantCV beyond quantification of size-based morphometric features, we developed several landmarking functions. Landmarks are generally geometric points located along the contours of a shape that correspond to homologous biological features that can be compared

between subjects (Bookstein, 1991). Typical examples of landmarks include eyes between human subjects or suture joins in a skull. For a growing plant, potential landmarks include the tips of leaves and pedicel and branch angles. When specified *a priori*, landmarks should be assigned to provide adequate coverage of the shape morphology across a single dimensional plane (Bookstein, 1991). Additionally, the identification of landmark points should be repeatable and reliable across subjects while not altering their topological positions relative to other landmark positions (Bookstein, 1991). Type I landmarks provide the strongest support for homology because they are defined by underlying biological features, but it is problematic to assign Type I landmarks *a priori* when analyzing high-throughput plant imagery. To address this, PlantCV v2 contains functions to identify anatomical landmarks based upon the mathematical properties of object contours (Type II) and non-anatomical pseudo-landmarks/semilandmarks (Type III), as well as functions to rescale and analyze biologically relevant shape properties (Bookstein, 1991, 1997; Gunz, Mitteroecker & Bookstein, 2005; Gunz & Mitteroecker, 2013).

The ‘acute’ function identifies Type II landmarks by implementing a pseudo-landmark identification algorithm that operates using a modified form of chain coding (Freeman, 1961). Unlike standard chain coding methods that attempt to capture the absolute shape of a contour, the acute method operates by measuring the angle between a pixel coordinate and two neighboring pixels on opposite sides of it that fall within a set distance, or window, along the length of the contour. The two neighboring points are used to calculate an angle score for the center pixel. When the angle score is calculated for each position along the length of a contour, clusters of acute points can be identified, which can be segmented out by applying an angle threshold. The

middle position within each cluster of acute points is then identified for use as a pseudo-landmark (Fig. 4A). The ability to subjectively adjust the window size used for generating angle scores also helps to tailor analyses for identifying points of interest that may differ in resolution. For example, an analysis of leaf data might utilize a larger window size to identify the tips of lobes whereas smaller window sizes would be able to capture more minute patterns such as individual leaf serrations. Further segmentation can also be done using the average pixel values output (pt_vals) for each pseudo-landmark, which estimates the mean pixel intensity within the convex hull of each acute region based on the binary mask used in the analysis. The average pixel value output allows for concave landmarks (e.g. leaf axils and grass ligules) and convex landmarks (e.g. leaf tips and apices) on a contour to be differentiated in downstream analyses. Additionally, PlantCV v2 includes the 'acute_vertex' function that uses the same chain code-based pseudo-landmark identification algorithm used in the 'acute' function except that it uses an adjustable local search space criteria to reduce the number of angle calculations, which speeds up landmark identification.

For Type III landmarks, the 'x_axis_pseudolandmarks' and 'y_axis_pseudolandmarks' functions identify homologous points along a single dimension of an object (x-axis or y-axis) based on equidistant point locations within an object contour. The plant object is divided up into twenty equidistant bins, and the minimum and maximum extent of the object along the axis and the centroid of the object within each bin is calculated. These sixty points located along each axis possess the properties of semi/pseudo-landmark points (an equal number of reference points that are approximately geometrically homologous between subjects to be compared) that approximate the contour and shape of the object (Fig. 4B). Such semi/pseudo-landmarking strategies have

been utilized in cases where traditional homologous landmark points are difficult to assign or poorly represent the features of object shape (Bookstein, 1997; Gunz, Mitteroecker & Bookstein, 2005; Gunz & Mitteroecker, 2013).

Frequently, comparison of shape attributes requires rescaling of landmark points to eliminate the influence of size on the relative position of landmark points. The landmark functions in PlantCV output untransformed point values that can either be directly input into morphometric programs in R [shapes (Dryden & Mardia, 2016) or morpho (Schlager, Jefferis & Schlager, 2016)] or uniformly rescaled to a 0-1 coordinate system using the PlantCV 'scale_features' function. The location of landmark points can be used to examine multidimensional growth curves for a broad variety of study systems and tissue types and can be used to compare properties of plant shape throughout development or in response to differences in plant growth environment. An example of one such application is the 'landmark_reference_pt_dist' function. This function estimates the vertical, horizontal, Euclidean distance, and angle of landmark points from two landmarks (centroid of the plant object and centroid localized to the base of the plant). Preliminary evidence from a water limitation experiment performed using a *Setaria* recombinant inbred population indicates that vertical distance from rescaled leaf tip points identified by the 'acute_vertex' function to the centroid is decreased in response to water limitation and thus may provide a proximity measurement of plant turgor pressure (Fig. 4C and 4D).

Two-class or multiclass naive Bayes classifier

Pixel-level segmentation of images into two or more classes is not always straightforward using traditional image processing techniques. For example, two classes of features in an image may be

visually distinct but similar enough in color that simple thresholding is not sufficient to separate the two groups. Furthermore, even with methods that adjust for inconsistencies between images (e.g. white balancing and auto-thresholding functions), inconsistent lighting conditions in a growth chamber, greenhouse, or field can still make bulk processing of images with a single workflow difficult. Methods that utilize machine learning techniques are a promising approach to tackle these and other phenotyping challenges (Minervini, Abdelsamea & Tsafaris, 2014; Singh et al., 2016; Ubbens & Stavness, 2017; Atkinson et al., 2017; Pound et al., 2017). With PlantCV v2, we have started to integrate machine learning methods to detect features of interest (e.g. the plant), starting with a naive Bayes classifier (Abbasi & Fahlgren, 2016). The naive Bayes classifier can be trained using two different approaches for two-class or multiclass (two or more) segmentation problems. During the training phase using the ‘plantcv-train.py’ script, pixel RGB values for each input class are converted to the hue, saturation and value (HSV) color space. Kernel density estimation (KDE) is used to calculate a probability density function (PDF) from a vector of values for each HSV channel from each class. The output PDFs are used to parameterize the naive Bayes classifier function (‘naive_bayes_classifier’), which can be used to replace the thresholding steps in a PlantCV pipeline. The ‘naive_bayes_classifier’ function uses these PDFs to calculate the probability (using Bayes’ theorem) that a given pixel is in each class. The output of the ‘naive_bayes_classifier’ is a binary image for each class where the pixels are white if the probability the pixel was in the given class was highest of all classes and is black otherwise. A tutorial of how to implement naive Bayes plant detection into an image processing pipeline is online (http://plantcv.readthedocs.io/en/latest/machine_learning_tutorial/).

For the two-class approach, the training dataset includes color images and corresponding binary masks where the background is black and the foreground (plant or other target object) is white. PlantCV can be used to generate binary masks for the training set using the standard image processing methods and the new 'output_mask' function. It is important for the training dataset to be representative of the larger dataset. For example, if there are large fluctuations in light intensity throughout the day or plant color throughout the experiment, the training dataset should try to cover the range of variation. A random sample of 10% of the foreground pixels and the same number background pixels are used to build the PDFs.

To assess how well the two-class naive Bayes method identifies plant material in comparison to thresholding methods, we reanalyzed *Setaria* images (Fahlgren et al., 2015) using the naive Bayes classifier and compared the pixel area output to pipelines that utilize thresholding steps (Fig. 5). We used 99 training images (14 top view and 85 side view images) from a total of 6473 images. We found that the plant pixel area calculated by naive Bayes was highly correlated with that calculated from pipelines that use thresholding for both side-view images ($R^2=0.99$; Fig. 5A) and top-view images ($R^2=0.96$; Fig. 5B). Naive Bayes segmentation enabled use of pipelines that were both simpler (fewer steps) and more flexible: five new scripts were sufficient for processing the dataset (five categories of photo data), whereas nine threshold-based pipeline scripts had previously been required.

The multiclass naive Bayes approach requires a tab-delimited table for training where each column is a class (minimum two) and each cell is a comma-separated list of RGB pixel values from the column class. We currently use the Pixel Inspection Tool in ImageJ (Schneider,

Rasband & Eliceiri, 2012) to collect samples of pixel RGB values used to generate the training text file. As noted above for the two-class approach, it is important to adequately capture the variation in the image dataset for each class when generating the training text file to improve pixel classification. If images are consistent, only one image needs to be sampled for generating the training table; however, if they vary, several images may be needed. For complex backgrounds (or non-targeted objects), several classes may be required to capture all of the variation. Once the training table is generated, it is input into the ‘plantcv-train.py’ script to generate PDFs for each class. As an example, we used images of wheat leaves infected with wheat rust to collect pixel samples from four classes: non-plant background, unaffected leaf tissue, rust pustule, and chlorotic leaf tissue, and then used the naive Bayes classifier to segment the images into each class simultaneously (Fig. 6). This method can likely be used for a variety of applications, such as identifying a plant under variable lighting conditions or quantifying specific areas of stress on a plant.

In summary, the naive Bayes classifier offers several advantages over threshold-based segmentation: 1) two or more classes can be segmented simultaneously; 2) probabilistic segmentation can be more robust across images than fixed thresholds; and 3) classifier-based segmentation replaces multiple steps in threshold-based pipelines, reducing pipeline complexity.

CONCLUSIONS

The field of digital plant phenotyping is at an exciting stage of development where it is beginning to shift from a bottleneck to one that will have a positive impact on plant research, especially in agriculture. The Plant Image Analysis database currently lists over 150 tools that

can be used for plant phenotyping [<http://www.plant-image-analysis.org/>; (Lobet, Draye & Périlleux, 2013)]. Despite the abundance of software packages, long-term sustainability of individual projects may become an issue due to the lack of incentives for maintaining bioinformatics software developed in academia (Lobet, 2017). In a survey of corresponding authors of plant image analysis tools by Lobet, 60% either said the tool was no longer being maintained or did not respond (Lobet, 2017). To develop PlantCV as a sustainable project we have adopted an open, community-based development framework using GitHub as a central service for the organization of developer activities and the dissemination of information to users. We encourage contribution to the project by posting bug reports and issues, developing or revising analysis methods, adding or updating unit tests, writing documentation, and posting ideas for new features. We aim to periodically publish updates, such as the work presented here, to highlight the work of contributors to the PlantCV project.

There are several areas where we envision future PlantCV development. **Standards and interoperability:** Improved interoperability of PlantCV with data providers and downstream analysis tools will require adoption of community-based standards for data and metadata [e.g. Minimum Information About a Plant Phenotyping Experiment; (Ćwiek-Kupczyńska et al., 2016)]. Improved interoperability will make it easier to develop standardized tools for statistical analysis of image processing results, both within the PlantCV project or with tools from other projects. **New data sources:** Handling and analysis of data from specialized cameras that measure three-dimensional structure or hyperspectral reflectance will require development or integration of additional methods into PlantCV. **Machine learning:** Our goal is to develop additional tools for machine learning and collection of training data. In some cases, where these

methods can be implemented in a modular and reusable framework, they can be integrated directly into PlantCV. In other cases, PlantCV can be combined with new and existing tools. A recent example of this latter approach built on PlantCV, using its image preprocessing and segmentation functions alongside a modular framework for building convolutional neural networks (Ubbens & Stavness, 2017). As noted throughout, we see great potential for modular tools such as PlantCV and we welcome community feedback.

ACKNOWLEDGMENTS

We would like to thank Melinda Darnell, Leonardo Chavez, Kevin Reilly, and the staff of both the Danforth Center Facilities and Support Services group and the Plant Growth Facility for careful maintenance of the Danforth Center phenotyping facilities. We thank Katie Liberatore and Shahryar Kianian for images of wheat (*Triticum aestivum* L.). We would also like to thank all of the other people who have given us input on the PlantCV project in person or on GitHub.

REFERENCES

- Abbasi A., Fahlgren N. 2016. Naive Bayes pixel-level plant segmentation. In: *2016 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*. 1–4. DOI: 10.1109/WNYIPW.2016.7904790.
- Acosta-Gamboa LM., Liu S., Langley E., Campbell Z., Castro-Guerrero N., Mendoza-Cozatl D., Lorence A. 2017. Moderate to severe water limitation differentially affects the phenome and ionome of Arabidopsis. *Functional Plant Biology: FPB* 44:94–106. DOI: 10.1071/FP16172.
- Atkinson JA., Lobet G., Noll M., Meyer PE., Griffiths M., Wells DM. 2017. Combining

semi-automated image analysis techniques with machine learning algorithms to accelerate large scale genetic studies. *GigaScience*. DOI: 10.1093/gigascience/gix084.

Bookstein FL. 1991. *Morphometric tools for landmark data* Cambridge University Press. New York.

Bookstein FL. 1997. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.

Bradski G. 2000. The opencv library. *Doctor Dobbs Journal* 25:120–126.

Ćwiek-Kupczyńska H., Altmann T., Arend D., Arnaud E., Chen D., Cornut G., Fiorani F., Frohberg W., Junker A., Klukas C., Lange M., Mazurek C., Nafissi A., Neveu P., van Oeveren J., Pommier C., Poorter H., Rocca-Serra P., Sansone S-A., Scholz U., van Schriek M., Seren Ü., Usadel B., Weise S., Kersey P., Krajewski P. 2016. Measures for interoperability of phenotypic data: minimum information requirements and formatting. *Plant Methods* 12:44. DOI: 10.1186/s13007-016-0144-4.

Dryden IL., Mardia KV. 2016. *Statistical Shape Analysis: With Applications in R*. John Wiley & Sons.

Duarte M. 2015. Notes on Scientific Computing for Biomechanics and Motor Control. *GitHub repository*.

Dudley JT., Butte AJ. 2009. A quick guide for developing effective bioinformatics programming skills. *PLoS Computational Biology* 5:e1000589. DOI: 10.1371/journal.pcbi.1000589.

Fahlgren N., Feldman M., Gehan MA., Wilson MS., Shyu C., Bryant DW., Hill ST., McEntee CJ., Warnasooriya SN., Kumar I., Ficor T., Turnipseed S., Gilbert KB., Brutnell TP., Carrington JC., Mockler TC., Baxter I. 2015. A versatile phenotyping system and

analytics platform reveals diverse temporal responses to water availability in *Setaria*.

Molecular Plant 8:1520–1535. DOI: 10.1016/j.molp.2015.06.005.

Fahlgren N., Gehan MA., Baxter I. 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology* 24:93–99. DOI: 10.1016/j.pbi.2015.02.006.

Feldman MJ., Paul RE., Banan D., Barrett JF., Sebastian J., Yee M-C., Jiang H., Lipka AE., Brutnell TP., Dinneny JR., Leakey ADB., Baxter I. 2017. Time dependent genetic analysis links field and controlled environment phenotypes in the model C4 grass *Setaria*. *PLoS Genetics* 13:e1006841. DOI: 10.1371/journal.pgen.1006841.

Freeman H. 1961. On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronic Computers* EC-10:260–268. DOI: 10.1109/TEC.1961.5219197.

Furbank RT., Tester M. 2011. Phenomics--technologies to relieve the phenotyping bottleneck. *Trends in Plant Science* 16:635–644. DOI: 10.1016/j.tplants.2011.09.005.

Gonzalez RC., Woods RE. 2002. *Digital Image Processing*. Prentice Hall.

Gunz P., Mitteroecker P. 2013. Semilandmarks: a method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy* 24:103–109.

Gunz P., Mitteroecker P., Bookstein FL. 2005. Semilandmarks in Three Dimensions. In: Slice DE ed. *Modern Morphometrics in Physical Anthropology*. Developments in Primatology: Progress and Prospects. Springer US, 73–98. DOI: 10.1007/0-387-27614-9_3.

Hunter JD. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9:90–95. DOI: 10.1109/MCSE.2007.55.

Jones E., Oliphant T., Peterson P. 2014. *SciPy: Open source scientific tools for Python*.

Kaehler A., Bradski G. 2016. *Learning OpenCV 3: Computer Vision in C++ with the*

OpenCV Library. O'Reilly Media, Inc.

Kluyver T., Ragan-Kelley B., Pérez F., Granger B., Bussonnier M., Frederic J., Kelley K., Hamrick J., Grout J., Corlay S., Ivanov P., Avila D., Abdalla S., Willing C., Jupyter Development Team. 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Loizides F, Schmidt B eds. *Positioning and Power in Academic Publishing: Players, Agents and Agendas: Proceedings of the 20th International Conference on Electronic Publishing*. Amsterdam: IOS Press, 87–90. DOI: 10.3233/978-1-61499-649-1-87.

Liberti L., Lator C., Maculan N., Mucherino A. 2014. Euclidean distance geometry and applications. *SIAM Review* 56:3–69. DOI: 10.1137/120875909.

Lobet G. 2017. Image analysis in plant sciences: Publish then perish. *Trends in Plant Science*. DOI: 10.1016/j.tplants.2017.05.002.

Lobet G., Draye X., Périlleux C. 2013. An online database for plant image analysis software tools. *Plant Methods* 9:38. DOI: 10.1186/1746-4811-9-38.

Mangalam H. 2002. The Bio* toolkits--a brief overview. *Briefings in Bioinformatics* 3:296–302. DOI: 10.1093/bib/3.3.296.

McKinney W., Others. 2010. Data structures for statistical computing in python. In: *Proceedings of the 9th Python in Science Conference*. SciPy Austin, TX, 51–56.

Millman KJ., Aivazis M. 2011. Python for scientists and engineers. *Computing in Science & Engineering* 13:9–12. DOI: 10.1109/MCSE.2011.36.

Minervini M., Abdelsamea MM., Tsafaris SA. 2014. Image-based plant phenotyping with incremental learning and active contours. *Ecological Informatics* 23:35–48. DOI: 10.1016/j.ecoinf.2013.07.004.

Oliphant TE. 2007. Python for scientific computing. *Computing in Science & Engineering* 9:10–20. DOI: 10.1109/MCSE.2007.58.

Otsu N. 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9:62–66.

Pauli D., Chapman SC., Bart R., Topp CN., Lawrence-Dill CJ., Poland J., Gore MA. 2016. The quest for understanding phenotypic variation via integrated approaches in the field environment. *Plant Physiology* 172:622–634. DOI: 10.1104/pp.16.00592.

Perez-Riverol Y., Gatto L., Wang R., Sachsenberg T., Uszkoreit J., Leprevost F da V., Fufezan C., Ternent T., Eglen SJ., Katz DS., Pollard TJ., Konovalov A., Flight RM., Blin K., Vizcaíno JA. 2016. Ten simple rules for taking advantage of Git and GitHub. *PLoS Computational Biology* 12:e1004947. DOI: 10.1371/journal.pcbi.1004947.

Pound MP., Atkinson JA., Townsend AJ., Wilson MH., Griffiths M., Jackson AS., Bulat A., Tzimiropoulos G., Wells DM., Murchie EH., Pridmore TP., French AP. 2017. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *GigaScience*. DOI: 10.1093/gigascience/gix083.

R Core Team. 2017. R: A Language and Environment for Statistical Computing.

RStudio Team. 2016. *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.

Scharr H., Minervini M., French AP., Klukas C., Kramer DM., Liu X., Luengo I., Pape J-M., Polder G., Vukadinovic D., Yin X., Tsaftaris SA. 2016. Leaf segmentation in plant phenotyping: a collation study. *Machine Vision and Applications* 27:585–606. DOI: 10.1007/s00138-015-0737-3.

Schlager S., Jefferis G., Schlager MS. 2016. Package “Morpho.”

Schneider CA., Rasband WS., Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9:671–675.

Singh A., Ganapathysubramanian B., Singh AK., Sarkar S. 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science* 21:110–124. DOI: 10.1016/j.tplants.2015.10.015.

Tovar J., Hoyer JS., Lin A., Tielking A., Callen S., Castillo E., Miller M., Tessman M., Fahlgren N., Carrington J., Nusinow D., Gehan MA. 2017. Raspberry Pi powered imaging for plant phenotyping. *bioRxiv*:183822. DOI: 10.1101/183822.

