THE EVOLUTION OF THE MIR156/157 PATHWAY IN WOODY LEGUMES

Aaron R. Leichty

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

_____

Scott Poethig, Professor of Biology

Graduate Group Chairperson

_____

Michael Lampson, Associate Professor of Biology

Dissertation Committee

Dustin Brisson, Associate Professor of Biology
John Carlson, Professor of Molecular Genetics, Penn State University
Brian Gregory, Associate Professor of Biology
Doris Wagner, Professor of Biology

ACKNOWLEDGMENT

I am thankful to everyone who has helped me become a scientist. My early professors at Goshen College—Jody Saylor, Doug Schirch, and Dan Smith—taught me how to think critically and introduced me to molecular biology. David Pfennig, Karin Pfennig, and Corbin Jones taught me how to think and approach evolutionary questions and how to conduct field work. Though I didn't come to Penn for plant biology, the small group using plants as model systems helped me find the kingdom of life that most motivates. Brian Gregory, Doris Wagner and Kim Gallagher have all been models for how I want to continue my career in plant science, and have each contributed to my scientific approach. Dustin Brisson has always been generous with his time, even before my arrival at Penn and gave me the best rotation project a graduate student could ask for. John Carlson has traveled many miles to help advise me and I look forward to future miles traveled on collaborations.

There are many things I could say in gratitude towards Scott. Top among them is his almost singular role in inspiring me to start a career in plant science. It shouldn't be unclear to a farm kid whether they like plants or not, but for me it took Scott to make that clear. Scott has also invested a lot of time and resources that in the end will probably disproportionately benefit my career over his own. I will never forget this. He hasn't been the perfect advisor, but he has been as close as I could realistically ask for, and I deeply appreciate his cultivation of a nurturing lab environment. Members of this lab have been excellent colleagues and friends.

I'm also grateful to my family. It was never my parent's intention to raise a scientist. But to their dismay, all the years of working in the wood shop, cleaning the turkey shed, mushroom hunting, trapping, and quilting with the ladies at church, developed a set of skills that have served me well in science. I wouldn't change a thing about my childhood and I am thankful to my sister for her partnership in exploration during these early years.

Finally, Whitney has been a perfect partner during my entire career in science. She has always been willing and able to make the necessary sacrifices, and is an endless source of guidance on the sociology and psychology of science and graduate school. Lawyers are generally under-appreciated. Whitney's parents have also been a great source of support over the years. They're always willing to lend a hand and/or ear.

ABSTRACT

THE EVOLUTION OF THE MIR156/157 PATHWAY IN WOODY LEGUMES

Aaron R. Leichty

Scott Poethig

As plants age, the morphology and physiology of the shoot changes—a phenomenon known as vegetative phase change. Often these changes are important in the life history of the plant and contribute to shifting strategies in growth, defense, and competition. The temporally regulated microRNAs miR156 and miR157 are known to control many aspects of vegetative phase change making them ideal for examining how this pathway contributes to the evolution of life history strategies. In this work, I examine this question using two genera of woody legumes.

First, I use the New World Ant-Acacias (genus *Vachellia*) to examine if age-dependent changes in plant defenses evolve as a consequence of neutral constraints on development. I find that the miR156/157 pathway likely controls the age-dependent appearance of the swollen-thorn syndrome—a suite of traits important for interaction with mutualistic ants. This finding is in disagreement with the idea that the temporal regulation of the swollen-thorn syndrome has evolved as a consequence of selectively neutral processes such as ontogenetic drift and supports the idea that the timing of syndrome emergence is somehow adaptive. I discuss the possible agents of this selection and their relation to the miR156/157 pathway.

I also investigate the frequency of life-history heterochrony and its mechanism using plants of the genus *Acacia*, which undergo a conspicuous shift in leaf morphology during their life cycle. This transition from a compound leaf to a simple-horizontally expanded leaf, known as a phyllode has previously been linked to the miR156/157 pathway. Using this morphological marker of vegetative

phase change I examine the timing of this transition in 147 species within a phylogenetic context. I find that the evolution of a prolonged or persistent juvenile phase has evolved independently at least 7 times in *Acacia*. For two of these events the rate and magnitude of miR156 decline is reduced supporting the idea that these species are neotenous. I use genome sequencing to characterize the *MIR156/MIR157* gene families, and identify a gene with a mutation in a putative *cis*-regulatory element that may contribute to phenotypic differences between species.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

1. Introduction

1.1. Vegetative phase change

The life cycle of plants, like animals, is characterized by a series of developmental events that have significant impacts on survival and reproduction. Onset of reproductive competence and the eventual production of reproductive structures is one of the most important of these events. Plant development is also marked by changes in the morphology and physiology of vegetative structures produced by the shoot apex. These changes are referred to as vegetative phase change (Poethig, 1990) and in many organisms, impact survival, and ultimately whether a plant can reproduce.

Since the earliest descriptions of vegetative phase change (Hildebrand, 1875; Goebel, 1889; Goebel, 1900) a large number of traits have been described with age-dependent changes during development. In many instances the functional significance of these changes have been determined (Table 1.1). Yet, there are few instances where the molecular basis of these important changes in development are known.

The identification of the miR156/miR157 pathway and establishment of a role in controlling traits with age-dependent development in models like *Arabidopsis* and maize, open the door to mechanistic studies in ecological and evolutionary model systems. The traits that have been shown to be functionally important in plant ontogeny can now be studied through the mechanistic lens of miR156/miR157. Below I highlight what is known about this genetic regulatory pathway and speculate about the types of molecular changes that might yield different types of evolutionary events.

Table 1.1 Examples of age-dependent changes in vegetative traits and their functional significance.

Species/Group	Trait	Function	Reference
Eucalyptus nitens	leaf thickness and toughness	defense against mammalian herbivory	Loney et al., 2006
Populus angustifolia	phytochemistry	resistance to aphids	Holeski et al., 2009
Aciphylla aurea	leaf rigidity and spinescence	defense against herbivory	Clark & Burns, 2015
Ilex species	spinescence	defense against herbivory	Crawley, 1983; Supnick, 1983; Obeso, 1997
Eucalyptus species	leaf orientation	light capture for juvenile leaves, light avoidance for adult leaves	James & Bell, 2000
Pseudopanax ferox	orientation and morphology of leaves	defense against herbivory	Atkinson & Greenwood, 1989
Acacia melanoxylon	persistance of compound leaf morphology	water availability	Farrell and Ashton, 1978

1.2. Mechanisms of vegetative phase change

Work over the past twenty years has revealed that the microRNAs miR156 and miR157 are responsible for regulating many of the traits with age-dependent development. This discovery was facilitated primarily by work in *Arabidopsis* and maize (Park et al., 2005; Schwab et al., 2005; Wu and Poethig, 2006; Chuck et al., 2007). In *Arabidopsis*, early work demonstrated that miR156 and miR157 were master regulators of the vegetative phase transition. Overexpression of these miRNAs was capable of producing plants that produced juvenile-like leaves for all of development (Schwab et al., 2005; Wu and Poethig, 2006), and the reduction of miR156/157 via target mimicry constructs was capable of producing plants that entirely skipped production of juvenile leaves (Franco-Zorrilla et al., 2007; Wu et al., 2009).

In *Arabidopsis*, where the developmental dynamics of these microRNAs is best understood, miR156 and miR157 decline in the shoot apex during development, resulting in a base to tip gradient as the plant ages. The reduction of miR156 and miR157 in the apex and leaf primordia releases their protein coding targets, the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes from repression (Wu and Poethig, 2006; Wu et al., 2009, Xu et al., 2016b, He et al., 2018). In *Arabidopsis*, 10 miR156/157-targeted SPL transcription factors play divergent roles in the promotion of adult traits. For example, *SPL9*, *SPL13*, and *SPL15* are largely responsible for the age-dependent regulation of leaf shape and trichome distribution, whereas *SPL3*, *SPL4*, and *SPL5* have no impact on these traits but promote the identity of floral meristems (Xu et al., 2016b). More recent work has shown that the *MIR156* and *MIR157* genes

themselves can be quite divergent in their functions, with only a few genes having major roles in vegetative phase change (He et al., 2018).

These findings pointed to the pathway responsible for measuring developmental age, but did not reveal how age is measured and integrated into the pathway. Goebel originally hypothesized that the switch between juvenile and adult leaves was a consequence of changes in the nutritional status of the plant (Goebel, 1908). By this mechanism, a decrease in the root:shoot ratio during development allows leaves to progress further in their development. This hypotheses was supported by studies demonstrating a link between sugar and leaf complexity (e.g. Allsopp, 1953). More recent work demonstrated that a substance produced by leaves was capable of reducing miR156 levels (Yang et al., 2011) and later work found that sugar was this substance (Yang et al., 2013; Yu et al., 2013). The mechanism by which sugar represses miR156 remains relatively unexplored, but some work suggests that sugar acts on miR156 levels both transcriptionally and post-transcriptionally (Yang et al., 2013; Yu et al., 2013).

Additional work on miR156/157 regulation has also revealed the role of epigenetic silencing in their age-dependent expression pattern. In *Arabidopsis*, work examining regulation of *MIR156A* and *MIR156C* revealed that these genes are repressed by the temporal loss of H3K27ac and deposition of H3K27me3 in their promoters and across their gene bodies, an effect mediated by the PRC2 complex (Pico et al., 2015, Xu et al., 2016a). In addition to silencing, factors have been identified that promote the transcription of *MIR156A/C*. The histone variant H2A.Z promotes overall

levels of *MIR156A/C* expression (Choi et al., 2016; Xu et al., 2018), and chromatin remodelers have been shown to either promote *MIR156A* expression (Xu et al., 2016c) or repress both *MIR156A/C* (Xu et al., 2016a) depending on the remodeling complex.

To date, it still remains unclear how time is recorded and integrated into a readout of *MIR156/157* transcription. Sugar can repress miR156 levels, both transcriptionally and post-transcriptionally—a mechanism that is conserved across plants (Yang et al., 2013; Yu et al., 2013). What remains to be determined is how this mechanism of time keeping (i.e. the progressive increase in sugar concentration) integrates with the epigenetic mechanism acting to repress *MIR156/MIR157* transcription. Maybe more importantly, is sugar concentration the only mechanism of time keeping operating in this pathway? If there are other mechanisms, how do they interact to produce the varied lengths of phase change seen throughout the plant kingdom.

The past decade of work in *Arabidopsis* outline a relatively simple model for how traits can be controlled in an age dependent manner. Yet, despite the extensive diversity in developmental timing seen between plants, there are no clear examples how these differences are mediated at the genetic level. Below, I discuss the potential ways in which the miR156/miR157 pathway might be modified to produce both diversity in the traits that change during development, but also differences in timing between species.

1.3. The evolution of the miR156/miR157 pathway

The miR156/miR157 pathway is highly conserved in plants. The relationship between miR156/miR157 and the *SPL* genes is as old as land plants (Cho et al., 2012) and their role as inhibitors of vegetative phase change are conserved within seed plants (Wang et al., 2011; Chen et al., 2013). Given this conservation, and the diversity of traits that are regulated in an age-depenent manner, it is surprising that no studies have focused on how this pathway contribute to phenotypic differences between populations and/or species.

There are two primary ways in which the miR156/miR157 pathway might contribute to phenotypic diversity. The first is through changes in the timing of development. Altering the timing or magnitude of miR156/miR157 decline is one obvious way to achieve this, particularly by *cis*-regulatory changes at specific *MIR156/MIR157* genes, or by changes to *trans*-acting factors of one or a few *MIR156/157* genes. So far, the *trans*-acting factors discovered in *Arabidopsis* are quite general in their function, with roles regulating many genes, and would likely have strong pleiotrophic effects on other aspects of development. Almost nothing is know about the *cis*-regulation of *MIR156/MIR157* genes, but such changes are generally thought of as less pleiotrophic (Prud'homme et al., 2007). Additionally, loss or gain of *MIR156/157* genes would also be a method to accelerate or prolong developmental transitions. Timing of development could also be altered through post-transcriptional processing of the *MIR156/MIR157* genes or via mechanisms that affect how well they repress their SPL targets.

Phenotypic diversity could also be generated from the miR156/157 pathway by the loss or gain of age-dependent regulation. Traits can have age-dependent regulation in one species, be present from the very beginning of development in another, or be lost entirely. The presence of extra floral nectaries (EFN) in *Acacia* are an example of such diversity, with some species producing fully functional EFN on leaf 1, others requiring weeks of development, and still others never producing them at all (personal observation). Such patterns of diversity could be driven by loss or gain of miR156/157 target sites in *SPL* genes that regulate trait development. A much less pleiotrophic mechanism would be the gain or loss of *SPL* regulation. I believe this is probably a common mechanism for the evolution of age-dependence and as knowledge of SPL-targets increases, this mechanism can be better evaluated.

In this thesis, I attempt to move our understanding away from speculation to actual examples of how the miR156/miR157 pathway evolves and contributes to diversity. To do this, I have utilized two genera of woody legumes, each well known for their ecological and evolutionary significance. What these systems lack in tools and resources, they make up for in extensive phenotypic and genetic diversity. I hope that this work will lay a foundation for years of future study.

2. Development and evolution of the swollen-thorn syndrome in Ant-Acacias

2.1. Abstract

Age-dependent changes in plant defenses against herbivores are widespread. In most cases adaptive and non-adaptive explanations for why plants evolve this life-history strategy have not been determined. Here I use the classic animal-plant mutualism between swollen thorn acacias (genus *Vachellia*) and ants of the genus *Pseudomyrmex* (Belt, 1874; Wheeler, 1942), to evaluate these alternatives. Species that exhibit the swollen thorn syndrome provide ants with refuge and food in the form of swollen stipular spines, protein-lipid rich "Beltian" bodies, and sugar-secreting extrafloral nectaries, in exchange for protection against herbivores, pathogens, and competitors (Janzen, 1966; Gonzalez-Teuber et al., 2014). I show that stipular spine swelling, extrafloral nectaries, and Beltian bodies are produced at a predictable time in shoot development. I further show that the elaboration of these structures is tightly associated with the temporal decline in the microRNAs miR156 and miR157 and a corresponding increase in their targets—the SPL transcription factors—many of which are differentially expressed in leaves and stipules. Shade delays both the decline in miR156/157 and the development of the swollen thorn syndrome, supporting the conclusion that these traits are controlled by the miR156-SPL pathway. Production of extrafloral nectaries by *Vachellia sp.* that do not house ants is also correlated with the decline in miR156/miR157. This latter result suggests that a pre-existing program for the regulation of vegetative phase change was co-opted during the evolution of the swollen thorn syndrome. Along, with genetic evidence from other model systems, these findings support the hypothesis that the age-dependent development of the swollen thorn syndrome is a genetically-regulated trait of adaptive origin rather than a consequence of physiological or developmental genetic constraints on when these traits can develop.

2.2. Background

Plants exhibit a wide diversity of morphological, chemical and behavioral defenses against herbivory. In most plants, these defenses change, appear, or disappear as plants age (Boege and Marquis 2005; Barton and Koricheva, 2010; Massad, 2013; Quintero et al., 2013). Why plants have these shifting patterns of defense remains an open question. Although, changes in defense are often thought to be adaptations to predictable changes in selection pressure during a plant's ontogeny, it is also possible that these patterns have no effect on fitness and are instead a consequence of underlying molecular and physical constraints on development (Barton and Boege, 2017). Tests of this neutral model are rare and difficult.

Selectively neutral changes in plant growth and size are known as 'ontogenetic drift' (Evans, 1972). Given that all plants increase in size as they develop, it is reasonable to assume that these changes might impinge on the types of defenses that can develop at different times in a plant's growth curve. For example, the stem derived domatia of many obligate ant-plants only develop once the stem is of a sufficient diameter (Brouat and McKey, 2001). However, it is also possible that the age-dependent development of specific traits is a consequence of the nature of the genetic regulatory networks necessary for trait development (Kittelmann et al., 2018). For example, the molecular pathway necessary for development of a trait may be incompatible with other pathways present in a particular tissue, or at a particular stage of development. An

understanding of the genetic regulatory networks that control plant defenses is therefore essential for determining how such defenses are constrained by development.