Ubbens JR., Stavness I. 2017. Deep Plant Phenomics: A deep learning platform for complex plant phenotyping tasks. *Frontiers in Plant Science* 8:1190. DOI: 10.3389/fpls.2017.01190.

van der Walt S., Colbert SC., Varoquaux G. 2011. The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering* 13:22–30. DOI: 10.1109/MCSE.2011.37.

van der Walt S., Schönberger JL., Nunez-Iglesias J., Boulogne F., Warner JD., Yager N., Gouillart E., Yu T., scikit-image contributors. 2014. scikit-image: image processing in Python. *PeerJ* 2:e453. DOI: 10.7717/peerj.453.

Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer New York.

Wilson G., Aruliah DA., Brown CT., Chue Hong NP., Davis M., Guy RT., Haddock SHD., Huff KD., Mitchell IM., Plumbley MD., Waugh B., White EP., Wilson P. 2014. Best practices for scientific computing. *PLoS Biology* 12:e1001745. DOI: 10.1371/journal.pbio.1001745.

Zack GW., Rogers WE., Latp SA. 1977. Automatic measurement of sister chromatid exchange frequency. *Journal of Histochemistry and Cytochemistry* 25:741–753.

FIGURES

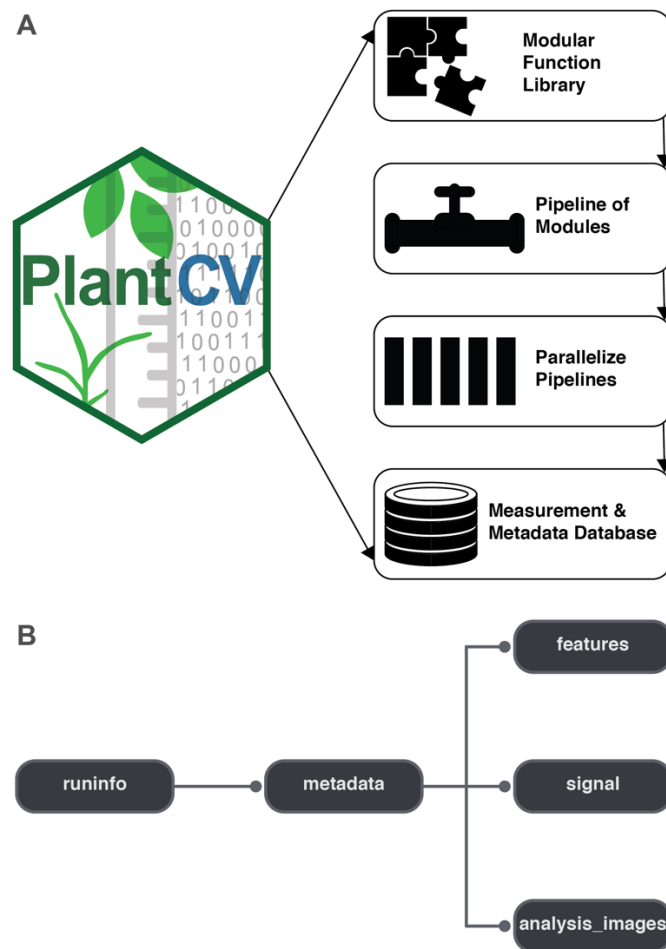


Figure 1: Diagram of the components of PlantCV.

A) PlantCV is an open-source, open-development suite of image analysis tools. PlantCV contains a library of modular Python functions that can be assembled into simple sequential or branching/merging processing pipelines. Image processing pipelines, which process single images (possibly containing multiple plants), can be deployed over large image sets using PlantCV parallelization, which outputs an SQLite database of both measurements and image/experimental metadata. B) Overview of the structure of the SQLite database.

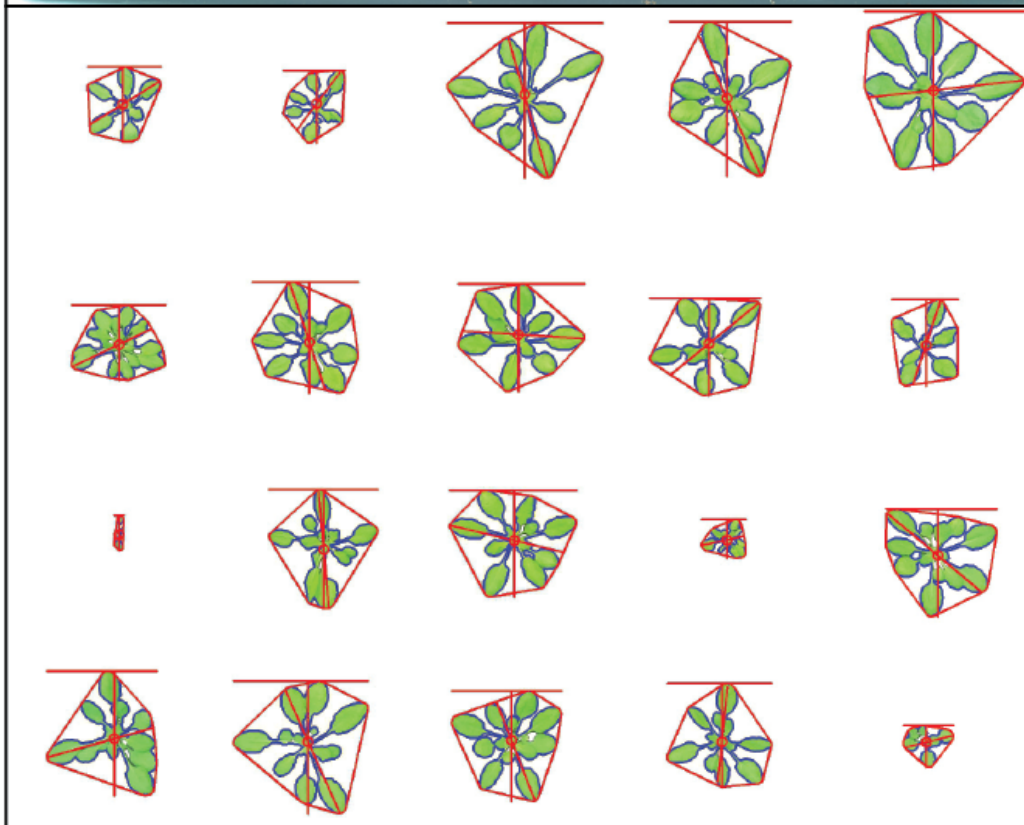


Figure 2: Analysis of images containing multiple plants.

New functions have been added to PlantCV v2 that enable individual plants from images containing multiple plants to be analyzed. The 'cluster_contours' function clusters contour objects using a flexible grid arrangement (approximate rows and columns defined by a user). The top image, produced by 'cluster_contours' in debug mode, highlights plants by their cluster group with unique colors on a sequential scale. The 'cluster_contours_split_img' function creates a new image for each cluster group. The resulting images of individual plants can be processed by standard PlantCV methods. In the bottom image, the 'cluster_contours_split_img' function was used to split the full image into individual plants. The shape of each plant was then analyzed with 'analyze_objects' and printed on a common image background.

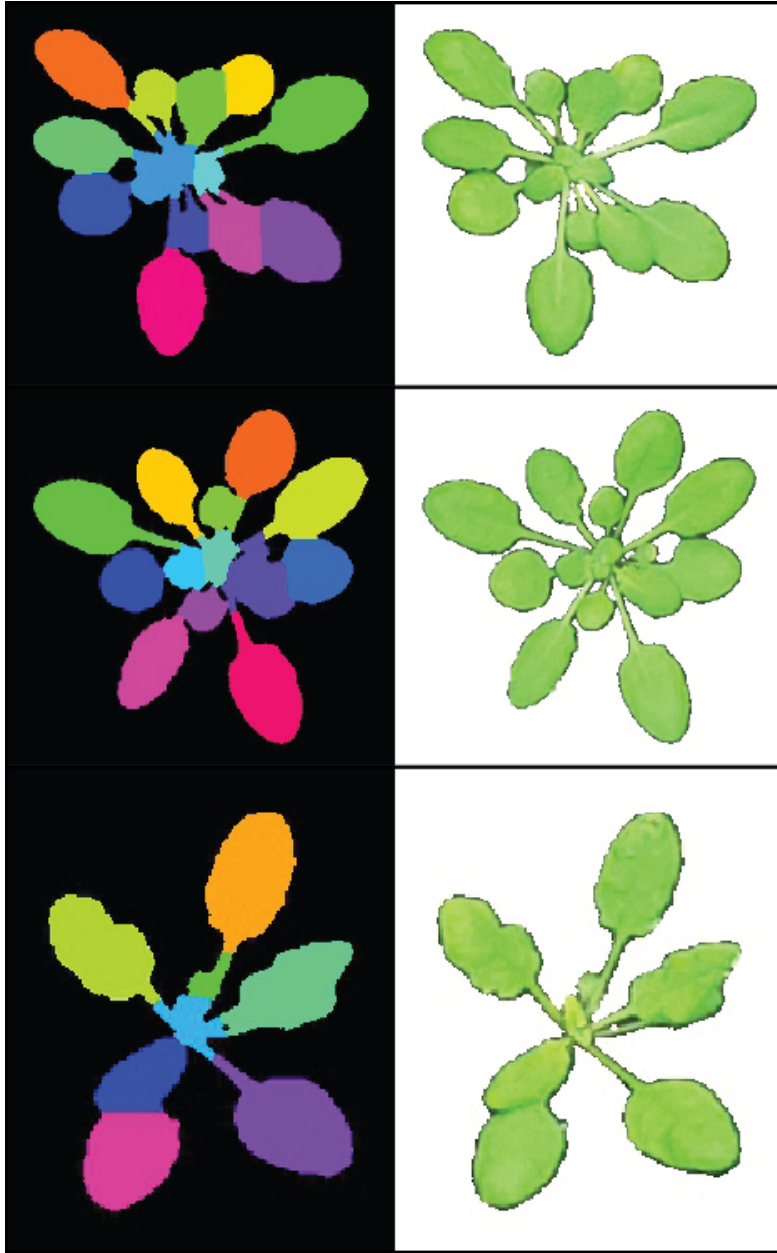


Figure 3: Leaf segmentation by a distance-based watershed transformation.

The watershed segmentation function can be used to segment and estimate the number of objects in an image. For the three example images, the watershed segmentation function was used to estimate the number of leaves for *Arabidopsis thaliana* (estimated leaf count for top: 13, middle: 14, and bottom: 8). Images shown are the output from the 'watershed_segmentation' function (left) and the segmented plants (right).

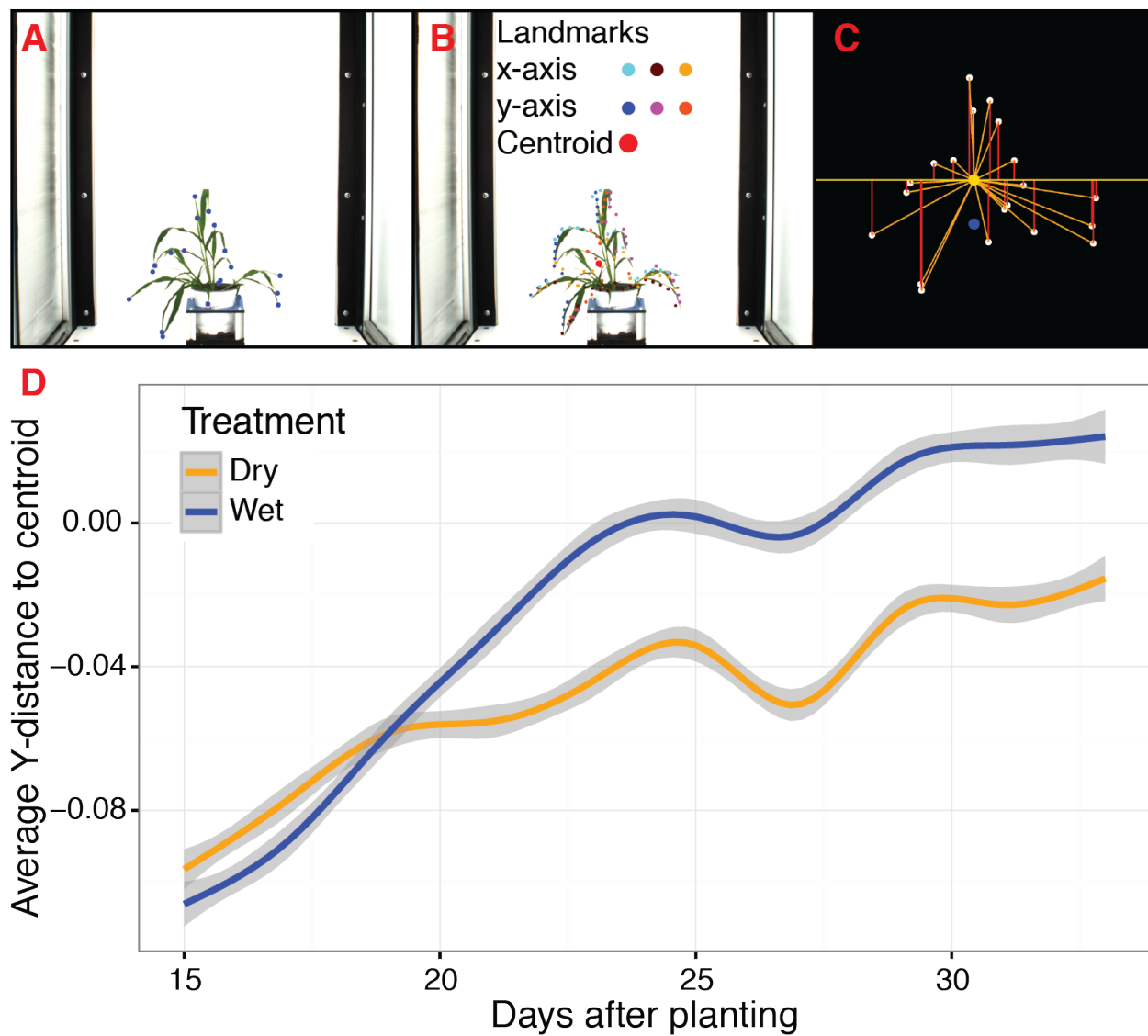


Figure 4: Landmark-based analysis of plant shape in PlantCV.

A) Automatic identification of leaf tip landmarks using the 'acute' and 'acute_vertex' functions (blue dots). B) Geometrically homologous semi/pseudo-landmarks across both the x- and y-axes. Semi/pseudo-landmarks identified by scanning the x-axis are denoted by light blue (top side of the contour), brown (bottom side of the contour), and light orange (centroid location of horizontal bins) dots. Semi/pseudo-landmarks identified by scanning the y-axis are denoted by dark blue (left side of the contour), pink (right side of the contour), and dark orange (centroid

location of vertical bins) dots. The plant centroid is plotted larger in red. C) A representation of the rescaled plant landmarks identified in panel (A). White points correspond to the leaf tips. The orange point is the location of the plant centroid. The blue point is the location of the plant centroid where the plant emerges from the soil. Red lines are the vertical distance from leaf tip points relative to the plant centroid. D) Analysis of the average scaled vertical distance from each leaf tip to the centroid diverges in response to water limitation.

Figure 5: Plant segmentation using a naive Bayes classifier.

Correlation between plant area (pixels) detected using thresholding pipelines (Fahlgren et al., 2015) on the x-axis compared to plant area detected using a trained naive Bayes classifier on the y-axis. A) Side-view images. B) Top-view images.

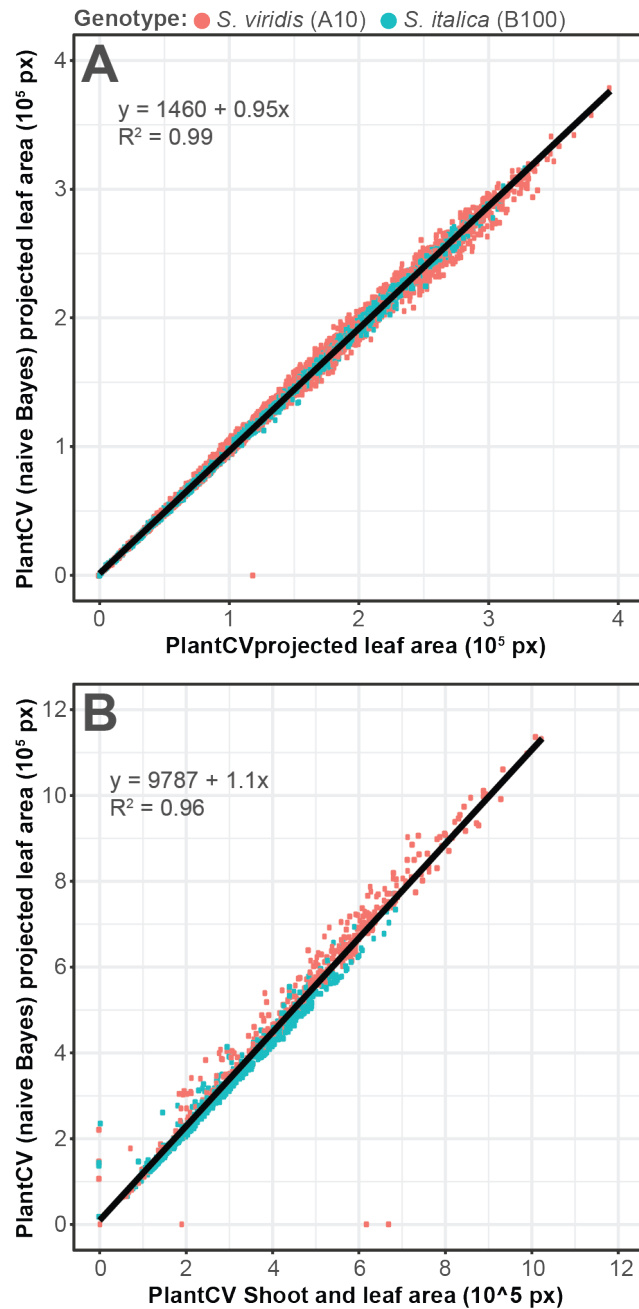
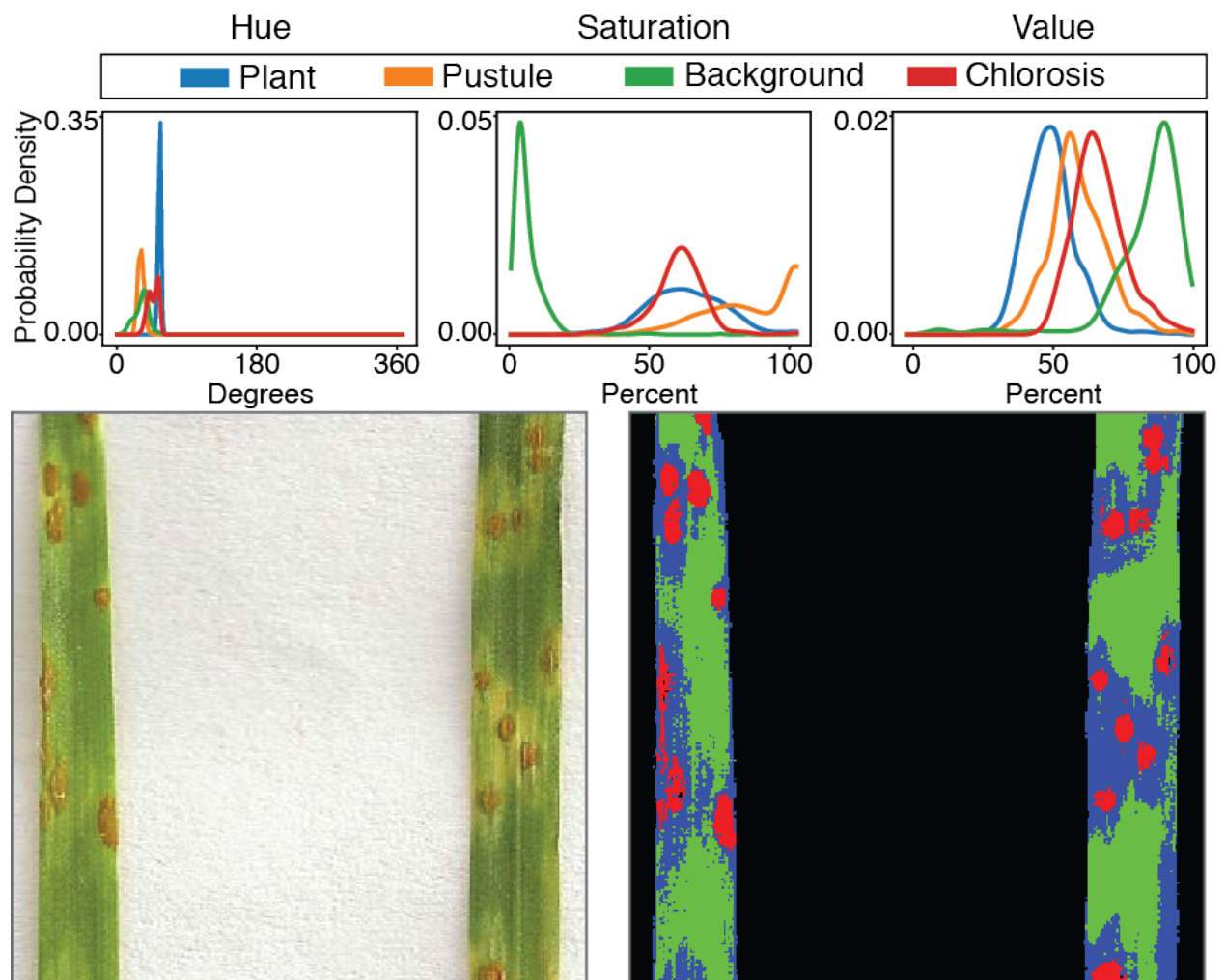


Figure 6: Simultaneous segmentation of four feature groups using the naive Bayes classifier.