One obvious candidate genetic regulatory network for developmental variation in plant defenses is the miR156-SPL pathway. This network of temporally-regulated microRNAs, miR156 and miR157, and their transcription factor targets, members of the Squamosa promoter-binding Protein Like (SPL) family, are responsible for age-dependent changes of multiple vegetative traits in *Arabidopsis* and maize (Wu and Poethig, 2006; Chuck et al., 2007). miR156/miR157-regulated *SPL* genes promote adult vegetative traits. These genes are repressed early in shoot development by high levels of miR156/miR157, and their abundance increases to varying extents as the levels of miR156 and miR157 decline in successive leaves (Wu and Poethig, 2006; Chuck et al., 2007; Wang et al., 2011).

I examined the role of this developmental regulatory network in a classic model system for the study of the evolution of indirect defenses and symbiotic relationships—the mutualism between Acacia plants (genus *Vachellia*) and ants of the genus *Pseudomyrmex*. In this system, the plants provide food via nectar secretion from nectaries and food bodies on leaves, and shelter in the form of hollow stipular spines at the base of a leaf (the “swollen thorn syndrome”). In exchange for these services, ants protect plants from competitors and herbivores.

50 years ago, Janzen (1967) reported that the swollen thorn syndrome is absent early in the development of *Vachellia cornigera.* However, the generality of this phenomenon and its molecular/physiological basis has never been investigated. Here, I show that the swollen thorn syndrome is temporally-regulated in multiple species of *Vachellia* and that this syndrome is likely under the regulation of the miR156-SPL pathway.

2.3. Methods

2.3.1 Plant material and growth conditions

V. cornigera seed were purchased online from a vendor in Florida (https://www.etsy.com/shop/MrNature). Two varieties of *V. collinsii* were used, one purchased online from a vendor in Belize (http://www.especies-seeds.com; *V. collinsii* Belize), and the other collected by Dan Janzen in Costa Rica (*V. collinsii* Costa Rica). *V. caven* (xDL-89-0115D), *V. constricta* (xDL-90-0431), *V. farnesiana* (xDL-90-0341), *V. pennatula* (xDL-96-0002), and *V. rigidula* (xDL-92-0153D) were all obtained from the Desert Legume Program at the University of Arizona. All plants were germinated in petri dishes on filter paper after clipping the seed coat. Seeds were then transferred to 3” pots containing Fafard#2 potting soil and ¼ teaspoon Osmocote applied on top of the soil. Plants were grown in a Conviron chamber maintained at 24°C, with 16 hrs. light/8 hrs. dark, and 190-220 µmol m^{-2} s^{-1} provided by warm white fluorescent lights. Plants were

periodically inoculated with *Amblyseius cucumeris* to prevent thrip infestation (Greenmethods.com).

Tissue samples for molecular analyses were collected using hand dissection. For *V. collinsii* Belize, leaf and stipule primoridia were removed when they were 1-3 mm in size and before they had become lignified. Samples were pooled for a given node, and consisted of 5-50 samples depending on node and tissue type. For *V. caven*, *V. farnesiana*, and *V. rigidula*, whole apices with leaf primordia less than 3 mm in size were sampled due limited seed supplies. Plant age was calculated from the date of transfer to soil. Samples were collected from 3-4 hours after onset of daylight. The age of first occurrence for SS, EFNs and BB represents the age of the plant when the leaf primordium displaying these structures was between 3-5 mm in size.

The shade treatment was conducted by growing plants under 85% shade cloth in the same chamber as un-shaded plants (above). This treatment reduced measured radiation levels to 20-30 μmol m^{-2} s^{-1} with no observable difference in ambient air temperature. For the leaf ablation experiments, one cotyledon and leaves 1 and 2 were removed at the time of full expansion. Plants were randomly assigned to each treatment group and grown on a single shelf in a walk-in chamber.

2.3.2 Confirmation of species identities

The identity of our samples of *V. collinsii* and *V. cornigera* was confirmed by Dan Janzen, based on morphology. I also used the sequence of the *MATK* chloroplast gene to perform a phylogenetic analysis of species/accessions used in this study. The coding region of this gene was determined for two samples of each accession using Sanger sequencing of PCR products following the methods of (Wojciechowski et al, 2004), with an additional internal primer used for sequencing a problematic region (trnK_3_seq, Table 5.3). Sequences were trimmed to the coding region of *MATK* and aligned using MAFFT (Katoh and Standley, 2013). A maximum likelihood tree was generated using RaxML v8.2 (Stamatakis, 2014) with the GTRGAMMA model of rate heterogeneity using the rapid bootstrapping mode with 1000 searches (Figure 5.1). *Prosopis glandulosa* was used as an outgroup and trimmed from the final tree using the Ape package (Paradis et al, 2004) in R (R Core Team, 2015).

2.3.3 Genome sequencing and identification of *MIR156*/*157* and SPL genes

DNA was extracted from leaves of a single *V. collinsii* Belize plant using the DNeasy Plant kit (Qiagen). A short insert library was prepared using the Illumina Truseq PCR-free kit. Briefly, genomic DNA was fragmented using a Biorupter (Diagenode) and size selected for a mean insert size of 450 bp using 1.5% gels on a Pippin Prep (Sage Science), and concentrated by column

purification (Thermo Fisher Scientific GeneJET PCR Purification kit) before beginning the Truseq protocol. The resulting library was sequenced at the University of Pennsylvania Next Generation Sequencing Core, in a single lane of a Hiseq 2500 using 250PE format.

Raw reads were merged using FLASH v1.2.11 with the default settings (Magoc and Salzberg, 2011). Both merged and unmerged reads were used to estimate the genome size using the Perl script, estimate_genome_size.pl (https://github.com/josephryan/estimate_genome_size.pl). The resulting reads were de novo assembled using MaSuRCA (Zimin et al., 2013). Superreads were assembled using a kmer of 127 bp. The MaSuRCA-generated assembly script was modified to allow reads larger than 200 bp.

To identify *MIR156* and *MIR157* containing scaffolds, the *V. collinsii* scaffolds were searched with BLASTn (Altschul et al., 1990; Camacho et al., 2009) for homology with a database of *MIR156* and *MIR157* hairpin sequences from 6 plant species. Data for the six species—*Arabidopsis thaliana*, *Gycine max*, *Malus domestica*, *Medicago truncatula*, *Populus trichocarpa*, *Vitis vinifera*—were obtained from miRBase v19 (Griffiths-Jones et al., 2006). The resulting sequences were checked for the existence of a stemloop structure in a 200-300 bp region centered on the putative miRNA using RNAfold (Lorenz et al., 2011).

Squamosa Protein Like (*SPL*) gene containing scaffolds were identified using tBLASTx with a database of all annotated *A. thaliana SPL* genes (Altschul

et al., 1990; Camacho et al., 2009). These scaffolds were then annotated with MAKER (Cantarel et al., 2008), using SNAP (Korf, 2004) for *ab initio* prediction, and a sampling of mRNA and protein sequences from species within the Fabaceae for homology based prediction. SNAP was run using the precompiled HMM models for *A. thaliana*. A final round of manual annotation was done to merge overlapping contigs producing identical transcripts and contigs corresponding to different haplotypes, to produce a final set of non-redundant *SPL* transcripts. Naming of *SPL* genes was based on the top blast hit between the predicted protein sequence and the *Arabidopsis* SPL family.

2.3.4 qPCR analysis of smRNA and mRNA abundance

Individual alignments of miR156, miR157, miR159, and miR168 sequences downloaded from miRBase v19 (Griffiths-Jones et al., 2006) were used to design primers specific to each gene family (Table 5.3). I utilized these primers with the stem-loop RT primers previously described for qPCR of smRNAs (Varkonyi-Gasic et al., 2007). The highly conserved microRNAs miR159 and miR168 were used as endogenous controls for smRNA qPCR, which were validated using the first (nodes 1-3) and last (nodes 9-12) time points. Total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich). cDNA was synthesized using Invitrogen SuperScript III following the methods of Varkonyi-Gasic et al. (2007). For small RNA qPCR, Platinum Taq (Invitrogen) was used with the Roche universal hydrolysis probe #21, and a three step amplification protocol. For mRNA qPCR and semi-quantitative PCR of *SP*L and

MIR156/*157* transcripts, cDNA was synthesized using Invitrogen SuperScript III and DNase digestion (Qiagen) following the manufacture's protocols. For these samples, a SYBR green master mix (Biotools) with a two-step protocol was used with *ACT2* as an endogenous control. Relative measures of abundance where calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). For semi-quantiative PCR, single gels were run for each gene individually. All gels were image processed identically using automated white balance and equalize adjustments. In cases where a primer set failed to amplify a predicted transcript, at least one additional primer pair was tested to confirm expression.

To validate the manual annotation of *SPL* genes, I cloned partial *SPL-like* transcripts from *V. collinsii* Belize using a pool of 9-12 mm stipule primordia from node 7. This RNA was used for 3' RACE using the SMARTer RACE 5'/3' kit (Clontech) with a degenerate primer designed from an alignment of the SBP-domain of all 10 miR156-targeted *SPL* genes in *A. thaliana* and blunt end cloning of the resulting multi-product PCR pool (Thermo Scientific CloneJET PCR Cloning Kit). Each *SPL-like* sequence was represented by at least 3 independent colonies. Validated genes are noted in Table S2.

2.3.5 Quantification and statistical analysis

Statistical tests were conducted in R (R Core Team). The validity of parametric tests was determined by analyzing the residuals from fitted linear models. When necessary, response variables were log transformed to meet

parametric assumptions. Analysis of covariance (ANCOVA) was used for testing the effect of shade treatment on miR156/157 and *SPL* transcript levels. Smoothed conditional means where calculated with a Loess smoother using ggplot2 (Wickham, 2016).

2.3.6 Data and software availability

The sequence data generated for this chapter have been deposited in GenBank. The genome assembly and sequencing reads are associated with the BioProject PRJNA470667. Gene sequences used for the phylogeny have the following accessions: KX959482-87 and MH324479-88. 3'RACE products of *SPL* transcripts have the following accessions: MH404175-MH404183.

2.4. Results

2.4.1 Development of the swollen thorn syndrome is age-dependent

Previous authors have noted that South American ant-acacias do not produce Beltian bodies (B), enlarged extra-floral nectaries (EFNs) on the petiole and rachis of leaves, and swollen stipular spines (SS) immediately after germination (Figure 2.1A -C) (Hocking, 1970; Janzen, 1967). However, there are no detailed characterizations for the developmental timing of these traits under controlled conditions. Consistent with Janzen's observations of plants in the wild, I found that *V. cornigera* and *V. collinsii* plants grown from seed in either a

greenhouse or growth chamber began producing these traits between node 5–15, and did so in sequence (Figure 2.1). Extrafloral nectaries appeared first, followed by the swelling and elongation of stipules, and then by Beltian bodies (Figure 2.1D). The timing and position of this transition differed between species and accessions, with *V. collinsii* Belize producing swollen stipules first, followed by *V. collinsii* Costa Rica and then *V. cornigera* Florida (Figure 2.1D, E). The development of enlarged stipules was particularly striking in *V. collinsii*, where it occurred within one node and involved a greater that 5-fold increase in stipule length (Figure 2.1C, F). Both of the *V. collinsii* accessions I examined continued to produce swollen stipules at subsequent nodes. However in *V. cornigera*, intermediate sized stipules were occasionally produced during the transition period (Figure 2.1F), and many individuals alternated between producing swollen and unswollen stipules as they developed.

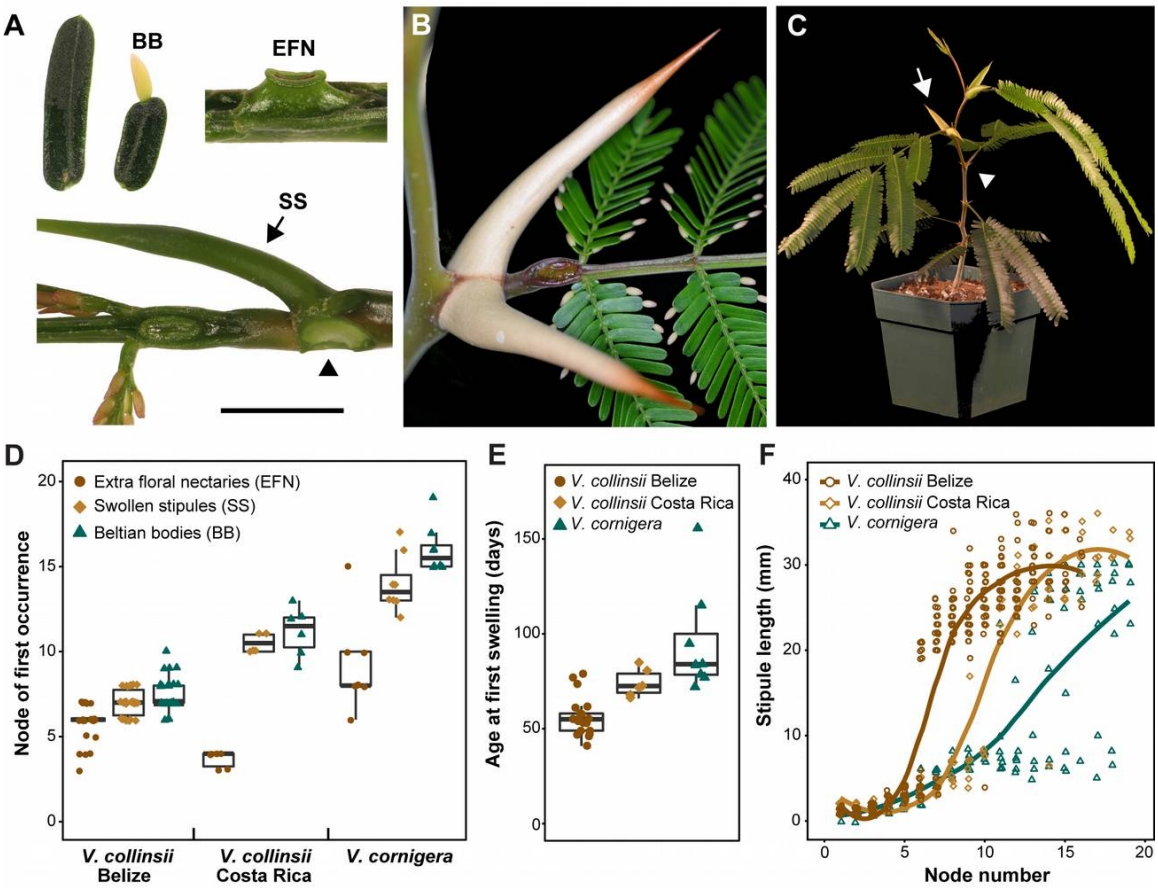


Figure 2.1. Ant-acacias have an early period in their life-cycle without traits necessary for mutualistic interactions. (A) The traits comprising the swollen thorn syndrome. BB: leaflets with and without a Beltian body growing from their tip. EFN: enlarged extrafloral nectary from leaf 8 of *V. collinsii* Belize. SS: late-stage swollen stipule primordia (arrow) showing relative position on leaf (arrowhead marks site of the second removed stipule). Scale bar equals 1 cm. **(B)** Relative position of mature syndrome traits in *V. collinsii* Costa Rica. **(C)** A two-month old *V. collinsii* Belize seedling showing the early transition from unswollen (arrowhead) to swollen (arrow) stipular spines. **(D)** Node of first occurrence for each syndrome trait. Boxes bound 1st and 3rd quartile, center line marks the median. **(E)** Plant age at the time of producing the first fully expanded swollen stipule. Boxes same as in D. **(F)** Stipule length as a function of node number. Curves represent conditional means using a Loess smoother.

2.4.2 The appearance of the swollen thorn syndrome is correlated with a decline in miR156/miR157

Many age-dependent changes in shoot identity are controlled by the microRNAs miR156 and miR157 and their targets Squamosa-promoter-binding Protein-Like (SPLs) transcription factors (Poethig, 2013). miR156/miR157-regulated *SPL* genes promote adult vegetative traits. These genes are repressed early in shoot development by high levels of miR156/miR157, and their expression increases to varying extents as the expression of miR156 and miR157 decline in successive leaves (Wu and Poethig, 2006; Chuck et al., 2007; Wang et al., 2011).