An example of the naive Bayes classifier used to assign pixels into 4 classes: background, unaffected plant tissue, chlorotic tissue, and wheat stem rust pustules. (Top) Probability density functions (PDFs) from the ‘plantcv-train.py’ script that show hue, saturation, and value color channel distributions of four classes estimated from training data. (Bottom) Example of a classified image with the original image (left) and merged pseudocolored image (right) with pixels classified by the ‘naive_bayes_classifier’ as background (black), unaffected leaf tissue (green), chlorotic leaf tissue (blue), and pustules (red).



Bibliography

- (1) M. Levine and R. Tjian, “Transcription regulation and animal diversity”, *Nature*, July 2003, **424**, 147–51, DOI: 10.1038/nature01763.
- (2) S. B. Carroll, “Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution”, *Cell*, July 2008, **134**, 25–36, DOI: 10.1016/j.cell.2008.06.030.
- (3) P. J. Wittkopp and G. Kalay, “Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence”, *Nat Rev Genet*, Jan. 2012, **13**, 59–69, DOI: 10.1038/nrg3095.
- (4) I. Rubio-Somoza, J. T. Cuperus, D. Weigel and J. C. Carrington, “Regulation and functional specialization of small RNA-target nodes during plant development”, *Curr Opin Plant Biol*, Oct. 2009, **12**, 622–7, DOI: 10.1016/j.pbi.2009.07.003.
- (5) I. Rubio-Somoza and D. Weigel, “MicroRNA networks and developmental plasticity in plants”, *Trends Plant Sci*, May 2011, **16**, 258–64, DOI: 10.1016/j.tplants.2011.03.001.
- (6) J. P. Fouracre and R. S. Poethig, “The role of small RNAs in vegetative shoot development”, *Curr Opin Plant Biol*, Feb. 2016, **29**, 64–72, DOI: 10.1016/j.pbi.2015.11.006.
- (7) B. C. Meyers, M. J. Axtell, B. Bartel, D. P. Bartel, D. Baulcombe, J. L. Bowman, X. Cao, J. C. Carrington, X. Chen, P. J. Green, S. Griffiths-Jones, S. E. Jacobsen, A. C. Mallory, R. A. Martienssen, R. S. Poethig, Y. Qi, H. Vaucheret, O. Voinnet, Y. Watanabe, D. Weigel

- and J. K. Zhu, “Criteria for annotation of plant microRNAs”, *Plant Cell*, Dec. 2008, **20**, 3186–90, DOI: 10.1105/tpc.108.064311.
- (8) M. J. Axtell, “Classification and comparison of small RNAs from plants”, *Annu Rev Plant Biol*, 2013, **64**, 137–59, DOI: 10.1146/annurev-arplant-050312-120043.
 - (9) R. C. Lee, R. L. Feinbaum and V. Ambros, “The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*”, *Cell*, Dec. 1993, **75**, 843–54, DOI: 10.1016/0092-8674(93)90529-Y.
 - (10) B. Wightman, I. Ha and G. Ruvkun, “Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*”, *Cell*, Dec. 1993, **75**, 855–62, DOI: 10.1038/338313a0.
 - (11) A. J. Hamilton and D. C. Baulcombe, “A species of small antisense RNA in posttranscriptional gene silencing in plants”, *Science*, Oct. 1999, **286**, 950–2, DOI: 10.1126/science.286.5441.950.
 - (12) E. C. Lai, “microRNAs: Runtts of the genome assert themselves”, *Curr Biol*, Dec. 2003, **13**, R925–36, DOI: 10.1016/j.cub.2003.11.017.
 - (13) M. A. Matzke and A. J. Matzke, “Planting the seeds of a new paradigm”, *PLOS Biol*, May 2004, **2**, E133, DOI: 10.1371/journal.pbio.0020133.
 - (14) D. Baulcombe, “Of maize and men, or peas and people: case histories to justify plants and other model systems”, *Nat Med*, Oct. 2008, **14**, 1046–9, DOI: 10.1038/nm1008-1046.
 - (15) V. Ambros, “The evolution of our thinking about microRNAs”, *Nat Med*, Oct. 2008, **14**, 1036–40, DOI: 10.1038/nm1008-1036.

- (16) Z. Xie, L. K. Johansen, A. M. Gustafson, K. D. Kasschau, A. D. Lellis, D. Zilberman, S. E. Jacobsen and J. C. Carrington, “Genetic and functional diversification of small RNA pathways in plants”, *PLOS Biol*, May 2004, **2**, E104, DOI: 10.1371/journal.pbio.0020104.
- (17) E. Bernstein, A. A. Caudy, S. M. Hammond and G. J. Hannon, “Role for a bidentate ribonuclease in the initiation step of RNA interference”, *Nature*, Jan. 2001, **409**, 363–6, DOI: 10.1038/35053110.
- (18) M. J. Axtell, J. O. Westholm and E. C. Lai, “Vive la différence: Biogenesis and evolution of microRNAs in plants and animals”, *Genome Biol*, 2011, **12**, 221, DOI: 10.1186/gb-2011-12-4-221.
- (19) H. C. Lee, L. Li, W. Gu, Z. Xue, S. K. Crosthwaite, A. Pertsemlidis, Z. A. Lewis, M. Freitag, E. U. Selker, C. C. Mello and Y. Liu, “Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi”, *Mol Cell*, June 2010, **38**, 803–14, DOI: 10.1016/j.molcel.2010.04.005.
- (20) N. Fahlgren, S. R. Bollmann, K. D. Kasschau, J. T. Cuperus, C. M. Press, C. M. Sullivan, E. J. Chapman, J. S. Hoyer, K. B. Gilbert, N. J. Grünwald and J. C. Carrington, “*Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs”, *PLOS One*, 2013, **8**, e77181, DOI: 10.1371/journal.pone.0077181.
- (21) M. A. Matzke and R. A. Mosher, “RNA-directed DNA methylation: An epigenetic pathway of increasing complexity”, *Nat Rev Genet*, June 2014, **15**, 394–408, DOI: 10.1038/nrg3683.
- (22) Z. Zhang, X. Liu, X. Guo, X. J. Wang and X. Zhang, “*Arabidopsis* AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing”, *Nat Plants*, Apr. 2016, **2**, 16049, DOI: 10.1038/nplants.2016.49.

- (23) H. Zhang, R. Xia, B. C. Meyers and V. Walbot, “Evolution, functions, and mysteries of plant ARGONAUTE proteins”, *Curr Opin Plant Biol*, Oct. 2015, **27**, 84–90, DOI: 10.1016/j.pbi.2015.06.011.
- (24) J. A. Law and S. E. Jacobsen, “Establishing, maintaining and modifying DNA methylation patterns in plants and animals”, *Nat Rev Genet*, Mar. 2010, **11**, 204–20, DOI: 10.1038/nrg2719.
- (25) R. K. Slotkin, M. Vaughn, F. Borges, M. Tanurdzić, J. D. Becker, J. A. Feijó and R. A. Martienssen, “Epigenetic reprogramming and small RNA silencing of transposable elements in pollen”, *Cell*, Feb. 2009, **136**, 461–72, DOI: 10.1016/j.cell.2008.12.038.
- (26) V. Olmedo-Monfil, N. Durán-Figueroa, M. Arteaga-Vázquez, E. Demesa-Arévalo, D. Autran, D. Grimanelli, R. K. Slotkin, R. A. Martienssen and J. P. Vielle-Calzada, “Control of female gamete formation by a small RNA pathway in *Arabidopsis*”, *Nature*, Mar. 2010, **464**, 628–32, DOI: 10.1038/nature08828.
- (27) J. Zhai, H. Zhang, S. Arikiti, K. Huang, G. L. Nan, V. Walbot and B. C. Meyers, “Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers”, *Proc Natl Acad Sci U S A*, Mar. 2015, **112**, 3146–51, DOI: 10.1073/pnas.1418918112.
- (28) C. Napoli, C. Lemieux and R. Jorgensen, “Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes *in trans*”, *Plant Cell*, Apr. 1990, **2**, 279–289, DOI: 10.1105/tpc.2.4.279.
- (29) A. R. van der Krol, L. A. Mur, P. de Lange, J. N. Mol and A. R. Stuitje, “Inhibition of flower pigmentation by antisense CHS genes: Promoter and minimal sequence requirements for the antisense effect”, *Plant Mol Biol*, Apr. 1990, **14**, 457–66, DOI: 10.1007/BF00027492.

- (30) M. A. Matzke, M. Primig, J. Trnovsky and A. J. Matzke, “Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants”, *EMBO J*, Mar. 1989, **8**, 643–9.
- (31) A. Deleris, J. Gallego-Bartolome, J. Bao, K. D. Kasschau, J. C. Carrington and O. Voinnet, “Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense”, *Science*, July 2006, **313**, 68–71, DOI: 10.1126/science.1128214.
- (32) T. Csorba, L. Kontra and J. Burgyán, “Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence”, *Virology*, May 2015, **479-480**, 85–103, DOI: 10.1016/j.virol.2015.02.028.
- (33) A. Carbonell and J. C. Carrington, “Antiviral roles of plant ARGONAUTES”, *Curr Opin Plant Biol*, Oct. 2015, **27**, 111–7, DOI: 10.1016/j.pbi.2015.06.013.
- (34) K. D. Kasschau, Z. Xie, E. Allen, C. Llave, E. J. Chapman, K. A. Krizan and J. C. Carrington, “P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function”, *Dev Cell*, Feb. 2003, **4**, 205–17, DOI: 10.1016/S1534-5807(03)00025-X.
- (35) E. J. Chapman, A. I. Prokhnevsky, K. Gopinath, V. V. Dolja and J. C. Carrington, “Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step”, *Genes Dev*, May 2004, **18**, 1179–86, DOI: 10.1101/gad.1201204.
- (36) T. Yifhar, I. Pekker, D. Peled, G. Friedlander, A. Pistunov, M. Sabban, G. Wachsman, J. P. Alvarez, Z. Amsellem and Y. Eshed, “Failure of the tomato *trans*-acting short interfering RNA program to regulate AUXIN RESPONSE FACTOR3 and ARF4 underlies the wiry leaf syndrome”, *Plant Cell*, Sept. 2012, **24**, 3575–89, DOI: 10.1105/tpc.112.100222.

- (37) N. P. Achkar, D. A. Cambiagno and P. A. Manavella, “miRNA biogenesis: A dynamic pathway”, *Trends Plant Sci*, Dec. 2016, **21**, 1034–1044, DOI: 10.1016/j.tplants.2016.09.003.
- (38) K. Bohmert, I. Camus, C. Bellini, D. Bouchez, M. Caboche and C. Benning, “*AGO1* defines a novel locus of *Arabidopsis* controlling leaf development”, *EMBO J*, Jan. 1998, **17**, 170–80, DOI: 10.1093/emboj/17.1.170.
- (39) S. E. Schauer, S. E. Jacobsen, D. W. Meinke and A. Ray, “*DICER-LIKE1*: Blind men and elephants in *Arabidopsis* development”, *Trends Plant Sci*, Nov. 2002, **7**, 487–91, DOI: 10.1016/S1360-1385(02)02355-5.
- (40) E. R. Havecker, L. M. Wallbridge, T. J. Hardcastle, M. S. Bush, K. A. Kelly, R. M. Dunn, F. Schwach, J. H. Doonan and D. C. Baulcombe, “The *Arabidopsis* RNA-directed DNA methylation Argonautes functionally diverge based on their expression and interaction with target loci”, *Plant Cell*, Feb. 2010, **22**, 321–34, DOI: 10.1105/tpc.109.072199.
- (41) X. Fang and Y. Qi, “RNAi in plants: An Argonaute-centered view”, *Plant Cell*, Feb. 2016, **28**, 272–85, DOI: 10.1105/tpc.15.00920.
- (42) X. Zhang, H. Zhao, S. Gao, W. C. Wang, S. Katiyar-Agarwal, H. D. Huang, N. Raikhel and H. Jin, “*Arabidopsis* Argonaute 2 regulates innate immunity via miRNA393*-mediated silencing of a Golgi-localized SNARE gene, *MEMB12*”, *Mol Cell*, May 2011, **42**, 356–66, DOI: 10.1016/j.molcel.2011.04.010.
- (43) W. Wei, Z. Ba, M. Gao, Y. Wu, Y. Ma, S. Amiard, C. I. White, J. M. Rendtlew Danielsen, Y. G. Yang and Y. Qi, “A role for small RNAs in DNA double-strand break repair”, *Cell*, Mar. 2012, **149**, 101–12, DOI: 10.1016/j.cell.2012.03.002.
- (44) M. Liu, Z. Ba, P. Costa-Nunes, W. Wei, L. Li, F. Kong, Y. Li, J. Chai, O. Pontes and Y. Qi, “IDN2 interacts with RPA and facilitates DNA double-strand break repair by homologous

- recombination in *Arabidopsis*”, *Plant Cell*, Mar. 2017, **29**, 589–599, DOI: 10.1105/tpc.16.00769.
- (45) A. Peragine, M. Yoshikawa, G. Wu, H. L. Albrecht and R. S. Poethig, “SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*”, *Genes Dev*, Oct. 2004, **18**, 2368–79, DOI: 10.1101/gad.1231804.
- (46) F. Vazquez, H. Vaucheret, R. Rajagopalan, C. Lepers, V. Gascioli, A. C. Mallory, J. L. Hilbert, D. P. Bartel and P. Cr  t  , “Endogenous *trans*-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs”, *Mol Cell*, Oct. 2004, **16**, 69–79, DOI: 10.1016/j.molcel.2004.09.028.
- (47) Q. Fei, R. Xia and B. C. Meyers, “Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks”, *Plant Cell*, July 2013, **25**, 2400–15, DOI: 10.1105/tpc.113.114652.
- (48) E. Allen, Z. Xie, A. M. Gustafson and J. C. Carrington, “microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants”, *Cell*, Apr. 2005, **121**, 207–21, DOI: 10.1016/j.cell.2005.04.004.
- (49) M. J. Axtell, C. Jan, R. Rajagopalan and D. P. Bartel, “A two-hit trigger for siRNA biogenesis in plants”, *Cell*, Nov. 2006, **127**, 565–77, DOI: 10.1016/j.cell.2006.09.032.
- (50) T. A. Montgomery, M. D. Howell, J. T. Cuperus, D. Li, J. E. Hansen, A. L. Alexander, E. J. Chapman, N. Fahlgren, E. Allen and J. C. Carrington, “Specificity of ARGONAUTE7-miR390 interaction and dual functionality in *TAS3 trans*-acting siRNA formation”, *Cell*, Apr. 2008, **133**, 128–41, DOI: 10.1016/j.cell.2008.02.033.
- (51) T. Elmayan, S. Balzergue, F. B  on, V. Bourdon, J. Daubremet, Y. Gu  net, P. Mourrain, J. C. Palauqui, S. Vernhettes, T. Vialle, K. Wostrikoff and H. Vaucheret, “*Arabidopsis* mutants

- impaired in cosuppression”, *Plant Cell*, Oct. 1998, **10**, 1747–58, DOI: 10.1105/tpc.10.10.1747.
- (52) P. Mourrain, C. Béclin, T. Elmayan, F. Feuerbach, C. Godon, J. B. Morel, D. Jouette, A. M. Lacombe, S. Nikic, N. Picault, K. Ré moué, M. Sanial, T. A. Vo and H. Vaucheret, “*Arabidopsis* *SGS2* and *SGS3* genes are required for posttranscriptional gene silencing and natural virus resistance”, *Cell*, May 2000, **101**, 533–42, DOI: 10.1016/S0092-8674(00)80863-6.
- (53) V. Jouannet and A. Maizel, “*Trans*-acting small interfering RNAs: Biogenesis, mode of action, and role in plant development”, in *MicroRNAs in plant development and stress responses*, ed. R. Sunkar, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 83–108, DOI: 10.1007/978-3-642-27384-1_5.
- (54) J. T. Cuperus, N. Fahlgren and J. C. Carrington, “Evolution and functional diversification of *MIRNA* genes”, *Plant Cell*, Feb. 2011, **23**, 431–42, DOI: 10.1105/tpc.110.082784.
- (55) R. A. Chávez Montes, F. de Fátima Rosas-Cárdenas, E. De Paoli, M. Accerbi, L. A. Rymarquis, G. Mahalingam, N. Marsch-Martínez, B. C. Meyers, P. J. Green and S. de Folter, “Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs”, *Nat Commun*, Apr. 2014, **5**, 3722, DOI: 10.1038/ncomms4722.
- (56) F. Berruezo, F. de Souza, P. I. Picca, S. I. Nemirovsky, L. Martínez Tosar, M. Rivero, A. N. Mentaberry and A. M. Zelada, “Sequencing of small RNAs of the fern *Pleopeltis minima* (Polypodiaceae) offers insight into the evolution of the microRNA repertoire in land plants”, *PLOS One*, 2017, **12**, e0177573, DOI: 10.1371/journal.pone.0177573.
- (57) C. A. Matthewman, C. G. Kawashima, D. Húska, T. Csorba, T. Dalmay and S. Kopriva, “miR395 is a general component of the sulfate assimilation regulatory network in *Arabidopsis*”, *FEBS Lett*, Sept. 2012, **586**, 3242–8, DOI: 10.1016/j.febslet.2012.06.044.

- (58) C. Ma, S. Burd and A. Lers, “miR408 is involved in abiotic stress responses in *Arabidopsis*”, *Plant J*, Oct. 2015, **84**, 169–87, DOI: 10.1111/tpj.12999.
- (59) B. J. Reinhart, E. G. Weinstein, M. W. Rhoades, B. Bartel and D. P. Bartel, “MicroRNAs in plants”, *Genes Dev*, July 2002, **16**, 1616–26, DOI: 10.1101/gad.1004402.
- (60) C. Llave, Z. Xie, K. D. Kasschau and J. C. Carrington, “Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA”, *Science*, Sept. 2002, **297**, 2053–6, DOI: 10.1126/science.1076311.
- (61) C. Llave, K. D. Kasschau, M. A. Rector and J. C. Carrington, “Endogenous and silencing-associated small RNAs in plants”, *Plant Cell*, July 2002, **14**, 1605–19, DOI: 10.1105/tpc.003210.
- (62) M. W. Rhoades, B. J. Reinhart, L. P. Lim, C. B. Burge, B. Bartel and D. P. Bartel, “Prediction of plant microRNA targets”, *Cell*, Aug. 2002, **110**, 513–20.
- (63) M. J. Axtell and D. P. Bartel, “Antiquity of microRNAs and their targets in land plants”, *Plant Cell*, June 2005, **17**, 1658–73, DOI: 10.1105/tpc.105.032185.
- (64) T. Arazi, M. Talmor-Neiman, R. Stav, M. Riese, P. Huijser and D. C. Baulcombe, “Cloning and characterization of micro-RNAs from moss”, *Plant J*, Sept. 2005, **43**, 837–48, DOI: 10.1111/j.1365-313X.2005.02499.x.
- (65) M. J. Axtell, J. A. Snyder and D. P. Bartel, “Common functions for diverse small RNAs of land plants”, *Plant Cell*, June 2007, **19**, 1750–69, DOI: 10.1105/tpc.107.051706.
- (66) M. W. Jones-Rhoades and D. P. Bartel, “Computational identification of plant microRNAs and their targets, including a stress-induced miRNA”, *Mol Cell*, June 2004, **14**, 787–99, DOI: 10.1016/j.molcel.2004.05.027.

- (67) L. Navarro, P. Dunoyer, F. Jay, B. Arnold, N. Dharmasiri, M. Estelle, O. Voinnet and J. D. Jones, “A plant miRNA contributes to antibacterial resistance by repressing auxin signaling”, *Science*, Apr. 2006, **312**, 436–9, DOI: 10.1126/science.1126088.
- (68) J. B. Song, S. Q. Huang, T. Dalmay and Z. M. Yang, “Regulation of leaf morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS”, *Plant Cell Physiol*, July 2012, **53**, 1283–94, DOI: 10.1093/pcp/pcs080.
- (69) S. Knauer, A. L. Holt, I. Rubio-Somoza, E. J. Tucker, A. Hinze, M. Pisch, M. Javelle, M. C. Timmermans, M. R. Tucker and T. Laux, “A protodermal miR394 signal defines a region of stem cell competence in the *Arabidopsis* shoot meristem”, *Dev Cell*, Jan. 2013, **24**, 125–32, DOI: 10.1016/j.devcel.2012.12.009.
- (70) Z. Xie, K. D. Kasschau and J. C. Carrington, “Negative feedback regulation of DICER-LIKE1 in *Arabidopsis* by microRNA-guided mRNA degradation”, *Curr Biol*, Apr. 2003, **13**, 784–9, DOI: 10.1016/S0960-9822(03)00281-1.
- (71) F. Vazquez, V. Gascioli, P. Crété and H. Vaucheret, “The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing”, *Curr Biol*, Feb. 2004, **14**, 346–51, DOI: 10.1016/j.cub.2004.01.035.
- (72) H. Vaucheret, F. Vazquez, P. Crété and D. P. Bartel, “The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development”, *Genes Dev*, May 2004, **18**, 1187–97, DOI: 10.1101/gad.1201404.
- (73) H. Vaucheret, A. C. Mallory and D. P. Bartel, “AGO1 homeostasis entails coexpression of *MIR168* and *AGO1* and preferential stabilization of miR168 by AGO1”, *Mol Cell*, Apr. 2006, **22**, 129–36, DOI: 10.1016/j.molcel.2006.03.011.

- (74) R. Rajagopalan, H. Vaucheret, J. Trejo and D. P. Bartel, “A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*”, *Genes Dev*, Dec. 2006, **20**, 3407–25, DOI: 10.1101/gad.1476406.
- (75) N. Fahlgren, M. D. Howell, K. D. Kasschau, E. J. Chapman, C. M. Sullivan, J. S. Cumbie, S. A. Givan, T. F. Law, S. R. Grant, J. L. Dangel and J. C. Carrington, “High-throughput sequencing of *Arabidopsis* microRNAs: Evidence for frequent birth and death of *MIRNA* genes”, *PLOS One*, Feb. 2007, **2**, e219, DOI: 10.1371/journal.pone.0000219.
- (76) N. Fahlgren, S. Jogdeo, K. D. Kasschau, C. M. Sullivan, E. J. Chapman, S. Laubinger, L. M. Smith, M. Dasenko, S. A. Givan, D. Weigel and J. C. Carrington, “MicroRNA gene evolution in *Arabidopsis lyrata* and *Arabidopsis thaliana*”, *Plant Cell*, Apr. 2010, **22**, 1074–89, DOI: 10.1105/tpc.110.073999.
- (77) Z. Ma, C. Coruh and M. J. Axtell, “*Arabidopsis lyrata* small RNAs: Transient *MIRNA* and small interfering RNA loci within the *Arabidopsis* genus”, *Plant Cell*, Apr. 2010, **22**, 1090–103, DOI: 10.1105/tpc.110.073882.
- (78) L. M. Smith, H. A. Burbano, X. Wang, J. Fitz, G. Wang, Y. Ural-Blimke and D. Weigel, “Rapid divergence and high diversity of miRNAs and miRNA targets in the *Camelineae*”, *Plant J*, Feb. 2015, **81**, 597–610, DOI: 10.1111/tpj.12754.
- (79) E. Allen, Z. Xie, A. M. Gustafson, G. H. Sung, J. W. Spatafora and J. C. Carrington, “Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*”, *Nat Genet*, Dec. 2004, **36**, 1282–90, DOI: 10.1038/ng1478.
- (80) F. F. Felippes, K. Schneeberger, T. DeZulian, D. H. Huson and D. Weigel, “Evolution of *Arabidopsis thaliana* microRNAs from random sequences”, *RNA*, Dec. 2008, **14**, 2455–9, DOI: 10.1261/rna.1149408.

- (81) M. J. Axtell, “Evolution of microRNAs and their targets: Are all microRNAs biologically relevant?”, *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, Nov. 2008, **1779**, 725–734, DOI: 10.1016/j.bbagr.2008.02.007.
- (82) K. Chen and N. Rajewsky, “The evolution of gene regulation by transcription factors and microRNAs”, *Nat Rev Genet*, Feb. 2007, **8**, 93–103, DOI: 10.1038/nrg1990.
- (83) E. Berezikov, “Evolution of microRNA diversity and regulation in animals”, *Nat Rev Genet*, Nov. 2011, **12**, 846–60, DOI: 10.1038/nrg3079.
- (84) A. E. Pasquinelli, B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, E. E. Ball, B. Degan, P. Müller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson and G. Ruvkun, “Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA”, *Nature*, Nov. 2000, **408**, 86–9, DOI: 10.1038/35040556.
- (85) B. J. Reinhart, F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvié, H. R. Horvitz and G. Ruvkun, “The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*”, *Nature*, Feb. 2000, **403**, 901–6, DOI: 10.1038/35002607.
- (86) R. C. Friedman, K. K. Farh, C. B. Burge and D. P. Bartel, “Most mammalian mRNAs are conserved targets of microRNAs”, *Genome Res*, Jan. 2009, **19**, 92–105, DOI: 10.1101/gr.082701.108.
- (87) E. M. Meyerowitz, “Plants and the logic of development”, *Genetics*, Jan. 1997, **145**, 5–9.
- (88) R. S. Poethig, “Vegetative phase change and shoot maturation in plants”, *Curr Top Dev Biol*, 2013, **105**, 125–52, DOI: 10.1016/B978-0-12-396968-2.00005-1.
- (89) D. H. Chitwood and N. R. Sinha, “Evolutionary and environmental forces sculpting leaf development”, *Curr Biol*, Apr. 2016, **26**, R297–306, DOI: 10.1016/j.cub.2016.02.033.

- (90) C. S. Jones, “An essay on juvenility, phase change, and heteroblasty in seed plants”, *International Journal of Plant Sciences*, Nov. 1999, **160**, S105–S111, DOI: 10.1086/314215.
- (91) G. Zotz, K. Wilhelm and A. Becker, “Heteroblasty—A review”, *Bot. Rev.*, June 2011, **77**, 109–151, DOI: 10.1007/s12229-010-9062-8.
- (92) C. Kuhlemeier and M. C. Timmermans, “The Sussex signal: Insights into leaf dorsiventral-ity”, *Development*, Sept. 2016, **143**, 3230–7, DOI: 10.1242/dev.131888.
- (93) A. Y. Husbands, D. H. Chitwood, Y. Plavskin and M. C. Timmermans, “Signals and prepatterns: New insights into organ polarity in plants”, *Genes Dev*, Sept. 2009, **23**, 1986–97, DOI: 10.1101/gad.1819909.
- (94) A. Izhaki and J. L. Bowman, “KANADI and class III HD-Zip gene families regulate embryo patterning and modulate auxin flow during embryogenesis in *Arabidopsis*”, *Plant Cell*, Feb. 2007, **19**, 495–508, DOI: 10.1105/tpc.106.047472.
- (95) J. W. Chandler, M. Cole, A. Flier, B. Grewe and W. Werr, “The AP2 transcription factors DORNROSCHE and DORNROSCHE-LIKE redundantly control *Arabidopsis* embryo patterning via interaction with PHAVOLUTA”, *Development*, May 2007, **134**, 1653–62, DOI: 10.1242/dev.001016.
- (96) Z. R. Smith and J. A. Long, “Control of *Arabidopsis* apical-basal embryo polarity by antagonistic transcription factors”, *Nature*, Mar. 2010, **464**, 423–6, DOI: 10.1038/nature08843.
- (97) J. R. McConnell and M. K. Barton, “Leaf polarity and meristem formation in *Arabidopsis*”, *Development*, Aug. 1998, **125**, 2935–42.
- (98) J. R. McConnell, J. Emery, Y. Eshed, N. Bao, J. Bowman and M. K. Barton, “Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots”, *Nature*, June 2001, **411**, 709–13, DOI: 10.1038/35079635.

- (99) A. C. Mallory, B. J. Reinhart, M. W. Jones-Rhoades, G. Tang, P. D. Zamore, M. K. Barton and D. P. Bartel, “MicroRNA control of *PHABULOSA* in leaf development: Importance of pairing to the microRNA 5' region”, *EMBO J*, Aug. 2004, **23**, 3356–64, DOI: 10.1038/sj.emboj.7600340.
- (100) H. Zhu, F. Hu, R. Wang, X. Zhou, S. H. Sze, L. W. Liou, A. Barefoot, M. Dickman and X. Zhang, “*Arabidopsis* ARGONAUTE10 specifically sequesters miR166/165 to regulate shoot apical meristem development”, *Cell*, Apr. 2011, **145**, 242–56, DOI: 10.1016/j.cell.2011.03.024.
- (101) Y. Yu, L. Ji, B. H. Le, J. Zhai, J. Chen, E. Luscher, L. Gao, C. Liu, X. Cao, B. Mo, J. Ma, B. C. Meyers and X. Chen, “ARGONAUTE10 promotes the degradation of miR165/6 through the SDN1 and SDN2 exonucleases in *Arabidopsis*”, *PLOS Biol*, Feb. 2017, **15**, e2001272, DOI: 10.1371/journal.pbio.2001272.
- (102) R. Brandt, Y. Xie, T. Musielak, M. Graeff, Y. D. Stierhof, H. Huang, C. M. Liu and S. Wenkel, “Control of stem cell homeostasis via interlocking microRNA and microProtein feedback loops”, *Mech Dev*, Jan. 2013, **130**, 25–33, DOI: 10.1016/j.mod.2012.06.007.
- (103) K. Lynn, A. Fernandez, M. Aida, J. Sedbrook, M. Tasaka, P. Masson and M. K. Barton, “The *PINHEAD/ZWILLE* gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the *ARGONAUTE1* gene”, *Development*, Feb. 1999, **126**, 469–81.
- (104) M. R. Tucker, A. Hinze, E. J. Tucker, S. Takada, G. Jürgens and T. Laux, “Vascular signalling mediated by *ZWILLE* potentiates *WUSCHEL* function during shoot meristem stem cell development in the *Arabidopsis* embryo”, *Development*, Sept. 2008, **135**, 2839–43, DOI: 10.1242/dev.023648.

- (105) A. Carlsbecker, J. Y. Lee, C. J. Roberts, J. Dettmer, S. Lehesranta, J. Zhou, O. Lindgren, M. A. Moreno-Risueno, A. Vatén, S. Thitamadee, A. Campilho, J. Sebastian, J. L. Bowman, Y. Helariutta and P. N. Benfey, “Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate”, *Nature*, May 2010, **465**, 316–21, DOI: 10.1038/nature08977.
- (106) J. Wu, Z. Yang, Y. Wang, L. Zheng, R. Ye, Y. Ji, S. Zhao, S. Ji, R. Liu, L. Xu, H. Zheng, Y. Zhou, X. Zhang, X. Cao, L. Xie, Z. Wu, Y. Qi and Y. Li, “Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA”, *eLife*, Feb. 2015, **4**, DOI: 10.7554/eLife.05733.
- (107) J. Wu, R. Yang, Z. Yang, S. Yao, S. Zhao, Y. Wang, P. Li, X. Song, L. Jin, T. Zhou, Y. Lan, L. Xie, X. Zhou, C. Chu, Y. Qi, X. Cao and Y. Li, “ROS accumulation and antiviral defence control by microRNA528 in rice”, *Nat Plants*, Jan. 2017, **3**, 16203, DOI: 10.1038/nplants.2016.203.
- (108) L. Zhai, W. Sun, K. Zhang, H. Jia, L. Liu, Z. Liu, F. Teng and Z. Zhang, “Identification and characterization of Argonaute gene family and meiosis-enriched Argonaute during sporogenesis in maize”, *J Integr Plant Biol*, Nov. 2014, **56**, 1042–52, DOI: 10.1111/jipb.12205.
- (109) Q. Fei, L. Yang, W. Liang, D. Zhang and B. C. Meyers, “Dynamic changes of small RNAs in rice spikelet development reveal specialized reproductive phaseRNA pathways”, *J Exp Bot*, Nov. 2016, **67**, 6037–6049, DOI: 10.1093/jxb/erw361.
- (110) D. Reinhardt, M. Frenz, T. Mandel and C. Kuhlemeier, “Microsurgical and laser ablation analysis of leaf positioning and dorsoventral patterning in tomato”, *Development*, Jan. 2005, **132**, 15–26, DOI: 10.1242/dev.01544.
- (111) I. M. Sussex, “Experiments on the cause of dorsiventrality in leaves”, *Nature*, Apr. 1951, **167**, 651–2.

- (112) F. T. Nogueira, S. Madi, D. H. Chitwood, M. T. Juarez and M. C. Timmermans, “Two small regulatory RNAs establish opposing fates of a developmental axis”, *Genes Dev*, Apr. 2007, **21**, 750–5, DOI: 10.1101/gad.1528607.
- (113) C. Hunter, H. Sun and R. S. Poethig, “The *Arabidopsis* heterochronic gene *ZIPPY* is an *ARGONAUTE* family member”, *Curr Biol*, Sept. 2003, **13**, 1734–9, DOI: 10.1016/j.cub.2003.09.004.
- (114) M. Yoshikawa, A. Peragine, M. Y. Park and R. S. Poethig, “A pathway for the biogenesis of *trans*-acting siRNAs in *Arabidopsis*”, *Genes Dev*, Sept. 2005, **19**, 2164–75, DOI: 10.1101/gad.1352605.
- (115) J. T. Cuperus, T. A. Montgomery, N. Fahlgren, R. T. Burke, T. Townsend, C. M. Sullivan and J. C. Carrington, “Identification of *MIR390a* precursor processing-defective mutants in *Arabidopsis* by direct genome sequencing”, *Proc Natl Acad Sci U S A*, Jan. 2010, **107**, 466–71, DOI: 10.1073/pnas.0913203107.
- (116) I. Pekker, J. P. Alvarez and Y. Eshed, “Auxin response factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity”, *Plant Cell*, Nov. 2005, **17**, 2899–910, DOI: 10.1105/tpc.105.034876.
- (117) D. H. Chitwood, F. T. Nogueira, M. D. Howell, T. A. Montgomery, J. C. Carrington and M. C. Timmermans, “Pattern formation via small RNA mobility”, *Genes Dev*, Mar. 2009, **23**, 549–54, DOI: 10.1101/gad.1770009.
- (118) R. Schwab, A. Maizel, V. Ruiz-Ferrer, D. Garcia, M. Bayer, M. Crespi, O. Voinnet and R. A. Martienssen, “Endogenous tasiRNAs mediate non-cell autonomous effects on gene regulation in *Arabidopsis thaliana*”, *PLOS One*, June 2009, **4**, e5980, DOI: 10.1371/journal.pone.0005980.

- (119) S. Simonini, J. Deb, L. Moubayidin, P. Stephenson, M. Valluru, A. Freire-Rios, K. Sorefan, D. Weijers, J. Friml and L. Østergaard, “A noncanonical auxin-sensing mechanism is required for organ morphogenesis in *Arabidopsis*”, *Genes Dev*, Oct. 2016, **30**, 2286–2296, DOI: 10.1101/gad.285361.116.
- (120) S. Simonini, S. Bencivenga, M. Trick and L. Østergaard, “Auxin-induced modulation of ETTIN activity orchestrates gene expression in *Arabidopsis*”, *Plant Cell*, Aug. 2017, DOI: 10.1105/tpc.17.00389.
- (121) R. A. Sessions and P. C. Zambryski, “*Arabidopsis* gynoecium structure in the wild and in *ettin* mutants”, *Development*, May 1995, **121**, 1519–32.
- (122) A. Sessions, J. L. Nemhauser, A. McColl, J. L. Roe, K. A. Feldmann and P. C. Zambryski, “*ETTIN* patterns the *Arabidopsis* floral meristem and reproductive organs”, *Development*, Nov. 1997, **124**, 4481–91.
- (123) Z. Su, L. Zhao, Y. Zhao, S. Li, S. Won, H. Cai, L. Wang, Z. Li, P. Chen, Y. Qin and X. Chen, “The THO complex non-cell-autonomously represses female germline specification through the *TAS3*-ARF3 module”, *Curr Biol*, June 2017, **27**, 1597–1609.e2, DOI: 10.1016/j.cub.2017.05.021.
- (124) E. Marin, V. Jouannet, A. Herz, A. S. Lokerse, D. Weijers, H. Vaucheret, L. Nussaume, M. D. Crespi and A. Maizel, “miR390, *Arabidopsis* *TAS3* tasiRNAs, and their *AUXIN RESPONSE FACTOR* targets define an autoregulatory network quantitatively regulating lateral root growth”, *Plant Cell*, Apr. 2010, **22**, 1104–17, DOI: 10.1105/tpc.109.072553.
- (125) E. K. Yoon, J. H. Yang, J. Lim, S. H. Kim, S. K. Kim and W. S. Lee, “Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in *Arabidopsis* lateral root development”, *Nucleic Acids Res*, Mar. 2010, **38**, 1382–91, DOI: 10.1093/nar/gkp1128.

- (126) Y. Plavskin, A. Nagashima, P. F. Perroud, M. Hasebe, R. S. Quatrano, G. S. Atwal and M. C. Timmermans, “Ancient *trans*-acting siRNAs confer robustness and sensitivity onto the auxin response”, *Dev Cell*, Feb. 2016, **36**, 276–89, DOI: 10.1016/j.devcel.2016.01.010.
- (127) R. Xia, J. Xu and B. C. Meyers, “The emergence, evolution, and diversification of the miR390-TAS3-ARF pathway in land plants”, *Plant Cell*, Apr. 2017, DOI: 10.1105/tpc.17.00185.
- (128) C. Hunter, M. R. Willmann, G. Wu, M. Yoshikawa, M. de la Luz Gutiérrez-Nava and S. R. Poethig, “*Trans*-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in *Arabidopsis*”, *Development*, Aug. 2006, **133**, 2973–81, DOI: 10.1242/dev.02491.
- (129) J. Klein, H. Saedler and P. Huijser, “A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*”, *Mol Gen Genet*, Jan. 1996, **250**, 7–16.
- (130) G. Wu and R. S. Poethig, “Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3”, *Development*, Sept. 2006, **133**, 3539–47, DOI: 10.1242/dev.02521.
- (131) L. Yang, S. R. Conway and R. S. Poethig, “Vegetative phase change is mediated by a leaf-derived signal that represses the transcription of miR156”, *Development*, Jan. 2011, **138**, 245–9, DOI: 10.1242/dev.058578.
- (132) L. Yang, M. Xu, Y. Koo, J. He and R. S. Poethig, “Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of *MIR156A* and *MIR156C*”, *eLife*, Mar. 2013, **2**, e00260, DOI: 10.7554/eLife.00260.