To determine if the development of BB, EFN, and SS might be regulated by this pathway, I used RT-qPCR to measure the abundance of miR156 and miR157 in 1-3 mm leaf and stipule primordia from *V. collinsii* Belize. I found that miR156 levels dropped over 8-fold in leaves and over 19-fold in stipules between node 1 and node 8, and then remained relatively constant after this point (Figure 2.2A & 2.2B). Over the same period miR157 dropped 3.5 and 6 fold in leaf and stipules, respectively, and continued to decline in later nodes (Figure 2.2A & 2.2B). In leaves, the gradual decline in miR156/miR157 was correlated with a gradual increase in the length of EFNs and an increase in the number of BB (Figure 2.2E). Stipule swelling occurred at the nodes with the lowest level of miR156 and miR157 (nodes 6-8) (Figure 2.2F). Conversely, levels of miR156 did not decline, and miR157 only dropped by two-fold in fully expanded leaves

across these same nodes (Figure 2.3A). Thus, miR156/miR157 only display a major change in gene expression at the developmental stage when the morphological fate of the leaf is being specified.

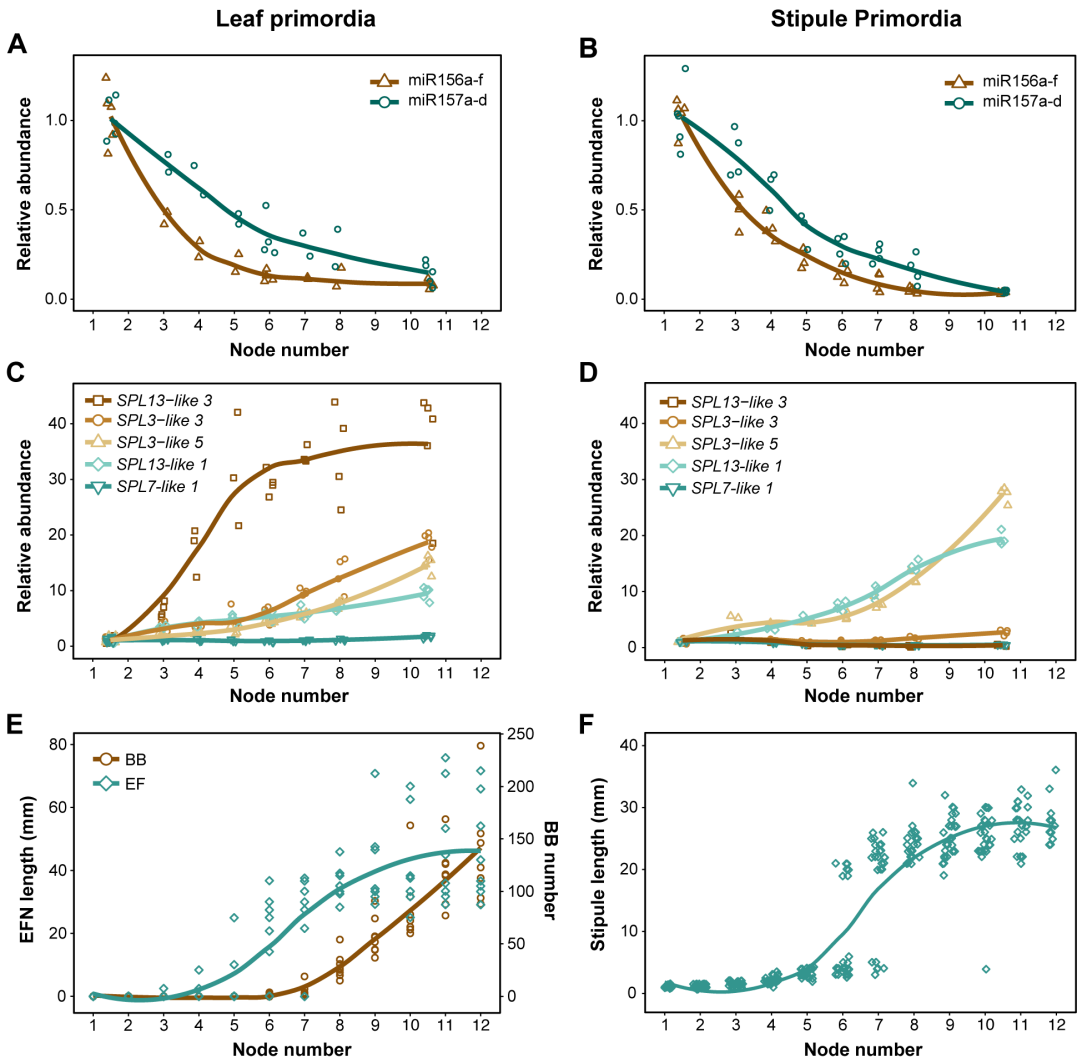


Figure 2.2. Variation in the abundance of miR156/miR157 and *SPL* transcripts is correlated with the appearance of the swollen thorn syndorome in *V. collinsii* Belize. (A,C, E) Leaf primordia, **(B, D, F)** Stipule primordia. Node number represents the position relative to the base of the shoot. **(A-B)** Relative abundance of the mature miR156a-f and miR157a-d small RNAs in leaf and stipule primordia. Curves represent conditional means using a Loess smoother. **(C-D)** Relative abundance of miR156/157-targeted (*SPL3-like 3*, *SPL3-like 5*, *SPL13-like 1*, *SPL13-like 3*) and untargeted (*SPL7-like 1*) transcripts in leaf and stipule primordia. Plotting as in A-B. **(E)** The length of EFNs and the number of BBs on leaves at successive nodes. Plotting as in A-B.

To characterize the molecular basis of this phenomenon in more detail it was necessary to identify the genes that encode miR156, miR157, and SPL transcription factors in *V. collinsii*. For this purpose, I sequenced a 450 bp insert genomic library of *V. collinsii* Belize using Illumina 250PE format. This generated 50X of overlapping reads, which were assembled using MaSuRCA (Zimin et al., 2013). The resulting assembly covered 89% of the 518 Mb genome in 122,266 contigs with an NG50 of 6,528 bp (Table 5.2). Manual annotation of this assembly revealed 8 putative *MIR156* genes, 5 putative *MIR157 genes*, and 23 putative *SPL* genes.

PCR primers were designed to a unique sequence within the predicted coding region of each *SPL* gene, and to the predicted hairpin region of *MIR156* and *MIR157* genes. These primers were then used to measure the abundance of these transcripts in leaf and stipule primordia at nodes 1-2 and 9-12, using semi-quantitative RT-PCR (Figure 2.3B & C). 12 of the 13 *MIR156/MIR157* genes I analyzed produced detectable transcripts, and all of these transcripts were less abundant at nodes 9-12 than at nodes 1-2 in either leaves or stipules, or in both organs. This is consistent with the abundance of the mature miR156 and miR157 transcripts at these nodes. All but one of the 17 *SPL* transcripts with a predicted miR156/157 target site was more abundant at nodes 9-12 than at nodes 1-2. In contrast, *SPL* transcripts that lacked a predicted miR156/157 target-site showed variable patterns of abundance, with the majority being equally abundant at

nodes 1-2 and nodes 9-12, or decreasing in abundance between these positions (Figure 2.3C).

I then performed a more detailed analysis of the expression patterns of four miR156/157-targeted transcripts (*SPL3-like 3*, *SPL3-like 5*, *SPL13-like 1*, and *SPL13-like 3*) and one non-targeted transcript (*SPL7-like 1*), using quantitative RT-PCR. *SPL7-like 1* was expressed at constant level in both leaf and stipule primordia across all sampled nodes (Figure 2.2C & D). However, miR156/157 targeted *SPL* had different patterns of abundance in different organs. *SPL13-like 3* increased nearly 30-fold in leaf primordia between node 1 and node 6, but was expressed at the same level in stipule primordia at different nodes. *SPL3-like 3* was expressed in a similar pattern, but increased less dramatically than *SPL13-like 3* in leaf primordia. *SPL3-like 5* and *SPL13-like 1* increased in both leaf and stipule primordia, but increased more in stipules than in leaves (Figure 2.2C & D). These patterns support the hypothesis that different components of the swollen thorn syndrome are regulated by different *SPL* genes, and that the absence of these traits early in shoot development is attributable to the repression of these genes by miR156 and miR157.

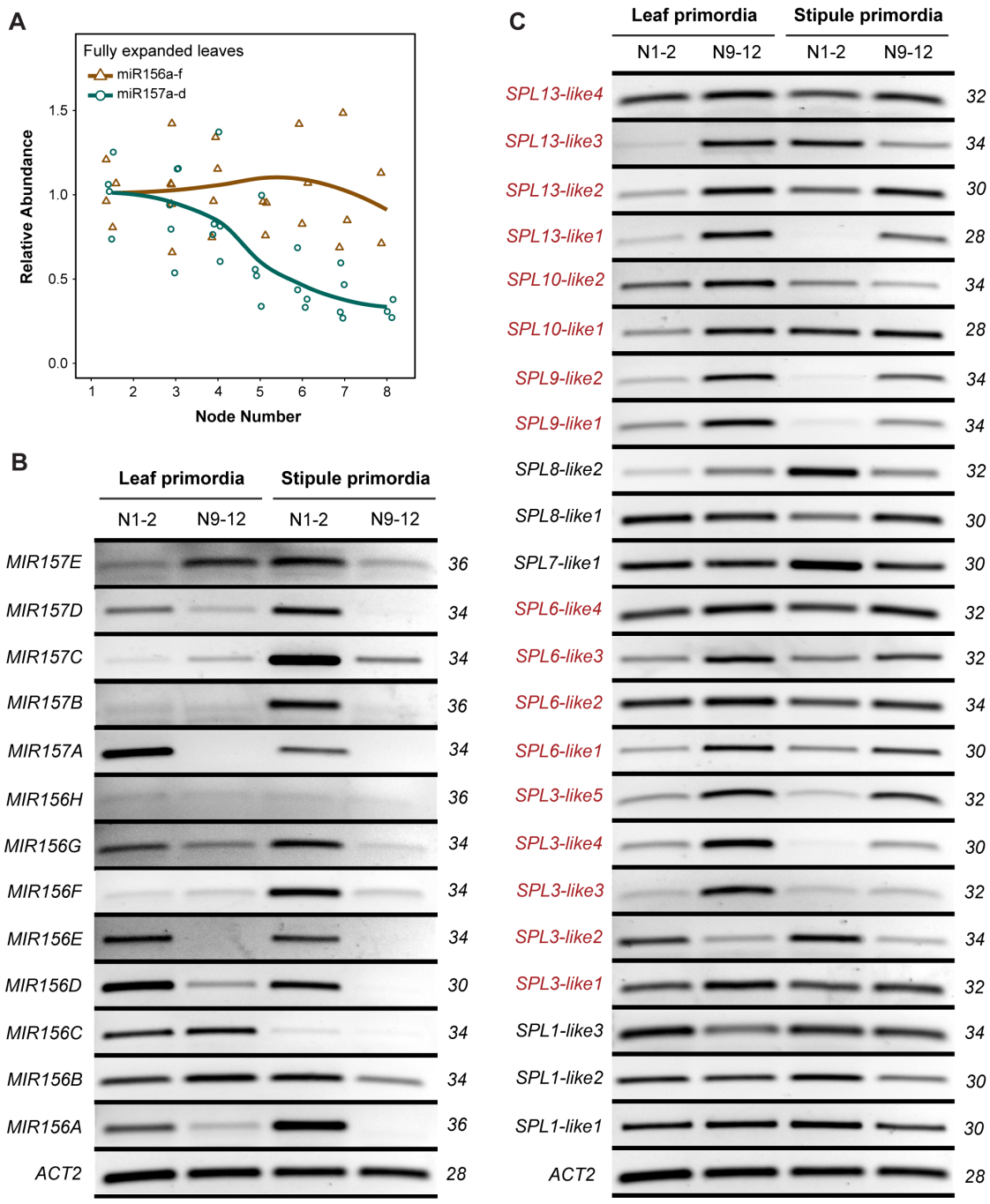


Figure 2.3. Semi-quantitative PCR of *MIR156/157* and *SPL* gene families in *V. collinsii* Belize. (A) Quantitative PCR of miR156 and miR157 in fully expanded leaves. Curves represent conditional means using a Loess smoother. **(B)** Semi-quantitative PCR of *MIR156/157* genes. Leaf and stipule primordia

were sampled at 1-3 mm in size for nodes 1 and 2 (columns 1 and 3) and nodes 9-12 (columns 2 and 4) and pools of 3-5 biological replicates were used for semi-quantitative PCR. Each gene was run on a single gel, allowing for direct comparisons of relative abundance between developmental stage and tissue type. Cycle numbers are on the right. **(C)** Semi-quantitative PCR of *SPL* genes. Red font represents those genes with a predicted miR156/157 target site. *ACT2*, the loading control, is identical between *MIR156/157* and *SPL* gels.

2.4.3 Shade delays the appearance of the swollen thorn syndrome and increases the abundance of miR156/miR157

Genetic analysis of the role of miR156/miR157 and their targets in the development of the swollen thorn syndrome is hindered by the lack of methods for inactivating gene function in *V. collinsii* and related species. As an alternative, I explored the observation that in the wild this syndrome develops more slowly in plants growing under shaded conditions (Janzen, 1967). To confirm this observation, *V. collinsii* (Belize) and *V. collinsii* (Costa Rica) were grown in a single growth chamber under full illumination or under a shade cloth-covered enclosure that reduced the light intensity by 85%. Both provenances produced SS significantly later under shaded conditions than under full illumination (Mann-Whitney U test: $P < 0.001$ and $P < 0.01$, respectively). Shade also reduced the number of BB and the length of the EFN in *V. collinsii* (Belize) (Figure 2.4A-B and Figure 2.5A).

To determine if shade affects the development of SS, BB and EFNs by modulating the activity of the miR156/miR157-SPL pathway, I measured the abundance of miR156 and miR157 in leaf primordia of *V. collinsii* Belize. These miRNAs were expressed in the same temporal pattern in shaded and un-shaded plants, but shaded plants had significantly higher levels of both miRNAs than un-shaded plants (Figure 2.4C & D: ANCOVA, $p < 0.001$ for both). Consistent with this observation, *SPL3-like 3* and *SPL3-like 5* were expressed at significantly lower levels in shaded plants than in unshaded plants (Figure 2.5B-C: ANCOVA,

$p < 0.01$ and $p < 0.001$, respectively). Furthermore, the abundance of these transcripts at the first node to produce BB in shaded plants (10.5 on average) was similar to their abundance at the corresponding note in unshaded plants (node 6.5 on average) (Figure 2.4A and Figure 2.5B & C). Similarly, although shade did not have a statistically significant effect on the abundance of the *SPL13-like 1* or *SPL13-like 3* transcripts their overall levels were lower in shaded plants (Figure 2.5D-E: ANCOVA, $p = 0.98$, $p = 0.08$, respectively). This was not true for *SPL7-like 1,* which had a similar level of abundance in both conditions across development (Figure 2.5F: ANCOVA, $p = 0.29$).