- (133) S. Yu, L. Cao, C. M. Zhou, T. Q. Zhang, H. Lian, Y. Sun, J. Wu, J. Huang, G. Wang and J. W. Wang, “Sugar is an endogenous cue for juvenile-to-adult phase transition in plants”, *eLife*, Mar. 2013, **2**, e00269, DOI: 10.7554/eLife.00269.
- (134) A. Telfer, K. M. Bollman and R. S. Poethig, “Phase change and the regulation of trichome distribution in *Arabidopsis thaliana*”, *Development*, Feb. 1997, **124**, 645–54.
- (135) S. H. Cho, C. Addo-Quaye, C. Coruh, M. A. Arif, Z. Ma, W. Frank and M. J. Axtell, “*Physcomitrella patens* DCL3 is required for 22-24 nt siRNA accumulation, suppression of retrotransposon-derived transcripts, and normal development”, *PLOS Genet*, Dec. 2008, **4**, e1000314, DOI: 10.1371/journal.pgen.1000314.
- (136) S. H. Cho, C. Coruh and M. J. Axtell, “miR156 and miR390 regulate tasiRNA accumulation and developmental timing in *Physcomitrella patens*”, *Plant Cell*, Dec. 2012, **24**, 4837–49, DOI: 10.1105/tpc.112.103176.
- (137) G. Wu, M. Y. Park, S. R. Conway, J. W. Wang, D. Weigel and R. S. Poethig, “The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*”, *Cell*, Aug. 2009, **138**, 750–9, DOI: 10.1016/j.cell.2009.06.031.
- (138) P. Huijser and M. Schmid, “The control of developmental phase transitions in plants”, *Development*, Oct. 2011, **138**, 4117–29, DOI: 10.1242/dev.063511.
- (139) S. P. Moose and P. H. Sisco, “*Glossy15* controls the epidermal juvenile-to-adult phase transition in maize”, *Plant Cell*, Oct. 1994, **6**, 1343–1355, DOI: 10.1105/tpc.6.10.1343.
- (140) S. P. Moose and P. H. Sisco, “*Glossy15*, an *APETALA2*-like gene from maize that regulates leaf epidermal cell identity”, *Genes Dev*, Dec. 1996, **10**, 3018–27, DOI: 10.1101/gad.10.23.3018.

- (141) N. Lauter, A. Kampani, S. Carlson, M. Goebel and S. P. Moose, “microRNA172 down-regulates *glossy15* to promote vegetative phase change in maize”, *Proc Natl Acad Sci U S A*, June 2005, **102**, 9412–7, DOI: 10.1073/pnas.0503927102.
- (142) J. W. Wang, B. Czech and D. Weigel, “miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*”, *Cell*, Aug. 2009, **138**, 738–49, DOI: 10.1016/j.cell.2009.06.014.
- (143) A. Yamaguchi, M. F. Wu, L. Yang, G. Wu, R. S. Poethig and D. Wagner, “The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of *LEAFY*, *FRUITFULL*, and *APETALA1*”, *Dev Cell*, Aug. 2009, **17**, 268–78, DOI: 10.1016/j.devcel.2009.06.007.
- (144) M. Shikata, T. Koyama, N. Mitsuda and M. Ohme-Takagi, “*Arabidopsis* SBP-box genes SPL10, SPL11 and SPL2 control morphological change in association with shoot maturation in the reproductive phase”, *Plant Cell Physiol*, Dec. 2009, **50**, 2133–45, DOI: 10.1093/pcp/pcp148.
- (145) S. Yu, V. C. Galvão, Y. C. Zhang, D. Horrer, T. Q. Zhang, Y. H. Hao, Y. Q. Feng, S. Wang, M. Schmid and J. W. Wang, “Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA PROMOTER BINDING-LIKE transcription factors”, *Plant Cell*, Aug. 2012, **24**, 3320–32, DOI: 10.1105/tpc.112.101014.
- (146) N. Yu, Q. W. Niu, K. H. Ng and N. H. Chua, “The role of miR156/SPLs modules in *Arabidopsis* lateral root development”, *Plant J*, Aug. 2015, **83**, 673–85, DOI: 10.1111/tpj.12919.

- (147) M. Xu, T. Hu, J. Zhao, M. Y. Park, K. W. Earley, G. Wu, L. Yang and R. S. Poethig, “Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes in *Arabidopsis thaliana*”, *PLOS Genet*, Aug. 2016, **12**, e1006263, DOI: 10.1371/journal.pgen.1006263.
- (148) R. Schwab, J. F. Palatnik, M. Riester, C. Schommer, M. Schmid and D. Weigel, “Specific effects of microRNAs on the plant transcriptome”, *Dev Cell*, Apr. 2005, **8**, 517–27, DOI: 10.1016/j.devcel.2005.01.018.
- (149) J. W. Wang, R. Schwab, B. Czech, E. Mica and D. Weigel, “Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*”, *Plant Cell*, May 2008, **20**, 1231–43, DOI: 10.1105/tpc.108.058180.
- (150) G. Chuck, A. M. Cigan, K. Saetern and S. Hake, “The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA”, *Nat Genet*, Apr. 2007, **39**, 544–9, DOI: 10.1038/ng2001.
- (151) K. H. Klempnauer, T. J. Gonda and J. M. Bishop, “Nucleotide sequence of the retroviral leukemia gene *v-myb* and its cellular progenitor *c-myb*: The architecture of a transduced oncogene”, *Cell*, Dec. 1982, **31**, 453–63, DOI: 10.1016/0092-8674(82)90138-6.
- (152) S. Ambawat, P. Sharma, N. R. Yadav and R. C. Yadav, “MYB transcription factor genes as regulators for plant responses: An overview”, *Physiol Mol Biol Plants*, July 2013, **19**, 307–21, DOI: 10.1007/s12298-013-0179-1.
- (153) X. Y. Xue, B. Zhao, L. M. Chao, D. Y. Chen, W. R. Cui, Y. B. Mao, L. J. Wang and X. Y. Chen, “Interaction between two timing microRNAs controls trichome distribution in *Arabidopsis*”, *PLOS Genet*, Apr. 2014, **10**, e1004266, DOI: 10.1371/journal.pgen.1004266.
- (154) E. Ioannidi, S. Rigas, D. Tsitsekian, G. Daras, A. Alatzas, A. Makris, G. Tanou, A. Argiriou, D. Alexandrou, S. Poethig, P. Hatzopoulos and A. K. Kanellis, “Trichome patterning control

- involves TTG1 interaction with SPL transcription factors”, *Plant Mol Biol*, Dec. 2016, **92**, 675–687, DOI: 10.1007/s11103-016-0538-8.
- (155) I. Rubio-Somoza, C. M. Zhou, A. Confraria, C. Martinho, P. von Born, E. Baena-Gonzalez, J. W. Wang and D. Weigel, “Temporal control of leaf complexity by miRNA-regulated licensing of protein complexes”, *Curr Biol*, Nov. 2014, **24**, 2714–9, DOI: 10.1016/j.cub.2014.09.058.
- (156) J. Doebley, A. Stec and L. Hubbard, “The evolution of apical dominance in maize”, *Nature*, Apr. 1997, **386**, 485–8, DOI: 10.1038/386485a0.
- (157) D. Luo, R. Carpenter, C. Vincent, L. Copsey and E. Coen, “Origin of floral asymmetry in *Antirrhinum*”, *Nature*, Oct. 1996, **383**, 794–9, DOI: 10.1038/383794a0.
- (158) S. Kosugi and Y. Ohashi, “PCF1 and PCF2 specifically bind to *cis* elements in the rice proliferating cell nuclear antigen gene”, *Plant Cell*, Sept. 1997, **9**, 1607–19, DOI: 10.1105/tpc.9.9.1607.
- (159) P. Cubas, N. Lauter, J. Doebley and E. Coen, “The TCP domain: A motif found in proteins regulating plant growth and development”, *Plant J*, Apr. 1999, **18**, 215–22, DOI: 10.1046/j.1365-3113X.1999.00444.x.
- (160) U. Nath, B. C. Crawford, R. Carpenter and E. Coen, “Genetic control of surface curvature”, *Science*, Feb. 2003, **299**, 1404–7, DOI: 10.1126/science.1079354.
- (161) M. Martín-Trillo and P. Cubas, “TCP genes: A family snapshot ten years later”, *Trends Plant Sci*, Jan. 2010, **15**, 31–9, DOI: 10.1016/j.tplants.2009.11.003.
- (162) M. Nicolas and P. Cubas, “TCP factors: New kids on the signaling block”, *Curr Opin Plant Biol*, Oct. 2016, **33**, 33–41, DOI: 10.1016/j.pbi.2016.05.006.

- (163) J. F. Palatnik, E. Allen, X. Wu, C. Schommer, R. Schwab, J. C. Carrington and D. Weigel, “Control of leaf morphogenesis by microRNAs”, *Nature*, Sept. 2003, **425**, 257–63, DOI: 10.1038/nature01958.
- (164) R. Sarojam, P. G. Sappl, A. Goldshmidt, I. Efroni, S. K. Floyd, Y. Eshed and J. L. Bowman, “Differentiating *Arabidopsis* shoots from leaves by combined YABBY activities”, *Plant Cell*, July 2010, **22**, 2113–30, DOI: 10.1105/tpc.110.075853.
- (165) J. P. Alvarez, C. Furumizu, I. Efroni, Y. Eshed and J. L. Bowman, “Active suppression of a leaf meristem orchestrates determinate leaf growth”, *eLife*, Oct. 2016, **5**, DOI: 10.7554/eLife.15023.
- (166) C. Schommer, J. F. Palatnik, P. Aggarwal, A. Chételat, P. Cubas, E. E. Farmer, U. Nath and D. Weigel, “Control of jasmonate biosynthesis and senescence by miR319 targets”, *PLOS Biol*, Sept. 2008, **6**, e230, DOI: 10.1371/journal.pbio.0060230.
- (167) T. Koyama, M. Furutani, M. Tasaka and M. Ohme-Takagi, “TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in *Arabidopsis*”, *Plant Cell*, Feb. 2007, **19**, 473–84, DOI: 10.1105/tpc.106.044792.
- (168) T. Koyama, N. Mitsuda, M. Seki, K. Shinozaki and M. Ohme-Takagi, “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*, Nov. 2010, **22**, 3574–88, DOI: 10.1105/tpc.110.075598.
- (169) C. Guo, Y. Xu, M. Shi, Y. Lai, X. Wu, H. Wang, Z. Zhu, R. S. Poethig and G. Wu, “Repression of miR156 by miR159 regulates the timing of the juvenile-to-adult transition in *Arabidopsis*”, *Plant Cell*, May 2017, DOI: 10.1105/tpc.16.00975.

- (170) I. Rubio-Somoza and D. Weigel, “Coordination of flower maturation by a regulatory circuit of three microRNAs”, *PLOS Genet*, Mar. 2013, **9**, e1003374, DOI: 10.1371/journal.pgen.1003374.
- (171) C. Schommer, J. M. Debernardi, E. G. Bresso, R. E. Rodriguez and J. F. Palatnik, “Repression of cell proliferation by miR319-regulated TCP4”, *Mol Plant*, Oct. 2014, **7**, 1533–44, DOI: 10.1093/mp/ssu084.
- (172) M. A. Mecchia, J. M. Debernardi, R. E. Rodriguez, C. Schommer and J. F. Palatnik, “MicroRNA miR396 and RDR6 synergistically regulate leaf development”, *Mech Dev*, Jan. 2013, **130**, 2–13, DOI: 10.1016/j.mod.2012.07.005.
- (173) H. Li, L. Xu, H. Wang, Z. Yuan, X. Cao, Z. Yang, D. Zhang, Y. Xu and H. Huang, “The putative RNA-dependent RNA polymerase RDR6 acts synergistically with *ASYMMETRIC LEAVES1* and 2 to repress *BREVIPEDICELLUS* and *MICRORNA165/166* in *Arabidopsis* leaf development”, *Plant Cell*, Aug. 2005, **17**, 2157–71, DOI: 10.1105/tpc.105.033449.
- (174) L. Xu, L. Yang, L. Pi, Q. Liu, Q. Ling, H. Wang, R. S. Poethig and H. Huang, “Genetic interaction between the *AS1-AS2* and *RDR6-SGS3-AGO7* pathways for leaf morphogenesis”, *Plant Cell Physiol*, July 2006, **47**, 853–63, DOI: 10.1093/pcp/pcj057.
- (175) D. Garcia, S. A. Collier, M. E. Byrne and R. A. Martienssen, “Specification of leaf polarity in *Arabidopsis* via the *trans*-acting siRNA pathway”, *Curr Biol*, May 2006, **16**, 933–8, DOI: 10.1016/j.cub.2006.03.064.
- (176) M. Iwasaki, H. Takahashi, H. Iwakawa, A. Nakagawa, T. Ishikawa, H. Tanaka, Y. Matsumura, I. Pekker, Y. Eshed, S. Vial-Pradel, T. Ito, Y. Watanabe, Y. Ueno, H. Fukazawa, S. Kojima, Y. Machida and C. Machida, “Dual regulation of *ETTIN (ARF3)* gene expression by *AS1-AS2*, which maintains the DNA methylation level, is involved in stabilization of

- leaf adaxial-abaxial partitioning in *Arabidopsis*”, *Development*, May 2013, **140**, 1958–69, DOI: 10.1242/dev.085365.
- (177) A. Y. Husbands, A. H. Benkovics, F. T. Nogueira, M. Lodha and M. C. Timmermans, “The ASYMMETRIC LEAVES complex employs multiple modes of regulation to affect adaxial-abaxial patterning and leaf complexity”, *Plant Cell*, Dec. 2015, **27**, 3321–35, DOI: 10.1105/tpc.15.00454.
- (178) J. L. Pruneda-Paz, G. Breton, D. H. Nagel, S. E. Kang, K. Bonaldi, C. J. Doherty, S. Ravelo, M. Galli, J. R. Ecker and S. A. Kay, “A genome-scale resource for the functional characterization of *Arabidopsis* transcription factors”, *Cell Rep*, July 2014, **8**, 622–32, DOI: 10.1016/j.celrep.2014.06.033.
- (179) J. L. Pruneda-Paz, G. Breton, A. Para and S. A. Kay, “A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock”, *Science*, Mar. 2009, **323**, 1481–5, DOI: 10.1126/science.1167206.
- (180) K. Century, T. L. Reuber and O. J. Ratcliffe, “Regulating the regulators: The future prospects for transcription-factor-based agricultural biotechnology products”, *Plant Physiol*, May 2008, **147**, 20–9, DOI: 10.1104/pp.108.117887.
- (181) J. F. Doebley, B. S. Gaut and B. D. Smith, “The molecular genetics of crop domestication”, *Cell*, Dec. 2006, **127**, 1309–21, DOI: 10.1016/j.cell.2006.12.006.
- (182) G. Swinnen, A. Goossens and L. Pauwels, “Lessons from domestication: Targeting *cis*-regulatory elements for crop improvement”, *Trends Plant Sci*, June 2016, **21**, 506–15, DOI: 10.1016/j.tplants.2016.01.014.
- (183) A. Carbonell, “Artificial small RNA-based strategies for effective and specific gene silencing in plants”, in *Plant gene silencing: Mechanisms and applications*, ed. T. Dalmay, Google-Books-ID: EscmDwAAQBAJ, CABI, May 2017, vol. 5, p. 110.

- (184) G. S. Chuck, C. Tobias, L. Sun, F. Kraemer, C. Li, D. Dibble, R. Arora, J. N. Bragg, J. P. Vogel, S. Singh, B. A. Simmons, M. Pauly and S. Hake, “Overexpression of the maize *Corngrass1* microRNA prevents flowering, improves digestibility, and increases starch content of switchgrass”, *Proc Natl Acad Sci U S A*, Oct. 2011, **108**, 17550–5, DOI: 10.1073/pnas.1113971108.
- (185) S. B. Preuss, R. Meister, Q. Xu, C. P. Urwin, F. A. Tripodi, S. E. Screen, V. S. Anil, S. Zhu, J. A. Morrell, G. Liu, O. J. Ratcliffe, T. L. Reuber, R. Khanna, B. S. Goldman, E. Bell, T. E. Ziegler, A. L. McClerren, T. G. Ruff and M. E. Petracek, “Expression of the *Arabidopsis thaliana* *BBX32* gene in soybean increases grain yield”, *PLOS One*, 2012, **7**, e30717, DOI: 10.1371/journal.pone.0030717.
- (186) M. L. Campos, Y. Yoshida, I. T. Major, D. de Oliveira Ferreira, S. M. Weraduwege, J. E. Froehlich, B. F. Johnson, D. M. Kramer, G. Jander, T. D. Sharkey and G. A. Howe, “Rewiring of jasmonate and phytochrome B signalling uncouples plant growth-defense tradeoffs”, *Nat Commun*, Aug. 2016, **7**, 12570, DOI: 10.1038/ncomms12570.
- (187) Z. Shi, J. Wang, X. Wan, G. Shen, X. Wang and J. Zhang, “Over-expression of rice *OsAGO7* gene induces upward curling of the leaf blade that enhanced erect-leaf habit”, *Planta*, June 2007, **226**, 99–108, DOI: 10.1007/s00425-006-0472-0.
- (188) D. Lin, Y. Xiang, Z. Xian and Z. Li, “Ectopic expression of *SLAGO7* alters leaf pattern and inflorescence architecture and increases fruit yield in tomato”, *Physiol Plant*, Aug. 2016, **157**, 490–506, DOI: 10.1111/pp1.12425.
- (189) A. M. Brunner and O. Nilsson, “Revisiting tree maturation and floral initiation in the poplar functional genomics era”, *New Phytologist*, Oct. 2004, **164**, 43–51, DOI: 10.1111/j.1469-8137.2004.01165.x.

- (190) J. W. Wang, M. Y. Park, L. J. Wang, Y. Koo, X. Y. Chen, D. Weigel and R. S. Poethig, “miRNA control of vegetative phase change in trees”, *PLOS Genet*, Feb. 2011, **7**, e1002012, DOI: 10.1371/journal.pgen.1002012.
- (191) C. Fu, R. Sunkar, C. Zhou, H. Shen, J. Y. Zhang, J. Matts, J. Wolf, D. G. Mann, C. N. Stewart, Jr, Y. Tang and Z. Y. Wang, “Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production”, *Plant Biotechnol J*, May 2012, **10**, 443–52, DOI: 10.1111/j.1467-7652.2011.00677.x.
- (192) H. Wang and H. Wang, “The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits”, *Mol Plant*, May 2015, **8**, 677–88, DOI: 10.1016/j.molp.2015.01.008.
- (193) J. Tang and C. Chu, “MicroRNAs in crop improvement: Fine-tuners for complex traits”, *Nat Plants*, June 2017, **3**, 17077, DOI: 10.1038/nplants.2017.77.
- (194) H. Garcia-Ruiz, A. Carbonell, J. S. Hoyer, N. Fahlgren, K. B. Gilbert, A. Takeda, A. Giampetruzzi, M. T. Garcia Ruiz, M. G. McGinn, N. Lowery, M. T. Martinez Baladejo and J. C. Carrington, “Roles and programming of *Arabidopsis* ARGONAUTE proteins during *Turnip mosaic virus* infection”, *PLOS Pathog*, Mar. 2015, **11**, e1004755, DOI: 10.1371/journal.ppat.1004755.
- (195) B. Deplancke, D. Dupuy, M. Vidal and A. J. Walhout, “A Gateway-compatible yeast one-hybrid system”, *Genome Res*, Oct. 2004, **14**, 2093–101, DOI: 10.1101/gr.2445504.
- (196) M. T. Weirauch, A. Yang, M. Albu, A. G. Cote, A. Montenegro-Montero, P. Drewe, H. S. Najafabadi, S. A. Lambert, I. Mann, K. Cook, H. Zheng, A. Goity, H. van Bakel, J. C. Lozano, M. Galli, M. G. Lewsey, E. Huang, T. Mukherjee, X. Chen, J. S. Reece-Hoyes, S. Govindarajan, G. Shaulsky, A. J. Walhout, F. Y. Bouget, G. Ratsch, L. F. Larrondo, J. R.

- Ecker and T. R. Hughes, “Determination and inference of eukaryotic transcription factor sequence specificity”, *Cell*, Sept. 2014, **158**, 1431–43, DOI: 10.1016/j.cell.2014.08.009.
- (197) J. M. Franco-Zorrilla, I. López-Vidriero, J. L. Carrasco, M. Godoy, P. Vera and R. Solano, “DNA-binding specificities of plant transcription factors and their potential to define target genes”, *Proc Natl Acad Sci U S A*, Feb. 2014, **111**, 2367–72, DOI: 10.1073/pnas.1316278111.
- (198) R. C. O’Malley, S. C. Huang, L. Song, M. G. Lewsey, A. Bartlett, J. R. Nery, M. Galli, A. Gallavotti and J. R. Ecker, “Cistrome and epicistrome features shape the regulatory DNA landscape”, *Cell*, Sept. 2016, **166**, 1598, DOI: 10.1016/j.cell.2016.08.063.
- (199) G. D. Stormo and Y. Zhao, “Determining the specificity of protein-DNA interactions”, *Nat Rev Genet*, Nov. 2010, **11**, 751–60, DOI: 10.1038/nrg2845.
- (200) A. M. Sullivan, A. A. Arsovski, J. Lempe, K. L. Bubbs, M. T. Weirauch, P. J. Sabo, R. Sandstrom, R. E. Thurman, S. Neph, A. P. Reynolds, A. B. Stergachis, B. Vernot, A. K. Johnson, E. Haugen, S. T. Sullivan, A. Thompson, F. V. Neri, 3rd, M. Weaver, M. Diegel, S. Mnaimneh, A. Yang, T. R. Hughes, J. L. Nemhauser, C. Queitsch and J. A. Stamatoyannopoulos, “Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*”, *Cell Rep*, Sept. 2014, **8**, 2015–30, DOI: 10.1016/j.celrep.2014.08.019.
- (201) N. Fahlgren, M. A. Gehan and I. Baxter, “Lights, camera, action: High-throughput plant phenotyping is ready for a close-up”, *Curr Opin Plant Biol*, Apr. 2015, **24**, 93–9, DOI: 10.1016/j.pbi.2015.02.006.

- (202) J. Bell and H. M. Dee, “Watching plants grow—a position paper on computer vision and *Arabidopsis thaliana*”, *IET Computer Vision*, Oct. 2016, DOI: 10.1049/iet-cvi.2016.0127.
- (203) E. Allen and M. D. Howell, “miRNAs in the biogenesis of *trans*-acting siRNAs in higher plants”, *Semin Cell Dev Biol*, Oct. 2010, **21**, 798–804, DOI: 10.1016/j.semcdb.2010.03.008.
- (204) L. Williams, C. C. Carles, K. S. Osmond and J. C. Fletcher, “A database analysis method identifies an endogenous *trans*-acting short-interfering RNA that targets the *Arabidopsis* *ARF2*, *ARF3*, and *ARF4* genes”, *Proc Natl Acad Sci U S A*, July 2005, **102**, 9703–8, DOI: 10.1073/pnas.0504029102.
- (205) N. Fahlgren, T. A. Montgomery, M. D. Howell, E. Allen, S. K. Dvorak, A. L. Alexander and J. C. Carrington, “Regulation of *AUXIN RESPONSE FACTOR3* by *TAS3* ta-siRNA affects developmental timing and patterning in *Arabidopsis*”, *Curr Biol*, May 2006, **16**, 939–44, DOI: 10.1016/j.cub.2006.03.065.
- (206) X. Adenot, T. Elmayan, D. Lauressergues, S. Boutet, N. Bouché, V. Gasciolli and H. Vaucheret, “DRB4-dependent *TAS3* *trans*-acting siRNAs control leaf morphology through *AGO7*”, *Curr Biol*, May 2006, **16**, 927–32, DOI: 10.1016/j.cub.2006.03.035.
- (207) D. S. Skopelitis, A. Y. Husbands and M. C. Timmermans, “Plant small RNAs as morphogens”, *Curr Opin Cell Biol*, Apr. 2012, **24**, 217–24, DOI: 10.1016/j.ceb.2011.12.006.
- (208) R. N. Douglas, D. Wiley, A. Sarkar, N. Springer, M. C. Timmermans and M. J. Scanlon, “*ragged seedling2* encodes an ARGONAUTE7-like protein required for mediolateral expansion, but not dorsiventrality, of maize leaves”, *Plant Cell*, May 2010, **22**, 1441–51, DOI: 10.1105/tpc.109.071613.

- (209) M. C. Dotto, K. A. Petsch, M. J. Aukerman, M. Beatty, M. Hammell and M. C. Timmermans, “Genome-wide analysis of *leafbladeless1*-regulated and phased small RNAs underscores the importance of the *TAS3* ta-siRNA pathway to maize development”, *PLOS Genet*, Dec. 2014, **10**, e1004826, DOI: 10.1371/journal.pgen.1004826.
- (210) J. Yan, X. Cai, J. Luo, S. Sato, Q. Jiang, J. Yang, X. Cao, X. Hu, S. Tabata, P. M. Gresshoff and D. Luo, “The *REDUCED LEAFLET* genes encode key components of the *trans*-acting small interfering RNA pathway and regulate compound leaf and flower development in *Lotus japonicus*”, *Plant Physiol*, Feb. 2010, **152**, 797–807, DOI: 10.1104/pp.109.140947.
- (211) C. Zhou, L. Han, C. Fu, J. Wen, X. Cheng, J. Nakashima, J. Ma, Y. Tang, Y. Tan, M. Tadege, K. S. Mysore, G. Xia and Z. Y. Wang, “The *trans*-acting short interfering RNA3 pathway and NO APICAL MERISTEM antagonistically regulate leaf margin development and lateral organ separation, as revealed by analysis of an *argonaute7/lobed leaflet1* mutant in *Medicago truncatula*”, *Plant Cell*, Dec. 2013, **25**, 4845–62, DOI: 10.1105/tpc.113.117788.
- (212) N. Baumberger and D. C. Baulcombe, “*Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs”, *Proc Natl Acad Sci U S A*, Aug. 2005, **102**, 11928–33, DOI: 10.1073/pnas.0505461102.
- (213) K. Bonaldi, Z. Li, S. E. Kang, G. Breton and J. L. Pruneda-Paz, “Novel cell surface luciferase reporter for high-throughput yeast one-hybrid screens”, *Nucleic Acids Res*, DOI: 10.1093/nar/gkx682.
- (214) A. J. Walhout, “What does biologically meaningful mean? A perspective on gene regulatory network validation”, *Genome Biol*, 2011, **12**, 109, DOI: 10.1186/gb-2011-12-4-109.

- (215) R. P. Birkenbihl, G. Jach, H. Saedler and P. Huijser, “Functional dissection of the plant-specific SBP-domain: Overlap of the DNA-binding and nuclear localization domains”, *J Mol Biol*, Sept. 2005, **352**, 585–96, DOI: 10.1016/j.jmb.2005.07.013.
- (216) J. N. Maloof, K. Nozue, M. R. Mumbach and C. M. Palmer, “LeafJ: An ImageJ plugin for semi-automated leaf shape measurement”, *J Vis Exp*, Jan. 2013, DOI: 10.3791/50028.
- (217) A. Carbonell, N. Fahlgren, H. Garcia-Ruiz, K. B. Gilbert, T. A. Montgomery, T. Nguyen, J. T. Cuperus and J. C. Carrington, “Functional analysis of three *Arabidopsis* ARGONAUTES using slicer-defective mutants”, *Plant Cell*, Sept. 2012, **24**, 3613–29, DOI: 10.1105/tpc.112.099945.
- (218) Z. Lu, H. Yu, G. Xiong, J. Wang, Y. Jiao, G. Liu, Y. Jing, X. Meng, X. Hu, Q. Qian, X. Fu, Y. Wang and J. Li, “Genome-wide binding analysis of the transcription activator IDEAL PLANT ARCHITECTURE1 reveals a complex network regulating rice plant architecture”, *Plant Cell*, Oct. 2013, **25**, 3743–59, DOI: 10.1105/tpc.113.113639.
- (219) C. E. Grant, T. L. Bailey and W. S. Noble, “FIMO: Scanning for occurrences of a given motif”, *Bioinformatics*, Apr. 2011, **27**, 1017–8, DOI: 10.1093/bioinformatics/btr064.
- (220) T. L. Bailey, M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li and W. S. Noble, “MEME Suite: Tools for motif discovery and searching”, *Nucleic Acids Res*, July 2009, **37**, W202–8, DOI: 10.1093/nar/gkp335.
- (221) J. T. Odell, F. Nagy and N. H. Chua, “Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter”, *Nature*, 1985, **313**, 810–2, DOI: 10.1038/313810a0.
- (222) A. Helfer, D. A. Nusinow, B. Y. Chow, A. R. Gehrke, M. L. Bulyk and S. A. Kay, “*LUX ARRHYTHMO* encodes a nighttime repressor of circadian gene expression in the

- Arabidopsis* core clock”, *Curr Biol*, Jan. 2011, **21**, 126–33, DOI: 10.1016/j.cub.2010.12.021.
- (223) M. D. Curtis and U. Grossniklaus, “A Gateway cloning vector set for high-throughput functional analysis of genes *in planta*”, *Plant Physiol*, Oct. 2003, **133**, 462–9, DOI: 10.1104/pp.103.027979.
- (224) M. Holsters, B. Silva, F. Van Vliet, C. Genetello, M. De Block, P. Dhaese, A. Depicker, D. Inzé, G. Engler, R. Villarroel, M. Van Montagu and J. Schell, “The functional organization of the nopaline *A. tumefaciens* plasmid pTiC58”, *Plasmid*, Mar. 1980, **3**, 212–230, DOI: 10.1016/0147-619X(80)90110-9.
- (225) S. J. Clough and A. F. Bent, “Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*”, *Plant J*, Dec. 1998, **16**, 735–43, DOI: 10.1046/j.1365-3113x.1998.00343.x.
- (226) K. Bomblies, “Whole mount GUS staining”, in *Arabidopsis: A Laboratory Manual*. Ed. D. Weigel and J. Glazebrook, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2002, pp. 243–245.
- (227) L. C. Strader, D. L. Wheeler, S. E. Christensen, J. C. Berens, J. D. Cohen, R. A. Rampey and B. Bartel, “Multiple facets of *Arabidopsis* seedling development require indole-3-butyric acid-derived auxin”, *Plant Cell*, Mar. 2011, **23**, 984–99, DOI: 10.1105/tpc.111.083071.
- (228) H. Vanhaeren, N. Gonzalez and D. Inzé, “A journey through a leaf: Phenomics analysis of leaf growth in *Arabidopsis thaliana*”, *Arabidopsis Book*, 2015, **13**, e0181, DOI: 10.1199/tab.0181.
- (229) N. J. Provart, J. Alonso, S. M. Assmann, D. Bergmann, S. M. Brady, J. Brkljacic, J. Browse, C. Chapple, V. Colot, S. Cutler, J. Dangl, D. Ehrhardt, J. D. Friesner, W. B. Frommer, E. Grotewold, E. Meyerowitz, J. Nemhauser, M. Nordborg, C. Pikaard, J. Shanklin, C.

- Somerville, M. Stitt, K. U. Torii, J. Waese, D. Wagner and P. McCourt, “50 years of *Arabidopsis* research: Highlights and future directions”, *New Phytol*, Feb. 2016, **209**, 921–44, DOI: 10.1111/nph.13687.
- (230) F. Apelt, D. Breuer, Z. Nikoloski, M. Stitt and F. Kragler, “Phytotyping(4D): A light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth”, *Plant J*, May 2015, **82**, 693–706, DOI: 10.1111/tpj.12833.
- (231) H. Tsukaya, “Leaf development”, *Arabidopsis Book*, 2013, **11**, e0163, DOI: 10.1199/tab.0163.
- (232) R. S. Poethig, “Small RNAs and developmental timing in plants”, *Curr Opin Genet Dev*, Aug. 2009, **19**, 374–8, DOI: 10.1016/j.gde.2009.06.001.
- (233) A. Tridgell, *Efficient algorithms for sorting and synchronization*, Australian National University Canberra, 1999.
- (234) N. Fahlgren, M. Feldman, M. A. Gehan, M. S. Wilson, C. Shyu, D. W. Bryant, S. T. Hill, C. J. McEntee, S. N. Warnasooriya, I. Kumar, T. Ficor, S. Turnipseed, K. B. Gilbert, T. P. Brutnell, J. C. Carrington, T. C. Mockler and I. Baxter, “A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in *Setaria*”, *Mol Plant*, Oct. 2015, **8**, 1520–35, DOI: 10.1016/j.molp.2015.06.005.
- (235) J.-M. Pape and C. Klukas, “3-D histogram-based segmentation and leaf detection for rosette plants”, in *European Conference on Computer Vision Workshops (ECCV)*, ed. L. Agapito, M. M. Bronstein and C. Rother, Springer International Publishing, Sept. 2014, pp. 61–74, DOI: 10.1007/978-3-319-16220-1_5.
- (236) X. Zhang, R. J. Hause, Jr and J. O. Borevitz, “Natural genetic variation for growth and development revealed by high-throughput phenotyping in *Arabidopsis thaliana*”, *G3*, Jan. 2012, **2**, 29–34, DOI: 10.1534/g3.111.001487.

- (237) M. Lièvre, N. Wuyts, S. J. Cookson, J. Bresson, M. Dapp, F. Vasseur, C. Massonnet, S. Tisné, M. Bettembourg, C. Balsera, A. Bédiée, F. Bouvery, M. Dauzat, G. Rolland, D. Vile and C. Granier, “Phenotyping the kinematics of leaf development in flowering plants: recommendations and pitfalls”, *Wiley Interdiscip Rev Dev Biol*, 2013, **2**, 809–21, DOI: 10.1002/wdev.119.
- (238) D. Leister, C. Varotto, P. Pesaresi, A. Niwergall and F. Salamini, “Large-scale evaluation of plant growth in *Arabidopsis thaliana* by non-invasive image analysis”, *Plant Physiology and Biochemistry*, Sept. 1999, **37**, 671–678, DOI: 10.1016/S0981-9428(00)80097-2.
- (239) M. Minervini, M. V. Giuffrida, P. Perata and S. A. Tsiftaris, “Phenotiki: An open software and hardware platform for affordable and easy image-based phenotyping of rosette-shaped plants”, *Plant J*, Apr. 2017, **90**, 204–216, DOI: 10.1111/tpj.13472.
- (240) M. Minervini, M. M. Abdelsamea and S. A. Tsiftaris, “Image-based plant phenotyping with incremental learning and active contours”, *Ecol Inform*, Sept. 2014, **23**, 35–48, DOI: 10.1016/j.ecoinf.2013.07.004.
- (241) H. Scharr, M. Minervini, A. P. French, C. Klukas, D. M. Kramer, X. Liu, I. Luengo, J.-M. Pape, G. Polder, D. Vukadinovic, X. Yin and S. A. Tsiftaris, “Leaf segmentation in plant phenotyping: A collation study”, *Mach Vis Appl*, May 2016, **27**, 585–606, DOI: 10.1007/s00138-015-0737-3.
- (242) O. L. Tessmer, Y. Jiao, J. A. Cruz, D. M. Kramer and J. Chen, “Functional approach to high-throughput plant growth analysis”, *BMC Syst Biol*, 2013, **7 Suppl 6**, S17, DOI: 10.1186/1752-0509-7-S6-S17.
- (243) A. M. Mutka and R. S. Bart, “Image-based phenotyping of plant disease symptoms”, *Front Plant Sci*, 2014, **5**, 734, DOI: 10.3389/fpls.2014.00734.

- (244) A. M. Mutka, S. J. Fentress, J. W. Sher, J. C. Berry, C. Pretz, D. A. Nusinow and R. Bart, “Quantitative, image-based phenotyping methods provide insight into spatial and temporal dimensions of plant disease”, *Plant Physiol*, Oct. 2016, **172**, 650–660, DOI: 10.1104/pp.16.00984.
- (245) N. An, C. M. Palmer, R. L. Baker, R. J. C. Markelz, J. Ta, M. F. Covington, J. N. Maloof, S. M. Welch and C. Weinig, “Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area”, *Comput Electron Agric*, Sept. 2016, **127**, 376–394, DOI: 10.1016/j.compag.2016.04.002.
- (246) N. An, S. M. Welch, R. J. C. Markelz, R. L. Baker, C. M. Palmer, J. Ta, J. N. Maloof and C. Weinig, “Quantifying time-series of leaf morphology using 2D and 3D photogrammetry methods for high-throughput plant phenotyping”, *Comput Electron Agric*, Apr. 2017, **135**, 222–232, DOI: 10.1016/j.compag.2017.02.001.
- (247) H. Cwiek-Kupczyńska, T. Altmann, D. Arend, E. Arnaud, D. Chen, G. Cornut, F. Fiorani, W. Frohberg, A. Junker, C. Klukas, M. Lange, C. Mazurek, A. Nafissi, P. Neveu, J. van Oeveren, C. Pommier, H. Poorter, P. Rocca-Serra, S. A. Sansone, U. Scholz, M. van Schriek, Ü. Seren, B. Usadel, S. Weise, P. Kersey and P. Krajewski, “Measures for interoperability of phenotypic data: Minimum information requirements and formatting”, *Plant Methods*, 2016, **12**, 44, DOI: 10.1186/s13007-016-0144-4.
- (248) D. Arend, M. Lange, J. Chen, C. Colmsee, S. Flemming, D. Hecht and U. Scholz, “e!DAL—a framework to store, share and publish research data”, *BMC Bioinformatics*, June 2014, **15**, 214, DOI: 10.1186/1471-2105-15-214.
- (249) D. Arend, A. Junker, U. Scholz, D. Schüller, J. Wylie and M. Lange, “PGP repository: A plant phenomics and genomics data publication infrastructure”, *Database*, 2016, **2016**, DOI: 10.1093/database/baw033.

- (250) C. J. Brien, B. Berger, H. Rabie and M. Tester, “Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems”, *Plant Methods*, Feb. 2013, **9**, 5, DOI: 10.1186/1746-4811-9-5.
- (251) M. Minervini, A. Fischbach, H. Scharr and S. A. Tsaftaris, “Finely-grained annotated datasets for image-based plant phenotyping”, *Pattern Recognition Lett*, 2016, DOI: 10.1016/j.patrec.2015.10.013.
- (252) M. Minervini, M. V. Giuffrida and S. Tsaftaris, “An interactive tool for semi-automated leaf annotation”, in *Proc Computer Vision Problems in Plant Phenotyping Workshop (CVPPP)*, British Machine Vision Association, 2015, pp. 6.1–6.13, DOI: 10.5244/C.29.CVPPP.6.
- (253) M. V. Giuffrida, M. Minervini and S. Tsaftaris, “Learning to count leaves in rosette plants”, in *Proc Computer Vision Problems in Plant Phenotyping Workshop (CVPPP)*, British Machine Vision Association, 2015, pp. 1.1–1.13, DOI: 10.5244/C.29.CVPPP.1.
- (254) J. A. Cruz, X. Yin, X. Liu, S. M. Imran, D. D. Morris, D. M. Kramer and J. Chen, “Multi-modality imagery database for plant phenotyping”, *Mach Vis Appl*, Nov. 2015, **27**, 735–749, DOI: 10.1007/s00138-015-0734-6.
- (255) X. Yin, X. Liu, J. Chen and D. M. Kramer, “Multi-leaf tracking from fluorescence plant videos”, in *Image Processing (ICIP), 2014 IEEE International Conference on*, IEEE, 2014, pp. 408–412, DOI: <http://doi.org/10.1109/ICIP.2014.7025081>.
- (256) X. Yin, X. Liu, J. Chen and D. M. Kramer, “Multi-leaf alignment from fluorescence plant images”, in *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, IEEE, 2014, pp. 437–444, DOI: 10.1109/WACV.2014.6836067.
- (257) X. Yin, x. Liu, J. Chen and D. M. Kramer, “Joint multi-leaf segmentation, alignment and tracking from fluorescence plant videos”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **PP**, 1–1, DOI: 10.1109/TPAMI.2017.2728065.