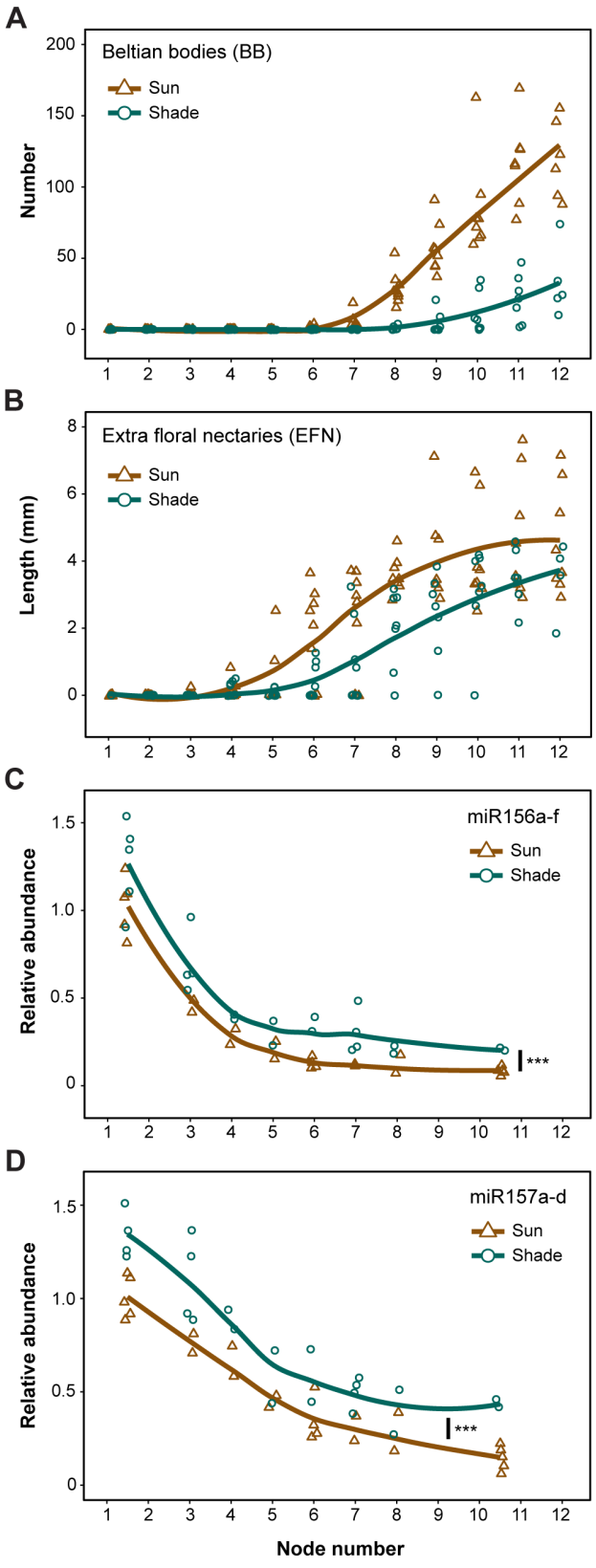

A
Beltian bodies (BB)
Sun
Shade
Number
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12
B
Extra floral nectaries (EFN)
Sun
Shade
Length (mm)
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12
C
miR156a-f
Sun
Shade
Relative abundance
1.5
1.0
0.5
0
***
1 2 3 4 5 6 7 8 9 10 11 12
D
miR157a-d
Sun
Shade
Relative abundance
1.5
1.0
0.5
0
***
1 2 3 4 5 6 7 8 9 10 11 12
Node number

Figure 2.4. Shade delays onset of swollen thorn syndrome and alters the temporal pattern of miR156/157-SPL pathway. (A-B) Number of Beltian bodies and length of EFN per node in shaded and unshaded conditions. Curves represent conditional means using a Loess smoother. **(C-D)** Relative abundance of miR156a-f and miR157a-d in leaf primordia. Differences between treatments were tested by ANCOVA and significance of these tests are indicated, ***p < 0.001. Plotting as in A.

The difference in the responsiveness of *SPL3-like* and *SPL13-like* may be due in part to differences in the mechanism by which they are repressed by miR156/miR157. In *Arabidopsis*, miR156/miR157 regulate *SPL13* expression at a translational level rather than by inducing transcript cleavage, and this may be true for *SPL13-like 1* and *SPL13-like 3* as well (Xu et al., 2016; He et al., 2018). The expression pattern of *SPL* genes is also regulated at a transcriptional level, and our data do not disentangle this level of regulation from the effect of miR156/miR157. Nevertheless, the correlated effect of shade on the development of SS, BB and EFN and on the expression of genes in the miR156/miR157-SPL pathway, support the hypothesis that the swollen thorn syndrome is regulated by this pathway.

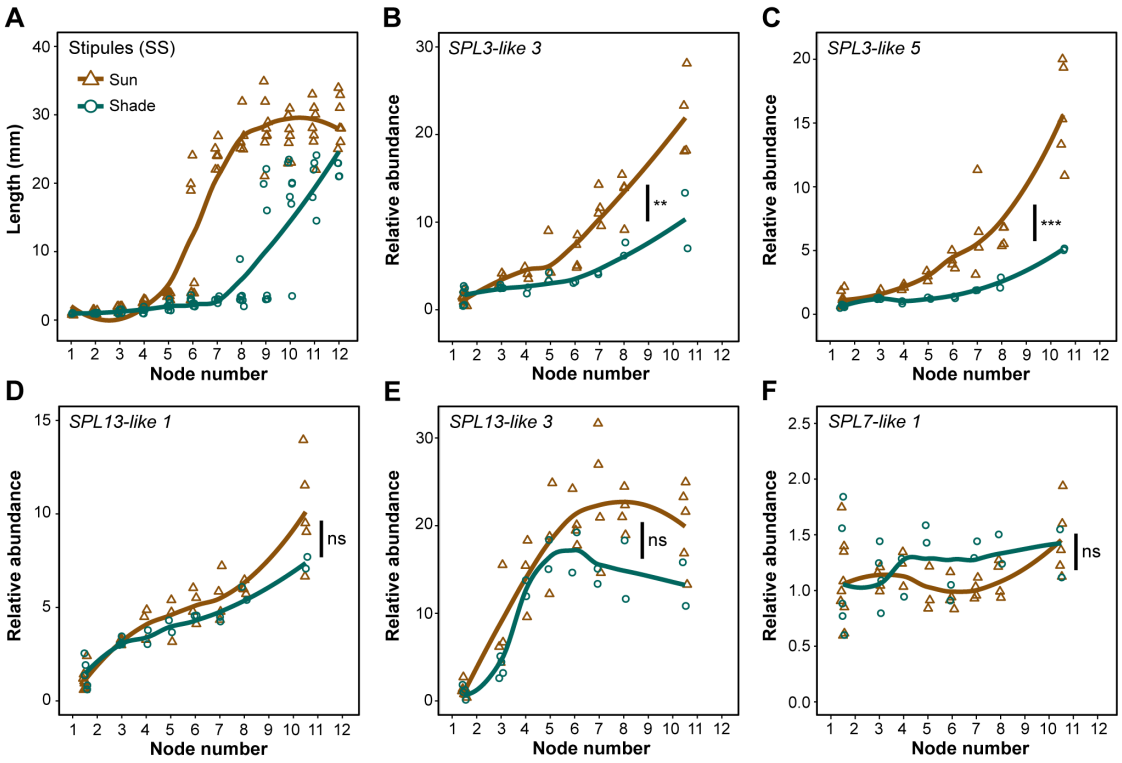


Figure 2.5. The effect of shade and ablation treatments on syndrome development. (A) Stipule length as a function of nodal position under sun and shade conditions. Curves represent conditional means using a Loess smoother. **(B-F)** Relative abundance of *SPL-like* transcripts in sun and shade conditions. Plotting as in A. Differences between treatments were tested by ANCOVA. Significance of these tests are indicated: ** = $p < 0.01$, *** = $p < 0.001$, ns = $p > 0.05$. Plotting as in A.

2.4.4 The swollen thorn syndrome likely evolved by co-opting a pre-existing regulatory pathway

The production of BB is a derived trait within *Vachellia* (Heil et al., 2004; Gomez-Acevedo et al., 2010). However, all species in this genus produce stipular spines and extrafloral nectaries (Figure 2.6A), although both of these structures remain relatively small in species that do not establish a relationship with ants. I reasoned that the swollen thorn syndrome may have evolved by co-opting a pathway that controls the development of these preexisting traits. This hypothesis was supported by our observation that EFN are not initially produced in *Vachellia* species that are closely related to *V. collinsii* (Figure 2.6A) but which do not produce enlarged SS and BB. Indeed, three of these species—*V. caven, V. farnesiana, and V. rigidula*—produced EFN even later than *V. collinsii* Belize (Figure 2.6B).

To determine if development of EFN in these species is correlated with the expression of genes in thet miR156/157-SPL pathway, I measured the abundance of miR156, miR157 and *SPL3-like 5* in shoot apices of *V. caven*, *V. farnesiana* and *V. rigidula* at 4-5 different times after planting. The samples collected at the final time point were taken from plants in which at least one prior leaf had had an EFN. The primers used to measure *SPL3-like 5* were identical to those used for *V. collinsii,* and sequencing of the resulting products revealed that the same gene was amplified in all 3 species. In every species, miR156 and miR157 declined over 4-fold and the expression of *SPL3-like 5* increased 15-20-

fold during the period preceding the production of the first EFN (Figure 2.6C and Figure 2.7A). Additionally, the expression pattern and the relative abundance of the *SPL3-like 5* transcript were correlated with the pattern of EFN development in different species. *V. caven* produced EFN first and had the fastest increase in *SPL3-like 5*, *V. farnesiana* produced EFN slightly later and displayed a slightly slower increase of *SPL3-like 5*, and *V. rigidula* produced EFN last and had the slowest increase of *SPL3-like 5* (Figure 2.6B & D). In contrast, abundance of *SPL7-like 1* showed no consistent pattern between species (Figure 2.7B). These results suggest that the miR156/miR57-SPL pathway coordinates the timing of vegetative development in many *Vachellia* species, and that this pre-existing regulatory network was co-opted during the evolution of the swollen thorn syndrome.

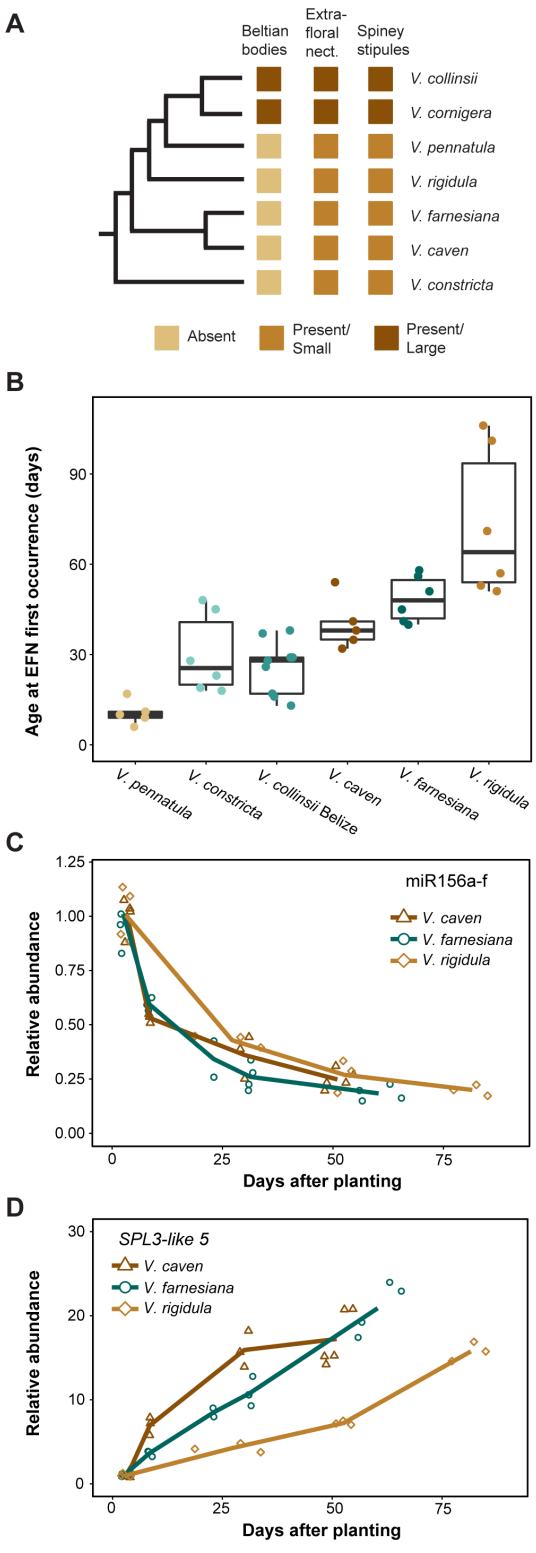

A
Beltian bodies
Extra-floral nect.
Spiney stipules
V. collinsii
V. cornigera
V. pennatula
V. rigidula
V. farnesiana
V. caven
V. constricta
Absent
Present/ Small
Present/ Large
B
Age at EFN first occurrence (days)
0
30
60
90
V. pennatula
V. constricta
V. collinsii Belize
V. caven
V. farnesiana
V. rigidula
C
Relative abundance
0.00
0.25
0.50
0.75
1.00
1.25
miR156a-f
V. caven
V. farnesiana
V. rigidula
0
25
50
75
Days after planting
D
Relative abundance
0
10
20
30
SPL3-like 5
V. caven
V. farnesiana
V. rigidula
0
25
50
75
Days after planting

Figure 2.6. Components of the swollen thorn syndrome show temporal regulation in non-ant-acacia (genus *Vachellia)*. (A) Phylogenetic distribution of syndrome traits for species used in this study. Non-ant-acacias = *V. pennatula*, *V. rigidula*, *V. farnesiana*, *V. caven*, *V. constricta*. **(B)** Timing of EFN first occurrence in *Vachellia* species. Boxes bound 1st and 3rd quartile, center line marks the median. **(C)** Abundance of miR156a-f in three non-ant-acacia species. Lines are plotted through the means of samples grouped by similar age. **(D)** Abundance of *SPL3-like 5* in three non-ant-acacia species. Plotting as in C.

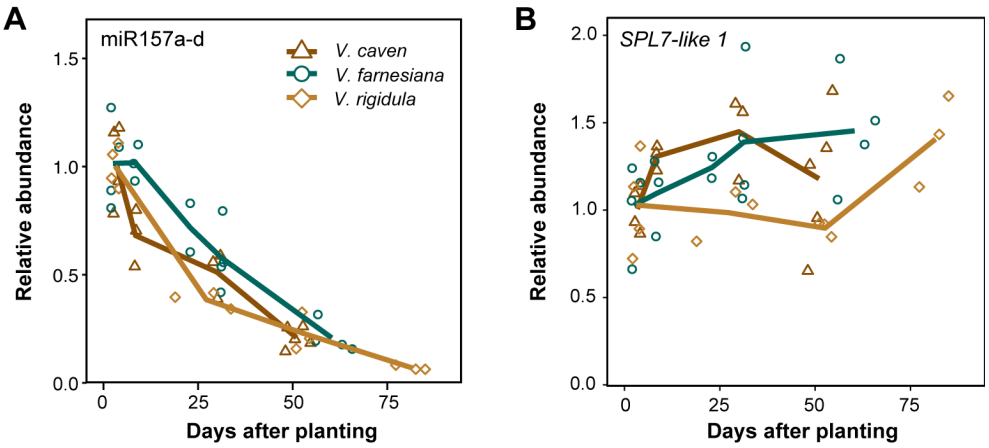


Figure 2.7. qPCR for non-ant-acacia species over development. (A) Relative abundance of miR157a-d as a function of age. Lines are plotted through the means of samples grouped by similar age. **(B)** *SPL7-like 1*. Same as A.

2.5. Discussion

The mutualistic relationship between swollen thorn “Acacias” and ants has been intensively studied for decades and has provided many new insights into fundamental questions in ecology and evolution (Heil and McKey, 2003; Mayer et al., 2014). However, most of this research has focused on the ecological dynamics between ants and plants, and much less is known about the molecular and developmental biology of how this interaction evolved. I am interested in the molecular mechanism of the swollen thorn syndrome. The organs that comprise this syndrome are morphologically and physiologically distinct, yet develop more-or-less simultaneously at a predictable time in shoot development. Because these traits are metabolically expensive but essential to plant survival, it is reasonable to expect that evolution has optimized the timing of this transition (Heil and McKey, 2003; Barton and Boege, 2017). Alternatively, the timing of this transition could be non-adaptive, the result of physical or developmental genetic constraints (Barton and Boege, 2017). Defining the molecular mechanism of the swollen thorn syndrome is therefore an interesting problem in plant development and evolution.

I found that *V. cornigera* and *V. collinsii* plants grown from seed in a controlled environment in the absence of ants begin producing SS, EFNs and BB in a stereotypical sequence at least one month after germination. This, and our evidence that the production of SS, EFNs and BB is tightly correlated with changes in the expression of miR156/miR157 and their SPL targets, strongly

suggest this syndrome is regulated by the vegetative phase change pathway. Our observation that EFNs in species closely related to *V. collinsii* and *V. cornigera* develop at approximately the same age as the swollen thorn syndrome in these ant-acacias, and that the appearance of EFNs is correlated with a change in miR156 and its targets, further suggests that the swollen thorn syndrome is a modification of a developmental pathway that exists in many, if not all species in this genus.

Given its value to the plant, it is interesting to consider why evolution placed the swollen thorn syndrome under temporal regulation. One leading hypothesis for temporal development of plant defenses posits that such patterns are selectively neutral, likely a consequence of developmental constraints related to plant size (Clark and Burns, 2015; Barton and Boege, 2017). This idea predicts that there is a minimum plant size necessary for producing these structures. For this to be correct, plants should be unable to produce syndrome traits immediately after germination. To date, the data on the traits controlled by the miR156-SPL pathway do not support this hypothesis. In *Arabidopsis*, reducing the abundance of miR156/157 using a target mimicry construct transforms the earliest juvenile leaves into adult leaves (Wu et al., 2009). A more natural example comes from the genus *Acacia*, where the adult leaf type, known as a phyllode, can be observed as early as leaf one or two in multiple species, and has been shown to be tightly correlated with the miR156-SPL pathway (Wang et al., 2011). These observations suggest that development of the swollen

thorn syndrome early in shoot development is theoretically possible and would evolve if selection was strong enough.

If the timing of the swollen thorn syndrome is driven by natural selection, I believe tradeoffs in resource allocation between whole plant growth and the production of structures associated with this syndrome is a likely cause (Heil and McKey, 2003; Barton and Boege, 2017). The situation is probably more complicated than this, however, given that *V. cornigera* seedlings do not host ants in their first year of growth, even though they begin to produce the swollen thorn syndrome within 1-2 months (Janzen, 1967). This disconnect may mean that ant foundresses require a critical mass of domatia and/or resources before they colonize a tree, or other temporally regulated factors, such as volatile organic compounds (VOCs) are required for selection. Foundress queens select a host using VOCs as cues of quality (Razo-Belman et al., 2017). These compounds are produced by fully expanded leaves, and our results indicate that miR156/miR157 decline more slowly in fully expanded leaves than they do in leaf and stipule primordia, possibly explaining the delay in colonization.

The evidence that the swollen thorn syndrome is regulated by the miR156/miR157-SPL pathway opens the door to more detailed questions about the mechanism of this phenomenon. For example, nectar secretion from EFNs in New World *Vachellia* species has been shown to depend on jasmonic acid (Heil et al., 2004). This is interesting because previous studies in rice (Hibara et al., 2016) and maize (Beydler et al., 2016) indicate that jasmonic acid promotes the

juvenile phase, possibly via regulation of miR156/157. Either jasmonic acid has the opposite function in New World *Vachellia* species (i.e. promotes the adult phase), or its effect on nectar secretion reflects an organ-specific function of this hormone. It will also be important to explore the functional significance of the observation that different *SPL* genes have different expression patterns in leaves and stipules. Are these genes required for the development of these structures and, if so, are they functionally distinct? This latter question will require the development of methods for manipulating gene expression in *Vachellia*, but it is reasonable to expect that these will become available in the near future.

3. Heterochrony and macroevolutionary patterns of leaf morphology in the genus *Acacia*

3.1. Abstract

Evolutionary change in the timing of developmental events—known as heterochrony—is thought to be a major mechanism of morphological evolution (De Beer, 1951; Gould, 1977; Gould, 1992; Raff and Wray, 1989). Many examples of heterochrony have been recorded, but the molecular mechanisms of these phenomena are largely unknown (Keyte and Smith, 2014). The identification of conserved timekeeping pathways—defined by genes that alter developmental timing when mutated ("heterochronic genes')—now make it possible to more explicitly address this gap in knowledge (Smith, 2003; Geuten and Coenen 2013). I examined the molecular mechanism of heterochrony in plants by examining the expression and the molecular evolution of the miR156/miR157 pathway in a group of neotenous plants in the genus *Acacia*. Previous work has demonstrated that the timing of miR156 and miR157 decline are correlated with a developmental transition in leaf morphology unique to *Acacia* species (Wang et al., 2011). This transition from the compound leaf typical of most legumes to a simple leaf has been lost in two groups of *Acacia*, the sections *Botrycephalae* and *Pulchellae* (Pedley, 1978). Using phylogenomics and growth experiments I demonstrated that these sections represent multiple independent origins of neoteny. Analysis of temporal patterns of miR156 and miR157 expression revealed that species from two different neotenous clades have delayed declines in these regulatory microRNAs. I used genome sequencing to characterize the genes encoding these miRNAs, and found that *MIR156* genes behave differently between normal and neotenous species. One of these genes, *MIR156D*, shows a high degree of conservation in its *cis*-regulatory elements compared with the key *MIR156* genes found in *Arabidopsis*. Phylogenetic shadowing of the *MIR156D* promoter identified a genetic variant in a GA-repeat element that has a near perfect association with neoteny in two independent clades. These findings offer insight into how the deeply conserved regulators of developmental time in plants—miR156 and miR157—contributed to macroevolutionary changes in shoot morphology.

3.2. Background

Gould popularized the concept of heterochrony over 40 years ago (Gould, 1977), and many phenotypic differences between species resulting from changes in developmental timing have been identified since then (Smith, 2003). However, the molecular basis for most of these changes remains unknown (Keyte and Smith, 2014). In animals, this lack of understanding is partly attributable to the fact that mechanisms of developmental timing are poorly conserved between taxa (Keyte and Smith, 2014). In contrast, work over the past 30 years has revealed that—at least in the case of shoot development—the mechanism used by land plants to keep time is broadly conserved (Wang et al., 2011; Cho et al., 2012; Cuperus et al., 2011).

In land plants, the microRNAs miR156 and miR157 target a family of plant specific transcription factors known as the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes. Although their role as inhibitors or promoters of developmental maturation differs at early branches in the land plant phylogeny (Cho et al., 2012), within flowering plants (Wang et al., 2011) and possibly seed plants more broadly (Chen et al., 2013) these small RNAs act to repress the adult vegetative phase of development. In *Arabidopsis*, recent work has revealed that the age-dependent decline in these microRNAs is mediated by the progressive accumulation of histone modifications at a few important *MIR156* genes (Pico et al., 2015; Xu et al., 2016a). This age-dependent repression of *MIR156* (and possibly *MIR157*) genes results in a temporal decrease in the

abundance of miR156 at the shoot apex, allowing its SPL targets to initiate changes in the development of various morphological and physiological traits..

The evidence that the miR156/SPL pathway is conserved across flowering plants opens the door to comparative studies of the evolution of this timing mechanism (Geuten and Coenen 2013). There are a number of ways in which such studies could be conducted. On the one hand, the genetic and genomic resources associated with the family Brassicaceae, of which *Arabidopsis thaliana* is a part, makes this group an obvious choice for research on how variation in the timing of vegetative development has evolved (Mitchell-Olds, 2001). However, *Arabidopsis* and many of its relatives undergo relative minor changes in morphology and physiology during shoot development. This may be due in part to their relatively short life cycle. In contrast, many perennials and woody species have prolonged juvenile phases (Brink, 1962; Poethig,1990) and undergo striking changes during shoot development, making them better models for examining the interplay between the ecological factors influencing changes in developmental timing, the evolution of species differences, and the molecular basis of these phenomena.

To this end, I chose the genus *Acacia*, one of the taxa in which morphologically distint juvenile and adult phases of vegetative development was first described (Goebel, 1990). This group of plants undergoes a conspicuous developmental transition in leaf morphology during development. Plants begin development by producing a pinnately or bipinnate compound leaf, but as the

plant ages it transitions to producing a simple leaf that is expanded in the horizontal plane (Figure 3.1). Although this simple leaf has traditionally been referred to as a phyllode because it was thought to be a modified petiole, this hypothesis is no longer accepted (Kaplan, 1980). Importantly, variation in the expression of miR156 and miR157 is tightly correlated with this developmental transition in leaf morphology (Wang et al., 2011).

The production of phyllodes is a derived trait within the legume subfamily Mimosoideae, and is the ancestral condition in *Acacia* (Murphy et al., 2003). However, some species of *Acacia* never produce phyllodes, and continue to produce bipinnately compound leaves throughout their entire life cycle (hereafter, non-phyllodinous). Traditionally, such species have been classified into two sections of *Acacia*, the *Botrycephalae* and the *Pulchellae* (Pedley, 1986). Because species within these groups are nested with species that produce phyllodes (hereafter, phyllodinous) on modern DNA phylogenies, these species are considered neotenous (Murphy et al., 2003; Murphy, 2008). Even before the production of molecular phylogenies, Pedley (1986) hypothesized that members of the *Botrycephalae* are neotenous based on the observation that some member of this group, such as *A. latisepala,* sometimes produce phyllodes late in shoot development.

In this work I used the extensive variation in the duration of the juvenile phase in *Acacia* species to examine whether the miR156/miR157 pathway has potentially contributed to these macroevolutionary patterns. Along the way I

identified new instances of neoteny in the *Acacia,* and obtained new insights into how species have modified the miR156/miR157 pathway to generate variation in developmental time.

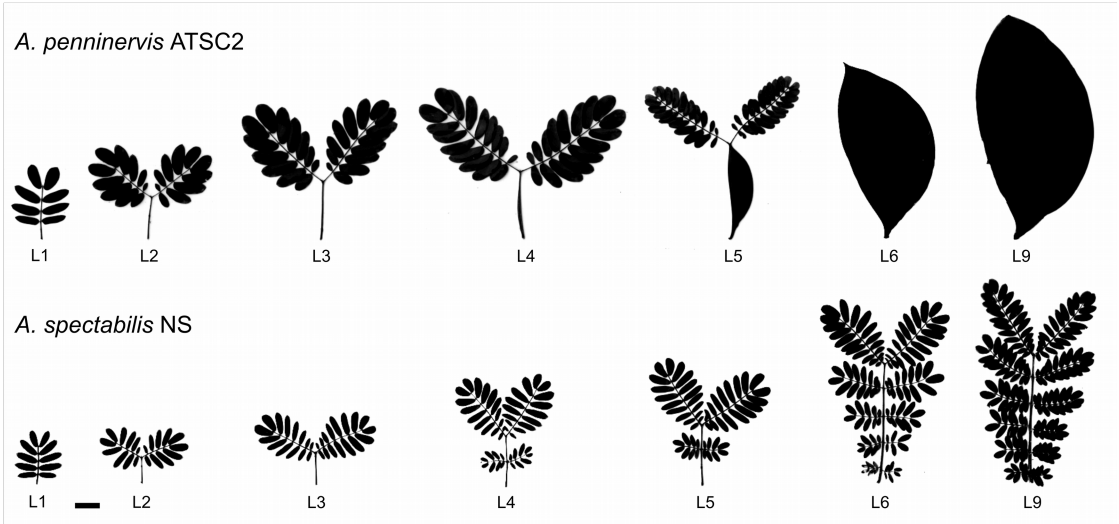


Figure 3.1. Two types of leaf morphology transitions in the genus *Acacia*. Most species of *Acacia* exhibit the pattern illustrated by *A. penninervis* (top row) where the leaf axis shifts from producing a dissected leaf blade in the horizontal plane, to an undissected leaf blade expanded in the vertical plane (hereafter, phyllodinous). Some species exhibit the pattern illustrated by *A. spectabilis*, where leaves maintain the dissected mode of leaf morphogenesis throughout their entire life-cycle (hereafter, non-phyllodinous). Scale bar equals 1 cm.

3.3. Methods

3.3.1 Plant material

For genome sequencing and assembly, leaf tissue from a single individual for a given species and accession was used. A list of these samples is found in Appendix 5.4. For molecular analyses, all plants were grown in a greenhouse with supplemental light. Leaves were collected at full expansion, from 5-7 hours after sunrise. *A. penninervis* plants used for qPCR of shoot apices where grown under conditions identical to plants in Chapter 2. Apices where hand dissected 4-5 hours after onset of daylight when they were 1-3 mm in size. A biological replicate consisted of a pool of 8-30 apices.

3.3.2 Genome sequencing

Short-insert sequencing libraries were prepared using the Illumina Truseq PCR-free kit. Briefly, genomic DNA was fragmented using a Biorupter (Diagenode) and size selected using 1.5% gels on a Pippin Prep (Sage Science), and concentrated by column purification (Thermo Fisher Scientific GeneJET PCR Purification kit) before beginning the Truseq protocol. The resulting libraries were sequenced at the University of Pennsylvania Next Generation Sequencing Core (Philadelphia, PA) on a combination of MiSeq, HiSeq 2500 and HiSeq 4000 machines.

Long insert libraries were prepared at the Huntsman Genomics Core (University of Utah, Salt Lake City, UT) or the Center for Genome Research &

Biocomputing (Oregon State University, Corvallis, OR) using the Illumina Nextera Mate Pair kit. For *A. penninervis* libraries, tagemented genomic DNA was size fractioned on a SageELF (Sage Science) and 6 of the fractions (3, 4.3, 6.6, 8.2,10, 13.5 kb) were subjected to adapter addition. Libraries were sequenced across two lanes of a HiSeq 2500 at 125PE. *A. cultriformis* and *A. spectabilis* libraries were prepared using the gel-free method. Libraries were sequenced on a single lane of a HiSeq 3000 using a 150PE format.

3.3.3 Genome assembly – *P. lophanta*

Raw reads were first merged with FLASH v1.2.11 (Magoc and Salzberg, 2011). Because of low coverage and long read lengths MaSuRCA was used for assembly. The merged reads were used for assembly since MaSuRCA integrates adapter trimming and error correction into its assembly pipeline (Zimin et al., 2013). A kmer of 127 was used for superread assembly and the MaSuRCA assembly script was modified to accommodate reads longer than 200bp.

3.3.4 Genome assembly – *A. mearnsii*

Raw reads were trimmed for adapters with Cutadapt (Martin, 2011) and quality filtered-trimmed using Prinseq v0.20.4 (Schmieder and Edwards, 2011) using a 10 bp window from each end, a 1 bp step size, and a minimum quality score of 20. Additionally, reads were filtered to have a minimum mean quality of 25 across the entire read and a maximum of 1% ambiguous positions. Reads were error corrected using SOAPec with a kmer size of 23 (Li et al., 2010).

Assemblies were generated with both SOAPdenovo (Li et al., 2010) and Platanus (Kajitani et al., 2014).

3.3.5 Genome assembly – *A. rubida*

Raw reads were trimmed for adapters with Cutadapt (Martin, 2011) and quality filtered-trimmed using Prinseq v0.20.4 (Schmieder and Edwards, 2011) using a 15 bp window from each end, a 1 bp step size, and a minimum quality score of 25. Additionally, reads were filtered to have a minimum mean quality of 25 across the entire read and a maximum of 1% ambiguous positions. The resulting reads were then used for error correction using Musket v1.1 (Yongchao et al., 2013). Error corrected reads were then assembled using Platanus v1.2.4 (Kajitani et al., 2014) for contig construction, scaffolding, and gap filling.

3.3.6 Genome assembly – *A. cultriformis*, *A. penninervis*, *A. spectabilis*

Matepair libraries were adapter trimmed using NxTrim (O'Connell et al., 2015) which in addition to trimming the adapter from outward-facing matepair reads, also generates inward-facing paired-end reads, and single-end reads. These reads, and all paired-end reads were then trimmed, quality filtered, error corrected, and assembled as outlined for *A. rubida*.