- (258) J. Bell and H. M. Dee, “Aberystwyth Leaf Evaluation Dataset”, *Zenodo*, 2016, DOI: 10.5281/zenodo.168158.
- (259) A. Junker, M. M. Muraya, K. Weigelt-Fischer, F. Arana-Ceballos, C. Klukas, A. E. Melchinger, R. C. Meyer, D. Riewe and T. Altmann, “Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems”, *Front Plant Sci*, 2014, **5**, 770, DOI: 10.3389/fpls.2014.00770.
- (260) D. Arend, M. Lange, J. M. Pape, K. Weigelt-Fischer, F. Arana-Ceballos, I. Mücke, C. Klukas, T. Altmann, U. Scholz and A. Junker, “Quantitative monitoring of *Arabidopsis thaliana* growth and development using high-throughput plant phenotyping”, *Sci Data*, Aug. 2016, **3**, 160055, DOI: 10.1038/sdata.2016.55.
- (261) D. Kelly, A. Vatsa, W. Mayham, L. Ngô, A. Thompson and T. Kazic, “An opinion on imaging challenges in phenotyping field crops”, *Mach Vis Appl*, July 2016, **27**, 681–694, DOI: 10.1007/s00138-015-0728-4.
- (262) S. A. Tsaftaris, M. Minervini and H. Scharr, “Machine learning for plant phenotyping needs image processing”, *Trends Plant Sci*, Dec. 2016, **21**, 989–991, DOI: 10.1016/j.tplants.2016.10.002.
- (263) C. Massonnet, D. Vile, J. Fabre, M. A. Hannah, C. Caldana, J. Lisec, G. T. Beemster, R. C. Meyer, G. Messerli, J. T. Gronlund, J. Perkovic, E. Wigmore, S. May, M. W. Bevan, C. Meyer, S. Rubio-Díaz, D. Weigel, J. L. Micol, V. Buchanan-Wollaston, F. Fiorani, S. Walsh, B. Rinn, W. Gruissem, P. Hilson, L. Hennig, L. Willmitzer and C. Granier, “Probing the reproducibility of leaf growth and molecular phenotypes: a comparison of three *Arabidopsis* accessions cultivated in ten laboratories”, *Plant Physiol*, Apr. 2010, **152**, 2142–57, DOI: 10.1104/pp.109.148338.

- (264) D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo and D. Dennison, “Hidden technical debt in machine learning systems”, in *Advances in Neural Information Processing Systems* 28, ed. C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett, Curran Associates, Inc., 2015, pp. 2503–2511.
- (265) C. Harper and M. Siller, “OpenAG: A globally distributed network of food computing”, *IEEE Pervasive Computing*, Oct. 2015, **14**, 24–27, DOI: 10.1109/MPRV.2015.72.
- (266) F. Katagiri, D. Canelon-Suarez, K. Griffin, J. Petersen, R. K. Meyer, M. Siegle and K. Mase, “Design and construction of an inexpensive homemade plant growth chamber”, *PLOS One*, 2015, **10**, e0126826, DOI: 10.1371/journal.pone.0126826.
- (267) E. Castelló Ferrer, J. Rye, G. Brander, T. Savas, D. Chambers, H. England and C. Harper, “Personal Food Computer: A new device for controlled-environment agriculture”, in *Future Technologies Conference (FTC)*, IEEE, June 2017.
- (268) W. Engelmann, K. Simon and C. J. Phen, “Leaf movement rhythm in *Arabidopsis thaliana*”, *Zeitschrift für Naturforschung C*, 1992, **47**, 925–928, DOI: 10.1515/znc-1992-11-1222.
- (269) S. A. Tsaftaris and C. Noutsos, “Plant phenotyping with low cost digital cameras and image analytics”, in *Information Technologies in Environmental Engineering*, ed. D. I. N. Athanasiadis, P. A. E. Rizzoli, P. A. Mitkas and P. D.-I. J. M. Gómez, Springer Berlin Heidelberg, 2009, pp. 238–251, DOI: 10.1007/978-3-540-88351-7_18.
- (270) J. Kokorian, G. Polder, J. J. B. Keurentjes, D. Vreugdenhil and M. C. O. Guzman, “An ImageJ based measurement setup for automated phenotyping of plants”, in *Proc ImageJ User and Developer Conf*, Luxembourg, 2010, pp. 178–182.
- (271) A. Hartmann, T. Czauderna, R. Hoffmann, N. Stein and F. Schreiber, “HTPheno: an image analysis pipeline for high-throughput plant phenotyping”, *BMC Bioinformatics*, May 2011, **12**, 148, DOI: 10.1186/1471-2105-12-148.

- (272) J. Fabre, M. Dauzat, V. Nègre, N. Wuyts, A. Tireau, E. Gennari, P. Neveu, S. Tisné, C. Massonnet, I. Hummel and C. Granier, “PHENOPSIS DB: an information system for *Arabidopsis thaliana* phenotypic data in an environmental context”, *BMC Plant Biol*, May 2011, **11**, 77, DOI: 10.1186/1471-2229-11-77.
- (273) J. De Vylder, F. Vandenbussche, Y. Hu, W. Philips and D. Van Der Straeten, “Rosette Tracker: An open source image analysis tool for automatic quantification of genotype effects”, *Plant Physiol*, Nov. 2012, **160**, 1149–59, DOI: 10.1104/pp.112.202762.
- (274) M. Lièvre, C. Granier and Y. Guédon, “Identifying developmental phases in the *Arabidopsis thaliana* rosette using integrative segmentation models”, *New Phytol*, June 2016, **210**, 1466–78, DOI: 10.1111/nph.13861.
- (275) T. P. Pridmore, A. P. French and M. P. Pound, “What lies beneath: Underlying assumptions in bioimage analysis”, *Trends Plant Sci*, Dec. 2012, **17**, 688–92, DOI: 10.1016/j.tplants.2012.07.003.
- (276) M. Chalfie, H. R. Horvitz and J. E. Sulston, “Mutations that lead to reiterations in the cell lineages of *C. elegans*”, *Cell*, Apr. 1981, **24**, 59–69, DOI: 10.1016/0092-8674(81)90501-8.
- (277) V. Ambros and H. R. Horvitz, “Heterochronic mutants of the nematode *Caenorhabditis elegans*”, *Science*, Oct. 1984, **226**, 409–16, DOI: 10.1126/science.6494891.
- (278) A. Telfer and R. S. Poethig, “HASTY: a gene that regulates the timing of shoot maturation in *Arabidopsis thaliana*”, *Development*, May 1998, **125**, 1889–98.
- (279) G. Ruvkun, “The perfect storm of tiny RNAs”, *Nat Med*, Oct. 2008, **14**, 1041–5, DOI: 10.1038/nm1008-1041.
- (280) C. Nusslein-Volhard and E. Wieschaus, “Mutations affecting segment number and polarity in *Drosophila*”, *Nature*, Oct. 1980, **287**, 795–801.

- (281) E. Wieschaus and C. Nüsslein-Volhard, “The Heidelberg screen for pattern mutants of *Drosophila*: A personal account”, *Annu Rev Cell Dev Biol*, Oct. 2016, **32**, 1–46, DOI: 10.1146/annurev-cellbio-113015-023138.
- (282) E. B. Lewis, “Homeosis: The first 100 years”, *Trends Genet*, Oct. 1994, **10**, 341–3, DOI: 10.1007/978-1-4020-6345-9_28.
- (283) J. L. Bowman, D. R. Smyth and E. M. Meyerowitz, “Genetic interactions among floral homeotic genes of *Arabidopsis*”, *Development*, May 1991, **112**, 1–20.
- (284) J. L. Bowman, D. R. Smyth and E. M. Meyerowitz, “The ABC model of flower development: Then and now”, *Development*, Nov. 2012, **139**, 4095–8, DOI: 10.1242/dev.083972.
- (285) F. Slack and G. Ruvkun, “Temporal pattern formation by heterochronic genes”, *Annu Rev Genet*, 1997, **31**, 611–34, DOI: 10.1146/annurev.genet.31.1.611.
- (286) A. Veraksa, M. Del Campo and W. McGinnis, “Developmental patterning genes and their conserved functions: From model organisms to humans”, *Mol Genet Metab*, Feb. 2000, **69**, 85–100, DOI: 10.1006/mgme.2000.2963.
- (287) L. Wolpert, “Positional information and the spatial pattern of cellular differentiation”, *J Theor Biol*, Oct. 1969, **25**, 1–47, DOI: 10.1016/S0022-5193(69)80016-0.
- (288) J. W. Hong, D. A. Hendrix, D. Papatsenko and M. S. Levine, “How the Dorsal gradient works: Insights from postgenome technologies”, *Proc Natl Acad Sci U S A*, Dec. 2008, **105**, 20072–6, DOI: 10.1073/pnas.0806476105.
- (289) A. Stathopoulos and D. Iber, “Studies of morphogens: Keep calm and carry on”, *Development*, Oct. 2013, **140**, 4119–24, DOI: 10.1242/dev.095141.
- (290) A. M. Turing, “The chemical basis of morphogenesis”, *Phil Trans R Soc Lond B Biol Sci*, Aug. 1952, **237**, 37–72, DOI: 10.1098/rstb.1952.0012.

- (291) L. Wolpert, “One hundred years of positional information”, *Trends Genet*, Sept. 1996, **12**, 359–64, DOI: 10.1016/S0168-9525(96)80019-9.
- (292) R. P. Bhalerao and M. J. Bennett, “The case for morphogens in plants”, *Nat Cell Biol*, Nov. 2003, **5**, 939–43, DOI: 10.1038/ncb1103-939.
- (293) O. Leyser, “Auxin signaling”, *Plant Physiol*, Aug. 2017, DOI: 10.1104/pp.17.00765.
- (294) E. Levine, P. McHale and H. Levine, “Small regulatory RNAs may sharpen spatial expression patterns”, *PLOS Comput Biol*, Nov. 2007, **3**, e233, DOI: 10.1371/journal.pcbi.0030233.
- (295) M. H. Frank and D. H. Chitwood, “Plant chimeras: The good, the bad, and the ‘Bizzaria’”, *Dev Biol*, Nov. 2016, **419**, 41–53, DOI: 10.1016/j.ydbio.2016.07.003.
- (296) M. Nakata, N. Matsumoto, R. Tsugeki, E. Rikirsch, T. Laux and K. Okada, “Roles of the middle domain-specific *WUSCHEL-RELATED HOMEODOMAIN* genes in early development of leaves in *Arabidopsis*”, *Plant Cell*, Feb. 2012, **24**, 519–35, DOI: 10.1105/tpc.111.092858.
- (297) C. Guan, B. Wu, T. Yu, Q. Wang, N. T. Krogan, X. Liu and Y. Jiao, “Spatial auxin signaling controls leaf flattening in *Arabidopsis*”, *Curr Biol*, Sept. 2017, DOI: 10.1016/j.cub.2017.08.042.
- (298) D. C. Henderson, G. J. Muehlbauer and M. J. Scanlon, “Radial leaves of the maize mutant *ragged seedling2* retain dorsiventral anatomy”, *Dev Biol*, June 2005, **282**, 455–66, DOI: 10.1016/j.ydbio.2005.03.027.
- (299) T. J. Guilfoyle and G. Hagen, “Auxin response factors”, *Curr Opin Plant Biol*, Oct. 2007, **10**, 453–60, DOI: 10.1016/j.pbi.2007.08.014.

- (300) D. Reinhardt, E. R. Pesce, P. Stieger, T. Mandel, K. Baltensperger, M. Bennett, J. Traas, J. Friml and C. Kuhlemeier, “Regulation of phyllotaxis by polar auxin transport”, *Nature*, Nov. 2003, **426**, 255–60, DOI: 10.1038/nature02081.
- (301) C. Kuhlemeier, “Phyllotaxis”, *Curr Biol*, Sept. 2017, **27**, R882–R887, DOI: 10.1016/j.cub.2017.05.069.
- (302) J. Qi, Y. Wang, T. Yu, A. Cunha, B. Wu, T. Vernoux, E. Meyerowitz and Y. Jiao, “Auxin depletion from leaf primordia contributes to organ patterning”, *Proc Natl Acad Sci U S A*, Dec. 2014, **111**, 18769–74, DOI: 10.1073/pnas.1421878112.
- (303) J. Shi, J. Dong, J. Xue, H. Wang, Z. Yang, Y. Jiao, L. Xu and H. Huang, “Model for the role of auxin polar transport in patterning of the leaf adaxial-abaxial axis”, *Plant J*, Aug. 2017, DOI: 10.1111/tpj.13670.
- (304) M. P. Caggiano, X. Yu, N. Bhatia, A. Larsson, H. Ram, C. K. Ohno, P. Sappl, E. M. Meyerowitz, H. Jönsson and M. G. Heisler, “Cell type boundaries organize plant development”, *eLife*, Sept. 2017, **6**, DOI: 10.7554/eLife.27421.
- (305) P. Merelo, E. B. Paredes, M. G. Heisler and S. Wenkel, “The shady side of leaf development: The role of the REVOLUTA/KANADI1 module in leaf patterning and auxin-mediated growth promotion”, *Curr Opin Plant Biol*, Feb. 2017, **35**, 111–116, DOI: 10.1016/j.pbi.2016.11.016.
- (306) M. Levine, “Transcriptional enhancers in animal development and evolution”, *Curr Biol*, Sept. 2010, **20**, R754–63, DOI: 10.1016/j.cub.2010.06.070.
- (307) M. Rebeiz and M. Tsiantis, “Enhancer evolution and the origins of morphological novelty”, *Curr Opin Genet Dev*, Aug. 2017, **45**, 115–123, DOI: 10.1016/j.gde.2017.04.006.

- (308) A. Studer, Q. Zhao, J. Ross-Ibarra and J. Doebley, “Identification of a functional transposon insertion in the maize domestication gene *tb1*”, *Nat Genet*, Sept. 2011, **43**, 1160–3, DOI: 10.1038/ng.942.
- (309) M. Levine, C. Cattoglio and R. Tjian, “Looping back to leap forward: Transcription enters a new era”, *Cell*, Mar. 2014, **157**, 13–25, DOI: 10.1016/j.cell.2014.02.009.
- (310) N. Y. Rodriguez-Granados, J. S. Ramirez-Prado, A. Veluchamy, D. Latrasse, C. Raynaud, M. Crespi, F. Ariel and M. Benhamed, “Put your 3D glasses on: Plant chromatin is on show”, *J Exp Bot*, May 2016, **67**, 3205–21, DOI: 10.1093/jxb/erw168.
- (311) K. C. Dobi and F. Winston, “Analysis of transcriptional activation at a distance in *Saccharomyces cerevisiae*”, *Mol Cell Biol*, Aug. 2007, **27**, 5575–86, DOI: 10.1128/MCB.00459-07.
- (312) C. Liu and D. Weigel, “Chromatin in 3D: Progress and prospects for plants”, *Genome Biol*, Aug. 2015, **16**, 170, DOI: 10.1186/s13059-015-0738-6.
- (313) T. T. Hu, P. Pattyn, E. G. Bakker, J. Cao, J. F. Cheng, R. M. Clark, N. Fahlgren, J. A. Fawcett, J. Grimwood, H. Gundlach, G. Haberer, J. D. Hollister, S. Ossowski, R. P. Ottilar, A. A. Salamov, K. Schneeberger, M. Spannagl, X. Wang, L. Yang, M. E. Nasrallah, J. Bergelson, J. C. Carrington, B. S. Gaut, J. Schmutz, K. F. Mayer, Y. Van de Peer, I. V. Grigoriev, M. Nordborg, D. Weigel and Y. L. Guo, “The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change”, *Nat Genet*, May 2011, **43**, 476–81, DOI: 10.1038/ng.807.
- (314) C. Liu, Y. J. Cheng, J. W. Wang and D. Weigel, “Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*”, *Nat Plants*, Aug. 2017, DOI: 10.1038/s41477-017-0005-9.
- (315) S. Barolo, “How to tune an enhancer”, *Proc Natl Acad Sci U S A*, June 2016, **113**, 6330–1, DOI: 10.1073/pnas.1606109113.

- (316) J. W. Hong, D. A. Hendrix and M. S. Levine, “Shadow enhancers as a source of evolutionary novelty”, *Science*, Sept. 2008, **321**, 1314, DOI: 10.1126/science.1160631.
- (317) M. W. Perry, A. N. Boettiger, J. P. Bothma and M. Levine, “Shadow enhancers foster robustness of *Drosophila* gastrulation”, *Curr Biol*, Sept. 2010, **20**, 1562–7, DOI: 10.1016/j.cub.2010.07.043.
- (318) N. Frankel, G. K. Davis, D. Vargas, S. Wang, F. Payre and D. L. Stern, “Phenotypic robustness conferred by apparently redundant transcriptional enhancers”, *Nature*, July 2010, **466**, 490–3, DOI: 10.1038/nature09158.
- (319) E. A. Miska, E. Alvarez-Saavedra, A. L. Abbott, N. C. Lau, A. B. Hellman, S. M. McGonagle, D. P. Bartel, V. R. Ambros and H. R. Horvitz, “Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability”, *PLOS Genet*, Dec. 2007, **3**, e215, DOI: 10.1371/journal.pgen.0030215.
- (320) E. Alvarez-Saavedra and H. R. Horvitz, “Many families of *C. elegans* microRNAs are not essential for development or viability”, *Curr Biol*, Feb. 2010, **20**, 367–73, DOI: 10.1016/j.cub.2009.12.051.
- (321) N. Peláez and R. W. Carthew, “Biological robustness and the role of microRNAs: A network perspective”, *Curr Top Dev Biol*, 2012, **99**, 237–55, DOI: 10.1016/B978-0-12-387038-4.00009-4.
- (322) R. S. Poethig, “Phase change and the regulation of developmental timing in plants”, *Science*, July 2003, **301**, 334–6, DOI: 10.1126/science.1085328.
- (323) M. C. Canver, E. C. Smith, F. Sher, L. Pinello, N. E. Sanjana, O. Shalem, D. D. Chen, P. G. Schupp, D. S. Vinjamur, S. P. Garcia, S. Luc, R. Kurita, Y. Nakamura, Y. Fujiwara, T. Maeda, G. C. Yuan, F. Zhang, S. H. Orkin and D. E. Bauer, “*BCL11A* enhancer dissection

- by Cas9-mediated *in situ* saturating mutagenesis”, *Nature*, Nov. 2015, **527**, 192–7, DOI: 10.1038/nature15521.
- (324) D. Rodríguez-Leal, Z. H. Lemmon, J. Man, M. E. Bartlett and Z. B. Lippman, “Engineering quantitative trait variation for crop improvement by genome editing”, *Cell*, Oct. 2017, **171**, 470–480.e8, DOI: 10.1016/j.cell.2017.08.030.
- (325) G. R. Fink, “Anatomy of a revolution.”, *Genetics*, June 1998, **149**, 473–7.
- (326) L. Li, K. Ljung, G. Breton, R. J. Schmitz, J. Pruneda-Paz, C. Cowing-Zitron, B. J. Cole, L. J. Ivans, U. V. Pedmale, H. S. Jung, J. R. Ecker, S. A. Kay and J. Chory, “Linking photoreceptor excitation to changes in plant architecture”, *Genes Dev*, Apr. 2012, **26**, 785–90, DOI: 10.1101/gad.187849.112.
- (327) C. S. Gillmor, C. O. Silva-Ortega, M. R. Willmann, M. Buendía-Monreal and R. S. Poethig, “The *Arabidopsis* Mediator CDK8 module genes CCT (MED12) and GCT (MED13) are global regulators of developmental phase transitions”, *Development*, Dec. 2014, **141**, 4580–9, DOI: 10.1242/dev.111229.
- (328) S. Picó, M. I. Ortiz-Marchena, W. Merini and M. Calonje, “Deciphering the role of POLY-COMB REPRESSIVE COMPLEX1 variants in regulating the acquisition of flowering competence in *Arabidopsis*”, *Plant Physiol*, Aug. 2015, **168**, 1286–97, DOI: 10.1104/pp.15.00073.
- (329) M. Xu, T. Hu, M. R. Smith and R. S. Poethig, “Epigenetic regulation of vegetative phase change in *Arabidopsis*”, *Plant Cell*, Jan. 2016, **28**, 28–41, DOI: 10.1105/tpc.15.00854.
- (330) Y. Xu, C. Guo, B. Zhou, C. Li, H. Wang, B. Zheng, H. Ding, Z. Zhu, A. Peragine, Y. Cui, S. Poethig and G. Wu, “Regulation of vegetative phase change by SWI2/SNF2 chromatin remodeling ATPase BRAHMA”, *Plant Physiol*, Dec. 2016, **172**, 2416–2428, DOI: 10.1104/pp.16.01588.