3.3.7 Preparation of RAD-seq libraries and sequencing

Double digest restriction associated DNA sequencing (ddRAD-seq) libraries were prepared following the methods of Peterson et al. (2012). Briefly, genomic DNA was extracted using the DNeasy Plant Kit (Qiagen) and DNA was eluted into EB buffer (Qiagen). 500 ng of DNA was double digested with HindIII-HF and MfeI-HF (NEB) for 4 hours at 37°C at a volume of 50 ul. Digests were confirmed on 0.5% agarose gels. Half of the digest was used for adapter ligation with 6.5 pmol of each annealed adapter in a total volume of 50 ul. At this stage, 12 samples with unique P1 barcodes were pooled and cleaned using a 1X concentration of AMPure XP beads (Beckman Coulter). One microgram of the resulting pools was size selected on a Pippen Prep (Sage Science) for an insert size of 350 bp +/- 12% using 1.5% gels. Size selected pools were then amplified and indexed in 5-10 PCR reactions using Q5 polymerase (NEB). PCR reactions were then combined and bead-cleaned as above, and eluted into a final volume of 30 ul. Final concentrations were determined using a Qubit fluorometer, and sequencing-pools of 3-5 cleaned PCR pools were combined at equal mass (a total of 36-60 samples). Before Illumina sequencing, libraries were quality checked by blunt end cloning and Sanger sequencing using the CloneJET PCR Cloning Kit (Thermo Scientific). Libraries were sequenced at the University of Pennsylvania Next Generation Sequencing Core (Philadelphia, PA) on a HiSeq 2500 with 100SE format.

3.3.8 Phylogenetic Analyses

Pyrad was used for identifying nucleotide variants in the ddRAD-seq data (Eaton, 2014). Reads were demultiplexed and filtered using default settings. Clustering was performed at 85% similarity and a minimum coverage of 6 for any given locus. A variety of supermatrices were evaluated by changing the maximum number of individuals with a shared heterozygous position (i.e. the paralog filter) and the minimum number of samples for a cluster.

For each supermatrix, maximum likelihood trees were generated using RaxML v8.2 (Stamatakis, 2014) with the GTRGAMMA model of rate heterogeneity using the rapid bootstrapping mode with 100 searches. *Paraserianthes lophanta* was used as an outgroup and trimmed from the final tree using the Ape package in R (Pardis et al., 2004). A full list of samples used in the phylogeny can be found in Appendix 5.7.

3.3.9 qPCR analysis of smRNA and mRNA abundance

Total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich). cDNA was synthesized using Invitrogen SuperScript III following the manufactures specifications. For smRNA-qPCR, Platinum Taq (Invitrogen) was used with the Roche universal hydrolysis probe #21, and a three step amplification protocol. For mRNA-qPCR, SYBR Green was used with a three step amplification protocol (Bimak.com).

Determination of relative discrimination of different miR156 and miR157 assays was done using synthetic RNA oligos (IDT). Reverse transcription of these reactions was conducted as above, but with the addition of 1 ug of total RNA from *E. coli*.

3.3.10 Characterization of *MIR156*/*157* and *SPL* gene families

Identification and annotation of *MIR156/157* and *SPL* gene containing scaffolds in *A. penninervis* and *A. spectabilis* was done as described previously for *V. collinsii* (Methods, Chapter 2).

3.3.11 Analysis of *MIR156D* sequence variation

Regions upstream of the mature miR156 sequence in *MIR156D* were sequenced and cloned from a sampling of species using the primers X and Y. PCR products were column purified and blunt-end cloned using the CloneJET PCR Cloning Kit (Thermo Scientific). At least 4 independent colonies were Sanger sequenced for each individual.

3.3.12 Analysis of *MIR156*/*157* gene orthology

I used microsynteny analysis to determine orthology between *MIR156* and *MIR157* genes in *Arabidopsis thaliana* and *Medicago truncatula*. *M. truncatula* was used as it is the closest relative of *Acacia* with a high quality annotated genome. Three protein coding genes on each side of a given *MIR156/157* gene

were used with BLASTp against flanking genes from all *MIR156/157* genes in a given species. An e-value cutoff of 1e-5 was set for assigning homology.

3.4. Results

3.4.1 Neoteny is frequent in *Acacia*

To characterize the duration of the juvenile phase within *Acacia*, I grew 133 accessions representing 106 named phyllodinous species in a single growth chamber. The length of the juvenile (i.e. bipinnate) phase was normally distributed, although there is a long tail at the upper end of the distribution for species that have a longer than average juvenile phase (Figure 3.2). The average length of the juvenile phase was 49.4 days (median of 39.5), and ranged from 13 days in *Acacia confusa* to 195 days in *A. rubida* NS2.

To understand how the length of the juvenile phase evolved, I used a type of reduced representation sequencing known as double digest restriction-site associated sequencing (ddRAD-seq) (Peterson et al. 2012) to obtain sequence data on 203 individuals representing 150 species. This generated an average of 6.3 million raw reads (5.9 million after filtering), per species. Reads were clustered within an individual using a similarity threshold of 85%, which yielded an average of 37,646 clusters with a minimum depth of 6 and an average depth of 89. Heterozygosity for these clusters averaged 2.3% with an error rate of 0.12%. Clustering across individuals resulted in an average of 32,426 loci that

passed both the minimum coverage threshold and paralog filter. These loci represent an average of over 2.71 Mb per individual.

Using these data I created 10 supermatrices of concatenated loci using a different minimum number of species at a locus ("min") and the maximum number of species with the same heterozygous position in a locus ("maxSH"), also known as the paralog filter. Increasing the minimum number of species decreases the likelihood of clustering due to the random drop out of loci in some individuals, although the number of shared loci is also expected to decrease with phylogenetic distance. Decreasing the paralog filter reduces the number of closely related paralogs in the dataset, but also might act to eliminate informative loci if many closely related individuals are used. I therefore tried a range of different minimum species and two different maxSH cutoffs. These supermatrices ranged from over 23,000 loci with 392,375 parsimony informative sites (PIS) in the min=30-maxSH=6 dataset, to 672 loci with 12,799 PIS in the min=150-maxSH=3 dataset (Table 3.1).

The phylogenies produced from these supermatrices were topologically similar, with all of the major clades maintaining identical relationships. Using the phylogeny from the min90-maxSH3 supermatrix, I mapped the duration of the juvenile phase onto the phylogeny and reconstructed ancestral states using a maximum likelihood framework (Revell, 2012; Revell, 2013). Species of *Acacia* that do not produce phyllodes were arbitrarily assigned a juvenile phase length of 1 year. I chose this value because coding these species with a larger value would

dilute our ability to observe smaller shifts along the phylogeny. Species that were previously described as having a long juvenile phase switched within one year, suggesting that the absence of phyllodes at this time is indicative of a permanent bipinnate condition.

Character mapping revealed at least 7 instances of neoteny. I classified a clade or species as neotenous if its juvenile phase was longer than 1 standard deviation from the mean of all species, and the tips of the branch were colored more strongly than reconstructed internal branches. Many of these events were known previously. For example, species with persistent juvenile foliage have been classified as either belonging to section *Botrycephalae* or *Pulchellae* based on inflorescence morphology (Pedley, 1986). *Botrycephalae* is known to be polyphyletic (Murphy et al., 2003; Miller et al., 2003; Brown et al., 2006), which is confirmed by our phylogeny, but the situation in the *Pulchellae* is unclear (Murphy, 2008). The clustering of *A. gilbertii*—a non-phyllodinous speciemember of the *Pulchellae*—with phyllodinous species indicates that the *Pulchellae* is also polyphyletic (Figure 3.3, H_1). Thus, at the very least, neoteny has evolved twice in this group. This is supported across all the trees I evaluated, with 100% bootstrap support for the branch separating two groups of neotenous *Pulchellae* in all trees (Figure 3.4A).

Another interesting finding was the position of *A. rubida* and *A. terminalis* within the tree. Of the species I grew, *A. rubida* had the longest juvenile phase. This is consistent with reports suggesting that *A. rubida* is allied with bipinnate

species, such as *A. latisepela*, which occasionally produce phyllodes (Acacia rubida, 2018). On the other hand, the exact relationship of *A. terminalis* with other members of the *Botrycephalae* have not been established. As predicted, *A. rubida* clusters with *A. latisepala.* However, it also clusters with the non-phyllodinous species *A. terminalis*, *A. elata*, and *A. jonesii* (Figure 3.3, H_2). The fact that *A. rubida* and *A. latisepala*—species that switch very slowly—are grouped with other species that never switch, is strong support for the idea that the mechanism of juvenilized foliage in many *Acacia* species is due to an extreme delay in the timing of phase change.

Support for the monophyly of the *A. rubida-A. terminalis* group in all of our phylogenies was less than that of the polyphyly of the *Pulchellae*. Bootstrap support for this relationship averaged 50% (Figure 3.4A), where in three of the trees (min150-maxSH3, min150-maxSH6, and min120-maxSH6) *A. terminalis* and *A. rubid*a were grouped with other species (Figure 3.4B). However as the number of loci in the dataset increased, the strength of the grouping between *A. rubida* and *A. terminalis* increased to 98-100% in the min30 datasets. Comparison of the bootstrap support on trees produced by maxSH3 and maxSH6 matrices for a given number of species revealed that relaxation of the paralog filter strong disfavored monophyly (Figure 3.4B, e.g. min90-maxSH3 versus min90-maxSH6). This suggests that the monophyly of this group is indeed real, and that uncertainty is introduced by inaccurate clustering of paralogous loci. In the future, our genome assembly of *A. rubida* may be helpful in resolving

which loci support each of the competing hypothesis for the relationships in this group (Section 3.4.6).

Another interesting group is the clade containing *A. neriifolia* (Figure 3.3, H_3). In our growth experiment, all 7 species in this clade required, on average, 142 day before switching, with the exception of one accession of *A. neriifolia* NS1, which only required 54 days. Relative to *A. rubida*, this group was the second slowest switching clade in the phylogeny. Bootstrap support for the monophyly of this clade was strong with all trees having 100% (Figure 3.4A).

Lastly, in agreement with work by Joe Miller (personal communication) I found that the largest group of species in the *Botrycephalae* is divided into two sections, one that groups with *A. podalyriifolia*, and one that doesn't (Figure 3.3, H_4). In particular, non-phyllodinous species such as *A. debilis* and *A. spectabilis*, group more strongly with phyllodinous species such as *A. podalyriifolia* and *A. jucunda* than with other non-phyllodinous species such as *A. mearnsii* and *A. dealbata*. Support for the polyphyly of this group of neotenous species was strong, with all but one tree having 100% bootstrap values for the branch dividing non-phyllodinous species into two clades (Figure 3.4A). This finding is exciting as it suggests that the *A. podalyriifolia* group is a reversion to phyllode production.

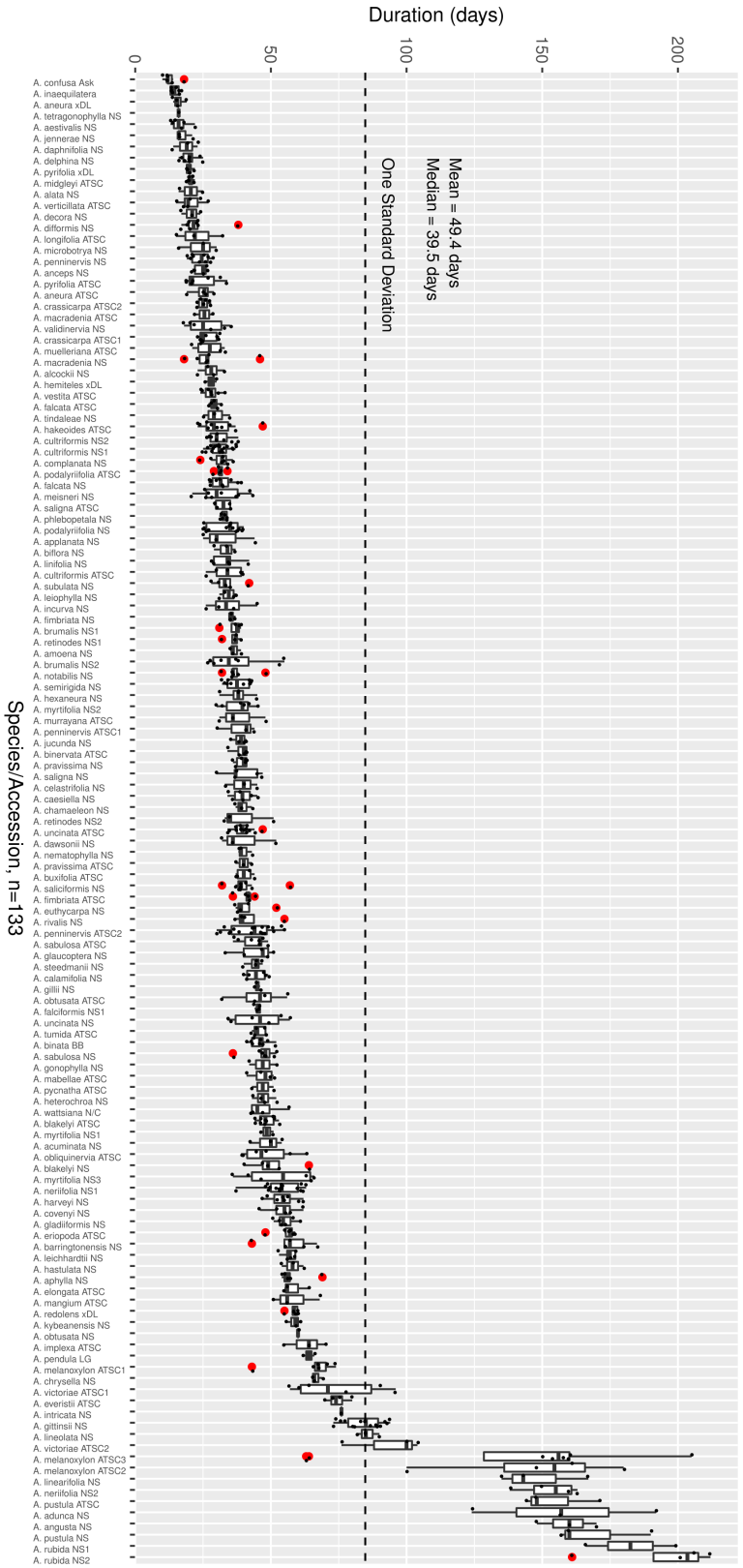

Juvenile Phase Length in Acacia
Duration (days)
0
50
100
150
200
One Standard Deviation
Mean = 49.4 days
Median = 39.5 days
Species/Accession, n=133
A. confusa Ask
A. inaequilatera
A. aneura xDL
A. tetragonophylla NS
A. aestivalis NS
A. jennerae NS
A. daphnifolia NS
A. delphina NS
A. pyrifolia xDL
A. midgleyi ATSC
A. alata NS
A. verticillata ATSC
A. decora NS
A. difformis NS
A. longifolia ATSC
A. microbotrya NS
A. penninervis NS
A. anceps NS
A. pyrifolia ATSC
A. aneura ATSC
A. crassicarpa ATSC2
A. macradenia ATSC
A. validinervia NS
A. crassicarpa ATSC1
A. muelleriana ATSC
A. macradenia NS
A. alcockii NS
A. hemiteles xDL
A. vestita ATSC
A. falcata ATSC
A. tindaleae NS
A. hakeoides ATSC
A. cultriformis NS2
A. cultriformis NS1
A. complanata NS
A. podalyriifolia ATSC
A. falcata NS
A. meisneri NS
A. saligna ATSC
A. phlebopetala NS
A. podalyriifolia NS
A. applanata NS
A. biflora NS
A. linifolia NS
A. cultriformis ATSC
A. subulata NS
A. leiophylla NS
A. incurva NS
A. fimbriata NS
A. brumalis NS1
A. retinodes NS1
A. amoena NS
A. brumalis NS2
A. notabilis NS
A. semirigida NS
A. hexaneura NS
A. myrtifolia NS2
A. murrayana ATSC
A. penninervis ATSC1
A. jucunda NS
A. binervata ATSC
A. pravissima NS
A. saligna NS
A. celastrifolia NS
A. caesiella NS
A. chamaeleon NS
A. retinodes NS2
A. uncinata ATSC
A. dawsonii NS
A. nematophylla NS
A. pravissima ATSC
A. buxifolia ATSC
A. saliciformis NS
A. fimbriata ATSC
A. euthycarpa NS
A. rivalis NS
A. penninervis ATSC2
A. sabulosa ATSC
A. glaucoptera NS
A. steedmanii NS
A. calamifolia NS
A. gillii NS
A. obtusata ATSC
A. falciformis NS1
A. uncinata NS
A. tumida ATSC
A. binata BB
A. sabulosa NS
A. gonophylla NS
A. mabellae ATSC
A. pycnatha ATSC
A. heterochroa NS
A. wattsiana N/C
A. blakelyi ATSC
A. myrtifolia NS1
A. acuminata NS
A. obliquinervia ATSC
A. blakelyi NS
A. myrtifolia NS3
A. neriifolia NS1
A. harveyi NS
A. covenyi NS
A. gladiiformis NS
A. eriopoda ATSC
A. barringtonensis NS
A. leichhardtii NS
A. hastulata NS
A. aphylla NS
A. elongata ATSC
A. mangium ATSC
A. redolens xDL
A. kybeanensis NS
A. obtusata NS
A. implexa ATSC
A. pendula LG
A. melanoxylon ATSC1
A. chrysella NS
A. victoriae ATSC1
A. everistii ATSC
A. intricata NS
A. gittinsii NS
A. lineolata NS
A. victoriae ATSC2
A. melanoxylon ATSC3
A. melanoxylon ATSC2
A. linearifolia NS
A. neriifolia NS2
A. pustula ATSC
A. adunca NS
A. angusta NS
A. pustula NS
A. rubida NS1
A. rubida NS2

Figure 3.2. Length of the juvenile phase in Acacia. 133 different species/accessions of Acacia were grown under common conditions and the time to the production of the first “stable” phyllode was measured. Boxes bound 1st and 3rd quartile, center line marks the median. Outlying points are marked in red.

Table 3.1. Size of supermatrices generated from the ddRAD-seq data.			
Super Matrix	**Number of loci after paralog filter**	**Total variable sites**	**Total parsimony informative sites**
paralog filter, maxSH = 3			
minimum = 30	15877	512073	237176
minimum = 60	5247	185922	87217
minimum = 90	2453	87084	41859
minimum = 120	1186	43016	21424
minimum = 150	672	25029	12799
paralog filter, maxSH = 6			
minimum = 30	23691	819753	392375
minimum = 60	9509	371449	182754
minimum = 90	4912	198715	99968
minimum = 120	2618	109838	57002
minimum = 150	1519	65384	34887

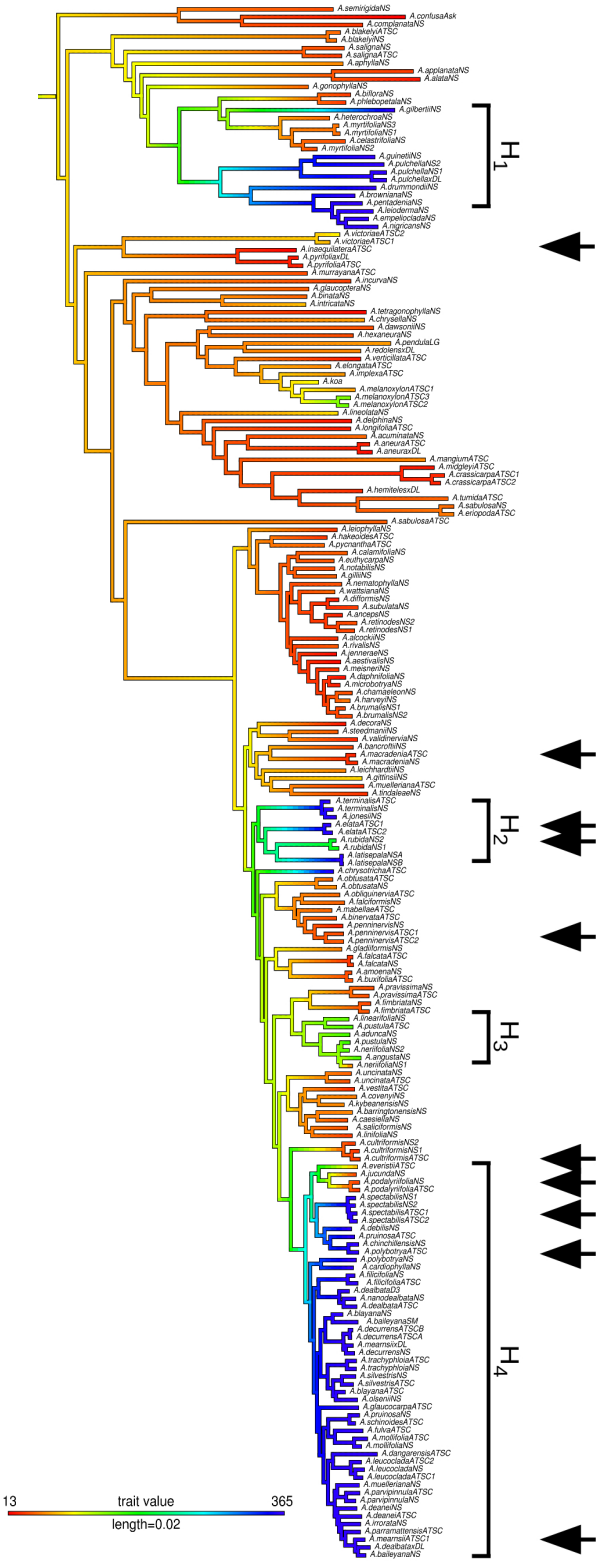

A.semirigidaNS
A.confusaAsk
A.complanataNS
A.blakelyiATSC
A.blakelyiNS
A.saligna NS
A.salignaATSC
A.aphyllaNS
A.applanataNS
A.alataNS
A.gonophyllaNS
A.biflora NS
A.phlebopetalaNS
A.gilbertiiNS
A.heterochroaNS
A.myrtifoliaNS3
A.myrtifoliaNS1
A.celastrifoliaNS
A.myrtifoliaNS2
A.guinetiiNS
A.pulchellaNS2
A.pulchellaNS1
A.pulchellaxDL
A.drummondiiNS
A.brownianaNS
A.pentadenaNS
A.leiodermaNS
A.ampelocladaNS
A.nigricansNS
A.victoriaeATSC2
A.victoriaeATSC1
A.inaequilateraATSC
A.pynfoliaxDL
A.pynfoliaATSC
A.murrayanaATSC
A.incurvaNS
A.glaucopteraNS
A.binataNS
A.intricataNS
A.tetragonophyllaNS
A.chysellaNS
A.dawsoniiNS
A.hexaneuraNS
A.pendulaLG
A.redolensDL
A.verticillataATSC
A.elongataATSC
A.implexaATSC
A.koa
A.melanoxylonATSC1
A.melanoxylonATSC3
A.melanoxylonATSC2
A.lineolataNS
A.delphinaNS
A.tonyfoliaATSC
A.acuminataNS
A.aneuraATSC
A.aneuraxDL
A.mangiumATSC
A.midgleyiATSC
A.crassicarpaATSC1
A.crassicarpaATSC2
A.hemiteles xDL
A.tumidaATSC
A.sabulosaNS
A.eriopodaATSC
A.sabulosaATSC
A.leiophyllaNS
A.hakeoidesATSC
A.lycnanthaATSC
A.calamifoliaNS
A.euthycarpaNS
A.notabilisNS
A.gilliiNS
A.rematophyllaNS
A.wattsianaNS
A.difformisNS
A.subulataNS
A.sneepsNS
A.retinodesNS2
A.retinodesNS1
A.alcockiiNS
A.rivalisNS
A.iennereaNS
A.aestivalisNS
A.mesnerNS
A.daphniifoliaNS
A.microbotryaNS
A.chamaeleonNS
A.harveyiNS
A.brumalisNS1
A.brumalisNS2
A.decoraNS
A.steedmaniiNS
A.validinerviaNS
A.bancroftiiNS
A.macradeniaATSC
A.macradeniaNS
A.leichhardtiiNS
A.gittinsiiNS
A.muellerianaATSC
A.tryptaleaeNS
A.terminalisATSC
A.terminalisNS
A.jonesiiNS
A.elataATSC1
A.elataATSC2
A.rubidaNS2
A.rubidaNS1
A.latisepalaNSA
A.latisepalaNSB
A.chrysotrichaATSC
A.obtusataATSC
A.obtusataNS
A.obliquinerviaATSC
A.falciformisNS
A.trabellataATSC
A.binervataATSC
A.penninerviaNS
A.penninerviaATSC1
A.penninerviaATSC2
A.gladiiformisNS
A.falcataATSC
A.falcataNS
A.amoenaNS
A.buxifoliaATSC
A.pravissimaNS
A.pravissimaATSC
A.timbriataNS
A.timbriataATSC
A.linearifoliaNS
A.pustulaATSC
A.aduncaNS
A.pustulaNS
A.nerifoliaNS2
A.angustaNS
A.nerifoliaNS1
A.uncinataNS
A.uncinataATSC
A.vestitaATSC
A.covenyiNS
A.kybeanensisNS
A.barringtonensisNS
A.caesiellaNS
A.salicifoliaNS
A.linifoliaNS
A.cultriformisNS2
A.cultriformisNS1
A.cultriformisATSC
A.everistiiATSC
A.ucunditaNS
A.podalyriifoliaNS
A.podalyriifoliaATSC
A.spectabilisNS1
A.spectabilisNS2
A.spectabilisATSC1
A.spectabilisATSC2
A.stebisiaNS
A.pruinosaATSC
A.chinchillensisNS
A.polybotryaATSC
A.polybotryaNS
A.cardiophyllaNS
A.lilicifoliaNS
A.lilicifoliaATSC
A.dealbataD3
A.nanodealbataNS
A.dealbataATSC
A.baiyannaNS
A.baileyanaSM
A.decurrensATSCB
A.decurrensATSCA
A.mearnsiixDL
A.decurrensNS
A.trachyphloiaATSC
A.trachyphloiaNS
A.silvestrisNS
A.silvestrisATSC
A.bjayanaATSC
A.obsenNS
A.glaucocarpaATSC
A.pruinosaNS
A.schinoidesATSC
A.fulvaATSC
A.mollifoliaATSC
A.mollifoliaNS
A.dangarensisATSC
A.leucocladaATSC2
A.leucocladaNS
A.leucocladaATSC1
A.muellerianaNS
A.parvipinnulaATSC
A.parvipinnulaNS
A.deanei NS
A.deaneiATSC
A.irrorataNS
A.parramattensisATSC
A.mearnsiiATSC1
A.dealbataxDL
A.baileyanaNS
H1
H2
H3
H4
13
trait value
365
length=0.02

Figure 3.3. Evolution of vegetative phase change in *Acacia*. A ddRAD-seq phylogeny of 200 individuals, representing 147 species. The color of branch tips represent the average length of the juvenile phase for the species. Color of internal branches was estimated using maximum likelihood. Species of *Acacia* that do not produce phyllodes were given a value of 365 days for the duration of the juvenile phase. Brackets with "H_1-H_4" represent clades where support for alternative evolutionary hypothesis was examined (Section 3.4.1). Arrows represent species used in Figures 3.3 and 3.4.

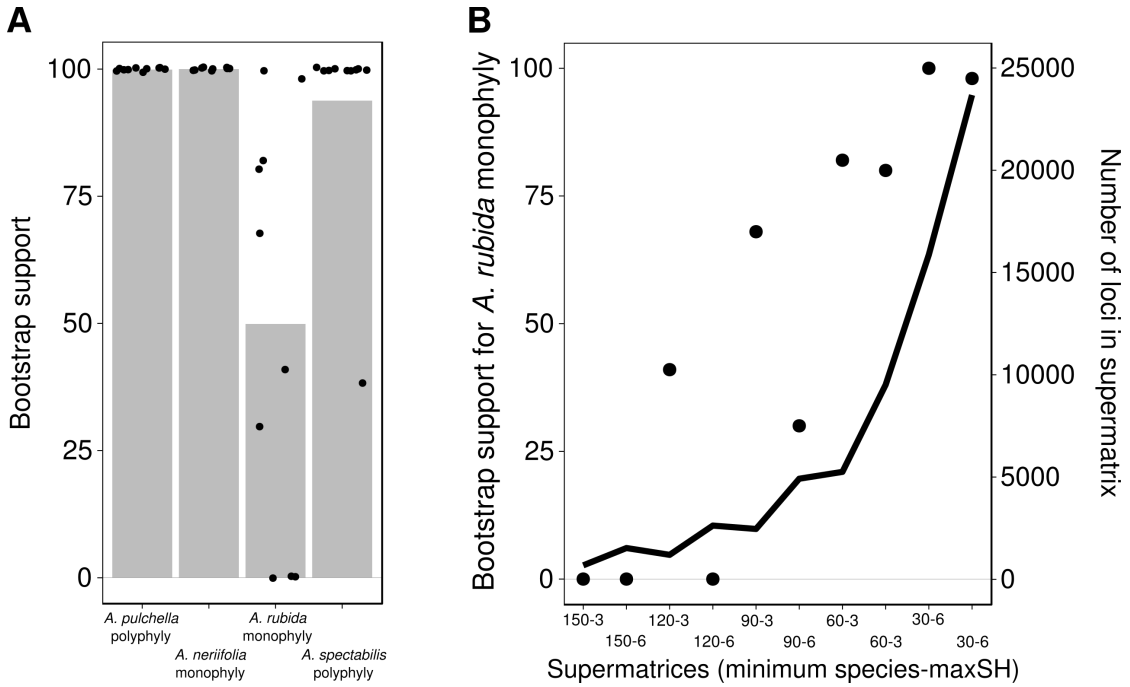


Figure 3.4. Analysis of support for key neotenous clades in Figure 3.3. (A) The average bootstrap support of 4 clades across 10 phylogenies with different levels of constraint on loci in supermatrix. "*A. pulchella* polyphyly" refers to the clade containing both *A. pulchella* and *A. gilbertii* with two independent origins of neoteny. "*A. neriifolia* monophyly" refers to the group of 7 species containing *A. neriifolia*. "A. rubida monophyly" refers to the possible single origin of neoteny in the clade containing *A. rubida* and *A. terminalis*. "*A. spectabilis* polyphyly" refers to the clade where *A. spectabilis* and *A. podalyriifolia* species group more closely than with other neotenous species such as *A. mearnsii*. **(B)** The relationship between degree of bootstrap support for the monophyly of *A. rubida-A.terminalis* species complex (dots on graph, left axis) and increasing number of loci in super matrix (line on graph, right axis), or stringency of paralog filter (e.g. line on graph, 90-3 vs 90-6).

3.4.2 miR156 levels in fully expanded leaves are associated with neoteny

Given previous evidence that a decrease in the abundance of miR156 and miR157 is strongly correlated with the switch from juvenile to adult leaves in *Acacia* (Wang et al., 2011), I decided to examine the role of these miRNAs in the neotenous events identified in this genus. As a starting point, I measured the abundance of these transcripts over the course of shoot development in fully expanded leaves from 9 different species. I sampled 3 ancestor-like species (*Vachellia farnesiana*, *Vachellia tortillis*, and *Paraserianthes falcata*) to determine how these microRNAs behave in genera within the Mimosoideae that do not possess phyllodes. In *Acacia*, I sampled three species that produce phyllodes within three months after germination (*A. victorieae*, *A. macradenia*, and *A. podalyriifolia*) and three species (*A. spectabilis*, *A. mearnsii*, and *A. elata*) that never produce phyllodes (Figure 3.3, arrows).

In the 3 phyllodinous species I examined, miR156/157 declined from 4-8 fold over the first 3 months of development, and 8-16 fold after 12-16 months (Figure 3.5). This is consistent with a previous study of two additional phyllodinous species (Wang et al., 2011). In ancestor-like species, the abundance of these miRNAs did not change significantly, or declined about 2 fold by 6-12 months. In non-phyllodinous *Acacia* species, miR156/miR157 declined 2-3 fold by 6 months of age, and declined no more than 3 fold by 12-16 months (Figure 3.5). Thus, the expression pattern of miR156/157 in non-phyllodinous

Acacia species more closely resembled the pattern in ancestor-like species than the pattern in phyllodinous *Acacia* species.