- (331) M. Buendía-Monreal and C. S. Gillmor, “Convergent repression of miR156 by sugar and the CDK8 module of *Arabidopsis* Mediator”, *Dev Biol*, Mar. 2017, **423**, 19–23, DOI: 10.1016/j.ydbio.2017.01.007.
- (332) A. Stief, S. Altmann, K. Hoffmann, B. D. Pant, W. R. Scheible and I. Bäurle, “*Arabidopsis* miR156 regulates tolerance to recurring environmental stress through SPL transcription factors”, *Plant Cell*, Apr. 2014, **26**, 1792–1807, DOI: 10.1105/tpc.114.123851.
- (333) I. G. Matsoukas, A. J. Massiah and B. Thomas, “Starch metabolism and antiflorigenic signals modulate the juvenile-to-adult phase transition in *Arabidopsis*”, *Plant Cell Environ*, Oct. 2013, **36**, 1802–11, DOI: 10.1111/pce.12088.
- (334) Y. H. Chew, B. Wenden, A. Flis, V. Mengin, J. Taylor, C. L. Davey, C. Tindal, H. Thomas, H. J. Ougham, P. de Reffye, M. Stitt, M. Williams, R. Muetzelfeldt, K. J. Halliday and A. J. Millar, “Multiscale digital *Arabidopsis* predicts individual organ and whole-organism growth”, *Proc Natl Acad Sci U S A*, Sept. 2014, **111**, E4127–36, DOI: 10.1073/pnas.1410238111.
- (335) A. L. Eveland and D. P. Jackson, “Sugars, signalling, and plant development”, *J Exp Bot*, May 2012, **63**, 3367–77, DOI: 10.1093/jxb/err379.
- (336) S. Yu, H. Lian and J. W. Wang, “Plant developmental transitions: The role of microRNAs and sugars”, *Curr Opin Plant Biol*, Oct. 2015, **27**, 1–7, DOI: 10.1016/j.pbi.2015.05.009.
- (337) M. R. Tucker, T. Okada, Y. Hu, A. Scholefield, J. M. Taylor and A. M. Koltunow, “Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in *Arabidopsis*”, *Development*, Apr. 2012, **139**, 1399–404, DOI: 10.1242/dev.075390.
- (338) M. J. Axtell, “Lost in translation? microRNAs at the rough ER”, *Trends Plant Sci*, Apr. 2017, **22**, 273–274, DOI: 10.1016/j.tplants.2017.03.002.

- (339) A. Carbonell, “Plant ARGONAUTES: Features, functions, and unknowns”, *Methods Mol Biol*, 2017, **1640**, 1–21, DOI: 10.1007/978-1-4939-7165-7_1.
- (340) W. Filipowicz, S. N. Bhattacharyya and N. Sonenberg, “Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight?”, *Nat Rev Genet*, Feb. 2008, **9**, 102–14, DOI: 10.1038/nrg2290.
- (341) E. Huntzinger and E. Izaurralde, “Gene silencing by microRNAs: Contributions of translational repression and mRNA decay”, *Nat Rev Genet*, Feb. 2011, **12**, 99–110, DOI: 10.1038/nrg2936.
- (342) H. Guo, N. T. Ingolia, J. S. Weissman and D. P. Bartel, “Mammalian microRNAs predominantly act to decrease target mRNA levels”, *Nature*, Aug. 2010, **466**, 835–40, DOI: 10.1038/nature09267.
- (343) S. W. Eichhorn, H. Guo, S. E. McGeary, R. A. Rodriguez-Mias, C. Shin, D. Baek, S. H. Hsu, K. Ghoshal, J. Villén and D. P. Bartel, “mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues”, *Mol Cell*, Oct. 2014, **56**, 104–15, DOI: 10.1016/j.molcel.2014.08.028.
- (344) S. Jonas and E. Izaurralde, “Towards a molecular understanding of microRNA-mediated gene silencing”, *Nat Rev Genet*, July 2015, **16**, 421–33, DOI: 10.1038/nrg3965.
- (345) D. V. Dugas and B. Bartel, “Sucrose induction of *Arabidopsis* miR398 represses two Cu/Zn superoxide dismutases”, *Plant Mol Biol*, July 2008, **67**, 403–17, DOI: 10.1007/s11103-008-9329-1.
- (346) L. Beauclair, A. Yu and N. Bouché, “microRNA-directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in *Arabidopsis*”, *Plant J*, May 2010, **62**, 454–62, DOI: 10.1111/j.1365-313X.2010.04162.x.

- (347) C. Brousse, Q. Liu, L. Beauclair, A. Deremetz, M. J. Axtell and N. Bouché, “A non-canonical plant microRNA target site”, *Nucleic Acids Res*, Apr. 2014, **42**, 5270–9, DOI: 10.1093/nar/gku157.
- (348) F. Wang, S. Polydore and M. J. Axtell, “More than meets the eye? Factors that affect target selection by plant miRNAs and heterochromatic siRNAs”, *Curr Opin Plant Biol*, Oct. 2015, **27**, 118–24, DOI: 10.1016/j.pbi.2015.06.012.
- (349) M. Gandikota, R. P. Birkenbihl, S. Höhmann, G. H. Cardon, H. Saedler and P. Huijser, “The miRNA156/157 recognition element in the 3' UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings”, *Plant J*, Feb. 2007, **49**, 683–93, DOI: 10.1111/j.1365-3113.2006.02983.x.
- (350) L. Yang, G. Wu and R. S. Poethig, “Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in *Arabidopsis*”, *Proc Natl Acad Sci U S A*, Jan. 2012, **109**, 315–20, DOI: 10.1073/pnas.1114673109.
- (351) Q. Liu, F. Wang and M. J. Axtell, “Analysis of complementarity requirements for plant microRNA targeting using a *Nicotiana benthamiana* quantitative transient assay”, *Plant Cell*, Feb. 2014, **26**, 741–53, DOI: 10.1105/tpc.113.120972.
- (352) C. Poulsen, H. Vaucheret and P. Brodersen, “Lessons on RNA silencing mechanisms in plants from eukaryotic Argonaute structures”, *Plant Cell*, Jan. 2013, **25**, 22–37, DOI: 10.1105/tpc.112.105643.
- (353) A. S. Flynt and E. C. Lai, “Biological principles of microRNA-mediated regulation: Shared themes amid diversity”, *Nat Rev Genet*, Nov. 2008, **9**, 831–42, DOI: 10.1038/nrg2455.
- (354) O. Hobert, “Gene regulation by transcription factors and microRNAs”, *Science*, Mar. 2008, **319**, 1785–6, DOI: 10.1126/science.1151651.

- (355) F. F. de Felippes, F. Ott and D. Weigel, “Comparative analysis of non-autonomous effects of tasiRNAs and miRNAs in *Arabidopsis thaliana*”, *Nucleic Acids Res*, Apr. 2011, **39**, 2880–9, DOI: 10.1093/nar/gkq1240.
- (356) L. Arribas-Hernández, A. Marchais, C. Poulsen, B. Haase, J. Hauptmann, V. Benes, G. Meister and P. Brodersen, “The slicer activity of ARGONAUTE1 is required specifically for the phasing, not production, of *trans*-acting short interfering RNAs in *Arabidopsis*”, *Plant Cell*, July 2016, **28**, 1563–80, DOI: 10.1105/tpc.16.00121.
- (357) M. Yoshikawa, T. Iki, Y. Tsutsui, K. Miyashita, R. S. Poethig, Y. Habu and M. Ishikawa, “3′ fragment of miR173-programmed RISC-cleaved RNA is protected from degradation in a complex with RISC and SGS3”, *Proc Natl Acad Sci U S A*, Mar. 2013, **110**, 4117–22, DOI: 10.1073/pnas.1217050110.
- (358) Z. Luo and Z. Chen, “Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in *Arabidopsis*”, *Plant Cell*, Mar. 2007, **19**, 943–58, DOI: 10.1105/tpc.106.045724.
- (359) K. Baeg, H. O. Iwakawa and Y. Tomari, “The poly(A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing”, *Nat Plants*, Mar. 2017, **3**, 17036, DOI: 10.1038/nplants.2017.36.
- (360) N. Kumakura, A. Takeda, Y. Fujioka, H. Motose, R. Takano and Y. Watanabe, “SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies”, *FEBS Lett*, Apr. 2009, **583**, 1261–6, DOI: 10.1016/j.febslet.2009.03.055.
- (361) V. Jouannet, A. B. Moreno, T. Elmayan, H. Vaucheret, M. D. Crespi and A. Maizel, “Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis”, *EMBO J*, Apr. 2012, **31**, 1704–13, DOI: 10.1038/emboj.2012.20.

- (362) S. Li, B. Le, X. Ma, S. Li, C. You, Y. Yu, B. Zhang, L. Liu, L. Gao, T. Shi, Y. Zhao, B. Mo, X. Cao and X. Chen, “Biogenesis of phased siRNAs on membrane-bound polysomes in *Arabidopsis*”, *eLife*, Dec. 2016, **5**, DOI: 10.7554/eLife.22750.
- (363) C. Y. Hou, W. C. Lee, H. C. Chou, A. P. Chen, S. J. Chou and H. M. Chen, “Global analysis of truncated RNA ends reveals new insights into ribosome stalling in plants”, *Plant Cell*, Oct. 2016, **28**, 2398–2416, DOI: 10.1105/tpc.16.00295.
- (364) H. M. Chen, L. T. Chen, K. Patel, Y. H. Li, D. C. Baulcombe and S. H. Wu, “22-nucleotide RNAs trigger secondary siRNA biogenesis in plants”, *Proc Natl Acad Sci U S A*, Aug. 2010, **107**, 15269–74, DOI: 10.1073/pnas.1001738107.
- (365) J. T. Cuperus, A. Carbonell, N. Fahlgren, H. Garcia-Ruiz, R. T. Burke, A. Takeda, C. M. Sullivan, S. D. Gilbert, T. A. Montgomery and J. C. Carrington, “Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in *Arabidopsis*”, *Nat Struct Mol Biol*, Aug. 2010, **17**, 997–1003, DOI: 10.1038/nsmb.1866.
- (366) P. A. Manavella, D. Koenig and D. Weigel, “Plant secondary siRNA production determined by microRNA-duplex structure”, *Proc Natl Acad Sci U S A*, Feb. 2012, **109**, 2461–6, DOI: 10.1073/pnas.1200169109.
- (367) T. Iki, M. Yoshikawa, T. Meshi and M. Ishikawa, “Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants”, *EMBO J*, Jan. 2012, **31**, 267–78, DOI: 10.1038/emboj.2011.395.
- (368) F. Frank, J. Hauver, N. Sonenberg and B. Nagar, “*Arabidopsis* Argonaute MID domains use their nucleotide specificity loop to sort small RNAs”, *EMBO J*, Aug. 2012, **31**, 3588–95, DOI: 10.1038/emboj.2012.204.

- (369) Y. Endo, H. O. Iwakawa and Y. Tomari, “*Arabidopsis* ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly”, *EMBO Rep*, July 2013, **14**, 652–8, DOI: 10.1038/embor.2013.73.
- (370) S. Nuthikattu, A. D. McCue, K. Panda, D. Fultz, C. DeFraia, E. N. Thomas and R. K. Slotkin, “The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs”, *Plant Physiol*, May 2013, **162**, 116–31, DOI: 10.1104/pp.113.216481.
- (371) A. D. McCue, K. Panda, S. Nuthikattu, S. G. Choudury, E. N. Thomas and R. K. Slotkin, “ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation”, *EMBO J*, Jan. 2015, **34**, 20–35, DOI: 10.15252/embj.201489499.
- (372) K. M. Creasey, J. Zhai, F. Borges, F. Van Ex, M. Regulski, B. C. Meyers and R. A. Martienssen, “miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*”, *Nature*, Apr. 2014, **508**, 411–5, DOI: 10.1038/nature13069.
- (373) N. Bao, K. W. Lye and M. K. Barton, “MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome”, *Dev Cell*, Nov. 2004, **7**, 653–62, DOI: 10.1016/j.devcel.2004.10.003.
- (374) M. D. Howell, N. Fahlgren, E. J. Chapman, J. S. Cumbie, C. M. Sullivan, S. A. Givan, K. D. Kasschau and J. C. Carrington, “Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting”, *Plant Cell*, Mar. 2007, **19**, 926–42, DOI: 10.1105/tpc.107.050062.
- (375) H. Kitano, “Systems biology: A brief overview”, *Science*, Mar. 2002, **295**, 1662–4, DOI: 10.1126/science.1069492.

- (376) M. Levine and C. Vicente, “An interview with Mike Levine”, *Development*, Oct. 2015, **142**, 3453–5, DOI: 10.1242/dev.130609.
- (377) P. N. Benfey and N. H. Chua, “The cauliflower mosaic virus 35S promoter: Combinatorial regulation of transcription in plants”, *Science*, Nov. 1990, **250**, 959–66, DOI: 10.1126/science.250.4983.959.
- (378) M. A. Busch, K. Bomblies and D. Weigel, “Activation of a floral homeotic gene in *Arabidopsis*”, *Science*, July 1999, **285**, 585–7, DOI: 10.1126/science.285.5427.585.
- (379) M. A. White, “Understanding how *cis*-regulatory function is encoded in DNA sequence using massively parallel reporter assays and designed sequences”, *Genomics*, Sept. 2015, **106**, 165–70, DOI: 10.1016/j.ygeno.2015.06.003.
- (380) G. Jürgens, R. A. T. Ruiz, T. Laux, U. Mayer and T. Berleth, “Early events in apical-basal pattern formation in *Arabidopsis*”, *Plant Mol Biol*, 1994, 95–103, DOI: 10.1007/978-3-642-78852-9_10.
- (381) J. R. McConnell and M. K. Barton, “Effect of mutations in the *PINHEAD* gene of *Arabidopsis* on the formation of shoot apical meristems”, *Dev Genet*, 1995, **16**, 358–366, DOI: 10.1002/dvg.1020160409.
- (382) A. Takeda, S. Iwasaki, T. Watanabe, M. Utsumi and Y. Watanabe, “The mechanism selecting the guide strand from small RNA duplexes is different among ARGONAUTE proteins”, *Plant Cell Physiol*, Apr. 2008, **49**, 493–500, DOI: 10.1093/pcp/pcn043.
- (383) A. C. Mallory, A. Hinze, M. R. Tucker, N. Bouché, V. Gascioli, T. Elmayan, D. Lauressergues, V. Jauvion, H. Vaucheret and T. Laux, “Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing”, *PLOS Genet*, Sept. 2009, **5**, e1000646, DOI: 10.1371/journal.pgen.1000646.

- (384) M. R. Tucker, F. Roodbarkelari, E. Truernit, N. M. Adamski, A. Hinze, B. Lohmüller, T. Würschum and T. Laux, “Accession-specific modifiers act with ZWILLE/ARGONAUTE10 to maintain shoot meristem stem cells during embryogenesis in *Arabidopsis*”, *BMC Genomics*, Nov. 2013, **14**, 809, DOI: 10.1186/1471-2164-14-809.
- (385) T. Z. Berardini, K. Bollman, H. Sun and R. S. Poethig, “Regulation of vegetative phase change in *Arabidopsis thaliana* by cyclophilin 40”, *Science*, Mar. 2001, **291**, 2405–7, DOI: 10.1126/science.1057144.
- (386) S. Tisne, F. Barbier and C. Granier, “The *ERECTA* gene controls spatial and temporal patterns of epidermal cell number and size in successive developing leaves of *Arabidopsis thaliana*”, *Ann Bot*, July 2011, **108**, 159–68, DOI: 10.1093/aob/mcr091.
- (387) K. L. Newman, A. G. Fernandez and M. K. Barton, “Regulation of axis determinacy by the *Arabidopsis PINHEAD* gene”, *Plant Cell*, Dec. 2002, **14**, 3029–42, DOI: 10.1105/tpc.005132.
- (388) Z. Zhang, F. Hu, M. W. Sung, C. Shu, C. Castillo-González, H. Koiwa, G. Tang, M. Dickman, P. Li and X. Zhang, “RISC-interacting clearing 3′–5′ exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in *Arabidopsis thaliana*”, *eLife*, May 2017, **6**, DOI: 10.7554/eLife.24466.
- (389) H. Garcia-Ruiz, A. Takeda, E. J. Chapman, C. M. Sullivan, N. Fahlgren, K. J. Brempelis and J. C. Carrington, “*Arabidopsis* RNA-dependent RNA polymerases and Dicer-like proteins in antiviral defense and small interfering RNA biogenesis during *Turnip mosaic virus* infection”, *Plant Cell*, Feb. 2010, **22**, 481–96, DOI: 10.1105/tpc.109.073056.
- (390) F. Qu, X. Ye and T. J. Morris, “*Arabidopsis* DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1”, *Proc Natl Acad Sci U S A*, Sept. 2008, **105**, 14732–7, DOI: 10.1073/pnas.0805760105.

- (391) M. Cao, P. Du, X. Wang, Y. Q. Yu, Y. H. Qiu, W. Li, A. Gal-On, C. Zhou, Y. Li and S. W. Ding, “Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*”, *Proc Natl Acad Sci U S A*, Oct. 2014, **111**, 14613–8, DOI: 10.1073/pnas.1407131111.
- (392) C. Brosseau and P. Moffett, “Functional and genetic analysis identify a role for *Arabidopsis* ARGONAUTE5 in antiviral RNA silencing”, *Plant Cell*, June 2015, **27**, 1742–54, DOI: 10.1105/tpc.15.00264.
- (393) K. Nonomura, A. Morohoshi, M. Nakano, M. Eiguchi, A. Miyao, H. Hirochika and N. Kurata, “A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice”, *Plant Cell*, Aug. 2007, **19**, 2583–94, DOI: 10.1105/tpc.107.053199.
- (394) J. J. Harvey, M. G. Lewsey, K. Patel, J. Westwood, S. Heimstädt, J. P. Carr and D. C. Baulcombe, “An antiviral defense role of AGO2 in plants”, *PLOS One*, Jan. 2011, **6**, e14639, DOI: 10.1371/journal.pone.0014639.
- (395) T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing and Jupyter Development Team, “Jupyter Notebooks—a publishing format for reproducible computational workflows”, in *Positioning and power in academic publishing: Players, agents and agendas*, ed. F. Loizides and B. Schmidt, 2016, pp. 87–90, DOI: 10.3233/978-1-61499-649-1-87.
- (396) G. Lobet, “Image analysis in plant sciences: Publish then perish”, *Trends Plant Sci*, May 2017, DOI: 10.1016/j.tplants.2017.05.002.