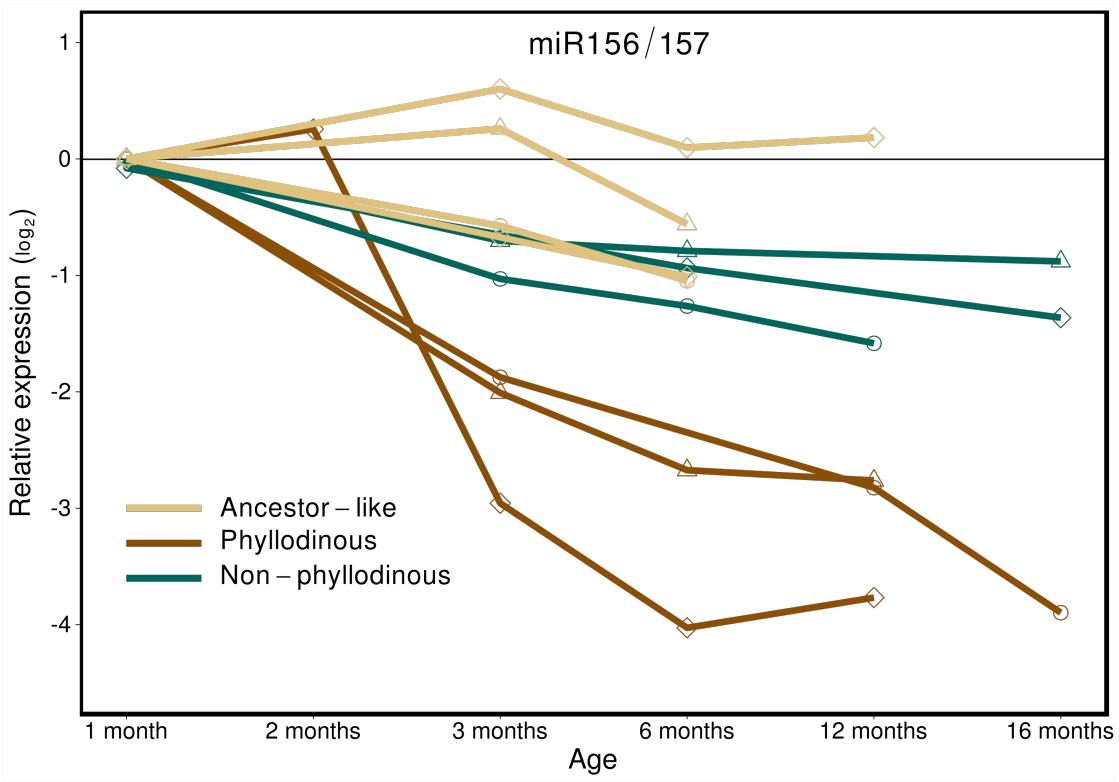


Figure 3.5. Combined levels of miR156 and miR157 in fully expanded leaves during development. Lines represent species means. See text for description of species used.

3.4.2 miR156 and miR157 have different expression patterns during development

The sequence of the mature miR156 and miR157 transcripts differ by two internal nucleotides. To determine if these transcripts can be discriminated by qPCR I tested a set of reverse transcription primers that primed from a variable position near the 3' end of the miRNAs and PCR primers that were specific to either miRNA or the reverse transcription primer. Using synthetic miRNAs our assays showed a high degree of specificity, with less than 1/10 of a percent of the qPCR signal being derived from the opposing miRNA at cycling conditions similar to those used to amplify endogenous levels of these miRNAs (Table 3.2). These assays were also highly discriminant against another miR156 variant (miR156v3) found in some *Acacia* species.

Using these assays, I tested if the difference in the expression pattern of miR156/157 in phyllodinous and non-phyllodinous species was due to miR156, miR157 or both miRNAs (Figure 3.6). During the first 3 months of growth, the level of miR156 declined 16-42 fold in phyllodinous species (*A. penninervis*, *A. macradenia*, *A. cultriformis*). However, in an ancestor-like species (*P. lophantha*), a late-phyllodinous (which had not yet switched to phyllode production at time of sampling) *Acacia* species (*A. rubida*), and three non-phyllodinous *Acacia* species (*A. mearnsii*, *A. spectabilis*, *A. polybotrya*), the level of miR156 did not drop more than 2 fold during this period. In contrast, miR157 behaved similarly in all 8

species. In the phyllodinous specie, miR157 dropped 3.7-13.9 fold, in the non-phyllodinous *Acacia* species, it dropped 2.6-3.7 fold. Thus, the dynamics of miR156 and miR157 expression are significantly different between phyllodinous and non-phyllodinous *Acacia* species (Tukey's HSD, $p < 0.001$ for both miR156 and miR157). In phyllodinous species, miR156 declines more than miR157, whereas in non-phyllodinous species, miR157 declines more than miR156. This difference suggests that miR156 may play a larger role in determining the phenotype of non-phyllodinous species than miR157.

Table 3.2 Relative discrimination of smRNA-qPCR assays.		
	Assay	
miRNA	miR156 F1/R1	miR157 F1/R1
miR156	100	0.001
miR157	0.021	100
miR156v3	0	0

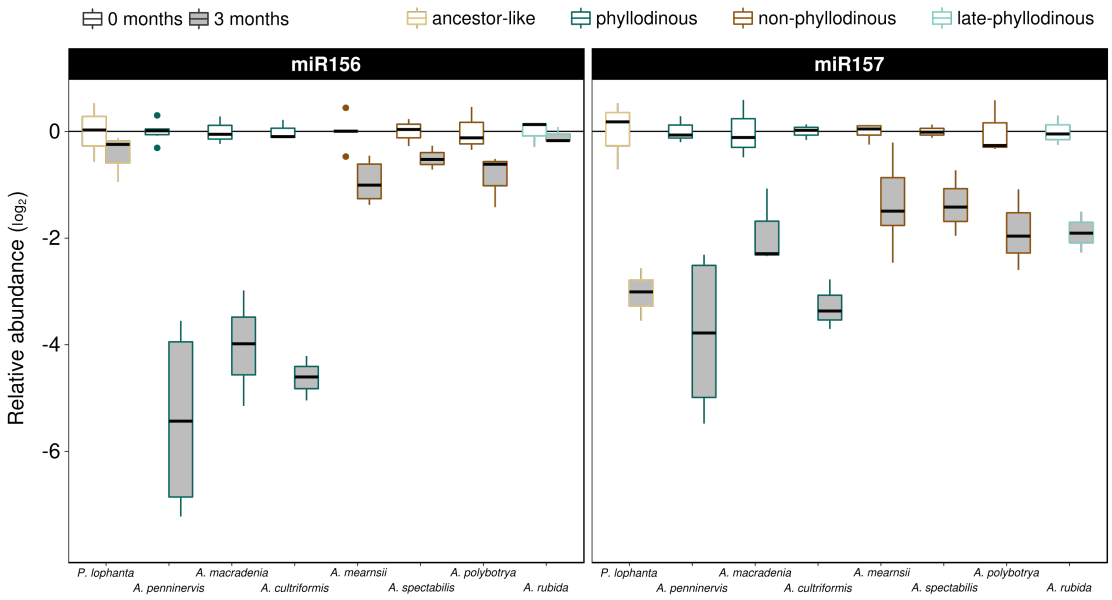


Figure 3.6. Changes in abundance of miR156 and miR157 during first 3 months of development. Boxes bound 1st and 3rd quartile, center line marks the median.

3.4.5 The absolute abundance of miR156 and miR157 is not higher in non-phyllodinous species

The determine if the difference in the morphology of phyllodinous and non-phyllodinous species is due to overall higher levels of miR156 or miR157 in non-phyllodinous species, I measured the abundance of these miRNAs in the first fully expanded leaf of 3 phyllodinous and 5 non-phyllodinous species. There was no consistent difference between these species, and the amounts of miR156 and miR157 was similar to that of miR159 and miR168, the two miRNAs used as endogenous controls for our qPCR (Figure 3.7).

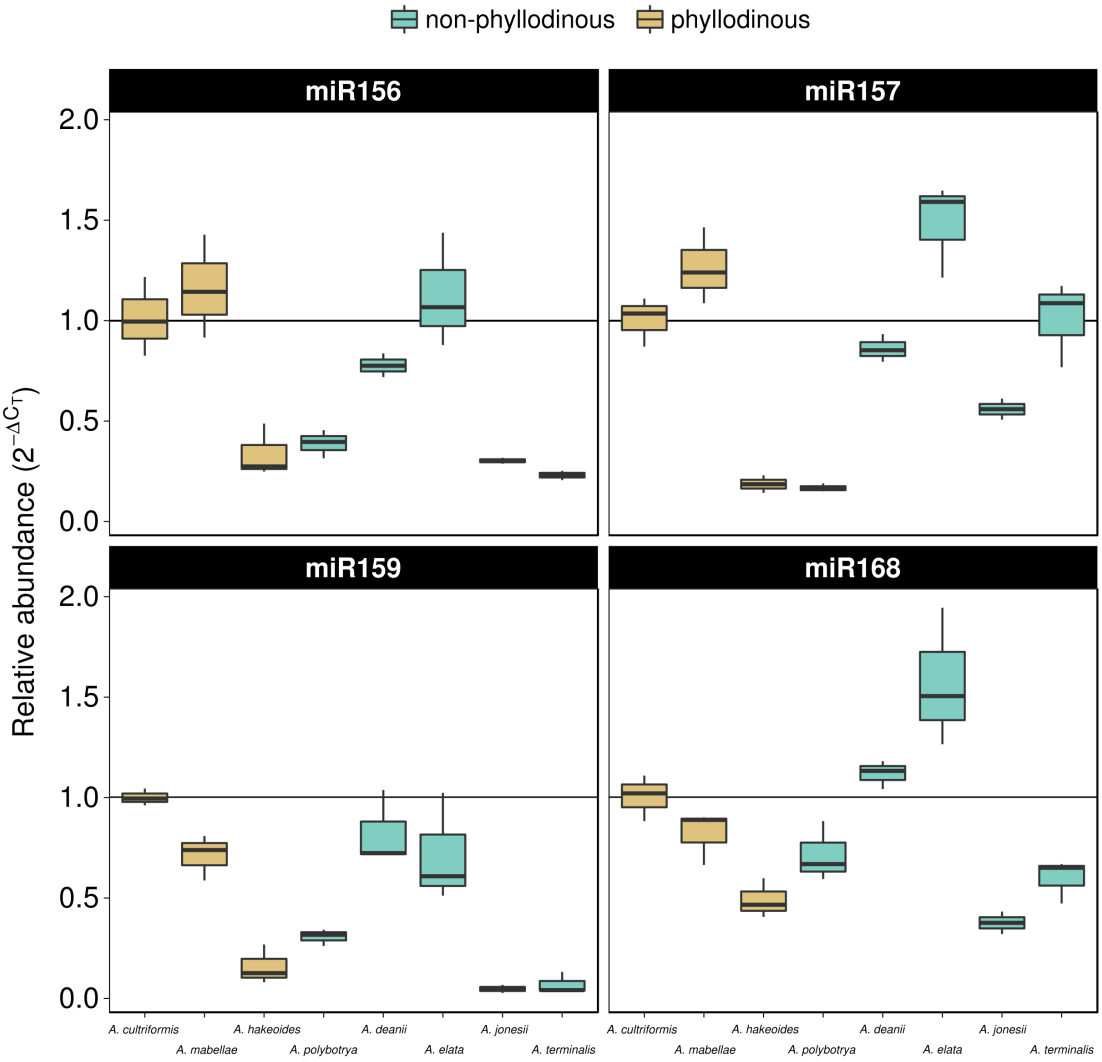


Figure 3.7. Levels of microRNAs in leaf 1 of phyllodinous and non-phyllodinous species. Levels of miR156, miR157, miR159, and miR168 were measured in leaf 1 of 3 phyllodinous and 5 non-phyllodinous species. Abundance was calculated relative to levels in *A. cultriformis*. Boxes bound 1st and 3rd quartile, center line marks the median.

3.4.6 *MIR156D* is responsible for much of the difference in miR156 levels between phyllodinous and non-phyllodinous species

The difference in the abundance of miR156/miR157 between phyllodinous and non-phyllodinous species could be due to changes in a common regulator of multiple *MIR156*/*MIR157* genes or to a change to a single gene. To evaluate these possibilities, I utilized whole genome sequencing of five *Acacia* species and one ancestor-like species (Table 3.3) to identify the genes that encode these miRNAs. The resulting assemblies revealed that the number of *MIR156/MIR157* genes in these species ranges from 14 in *P. lophanta* (an ancestor-like species) to 30 genes in *A. cultriformis* (Table 3.3). In combination with the genomic data from *V. collinsii* (Chapter 2), where I identified 13 *MIR156/157* genes, these data suggests that there has been a doubling of this gene family in *Acacia*. Although our estimated genome sizes of *Acacia* species are nearly double that of *P. lophanta* and *V. collinsii* (both ~0.5 Gb), I do not think the increase in number of *MIR156/157* genes is due to a whole genome duplication event because many of the genes that exist as single copies in *P. lophanta* or *V. collinsii* are present as tandem duplicates (up to 6 in the case of *MIR156B*) in *A. penninervis*. A full accounting of our sequencing effort can be found in Table 5.4.

Table 3.3. Summary statistics for genome sequencing.

Species	Estimated Genome Size (Gb)	Size of Assembly (% complete)	NG50 (bp)	Number of MIR156/157 genes
P. lophanta	0.50	0.61 (121.1)	18,034	14
A. rubida	1.28	0.83 (64.7)	132	22
A. penninervis	0.85	0.88 (103.9)	397,213	25
A. cultriformis	0.95	1.10 (116.3)	13,943	30
A. spectabilis	0.95	0.97 (102.5)	60,439	26
A. mearnsii	0.90	0.73 (81.0)	433	26

To determine which of these genes are most important for vegetative phase change, I examined their expression pattern in leaf primordia of *A. penninervis*. I designed 14 primer sets targeting all 25 of the identified *MIR156* and *MIR157* genes. All primer sets amplified a target of the expected size (Figure 3.8). *MIR156A.1-4*, *MIR156G*, *MIR156H.1-2*, *MIR157B*, and *MIR157E* were present at very low levels, which made it difficult to determine their expression pattern. *MIR156B.1-2,5-6* and *MIR157D* were relatively abundant, but increased from leaf 1 to leaf 9. *MIR156B.3-4*, *MIR156C*, *MIR156D*, *MIR156E.1-3*, *MIR156F*, *MIR157A*, and *MIR157C* decreased in abundance, consistent with a role in regulating the juvenile-to-adult transition. Although I did not control for difference in PCR efficiency between genes, I found that *MIR156D* consistently amplified with fewer cycles than all other genes, suggesting that it is the most highly expressed member of the gene family. This was also true for *V. collinsii* (Chapter 2).

I next examined how these genes behaved in non-phyllodinous species, specifically focusing on the *A. spectabilis* clade within the *Botrycephalae* (Figure 3.3, clade $H_{4)}$. In two phyllodinous species (*A. cultriformis*, *A. penninervis*), and two non-phyllodinous species (*A. mearnsii*, *A. spectabilis*), *MIR157* genes all declined during the first 3 months of development (Figure 3.9)—the exception being *MIR157D*, which increased in *A. mearnsii*. Importantly, two *MIR157* genes that are highly expressed and decrease temporally in *A. penninervis—MIR157A* and *MIR157C*—displayed steep drops in abundance across all species (Figure

3.9). This is consistent with our earlier finding that in non-phyllodinous species miR157 declines more than miR156 (Figure 3.6)

In contrast to *MIR157* genes, multiple *MIR156* genes behaved differently between phyllodinous and non-phyllodinous species. In *A. mearnsii*, the abundance of all *MIR156* transcripts (except *MIR156A*) increased. Only two genes consistently decreased in phyllodinous species and increased or remained the same in non-phyllodinous species, *MIR156C* and *MIR156D*. *MIR156C* increased more dramatically than *MIR156D* in non-phyllodinous species, but did not decline to the same extent as *MIR156D* in phyllodinous species. Sequence similarity and phylogenetic analyses suggest that *MIR156C* and *MIR156D* are Mimosoid-specific paralogs, as both genes are found in *V. collinsii,* but not other Papillonoid legumes (data not shown). I decided investigate the role of these two genes in the heterochronic event found in the *A. spectabilis*-*A. mearnsii* clade (Figure 3.3, clade H_4).

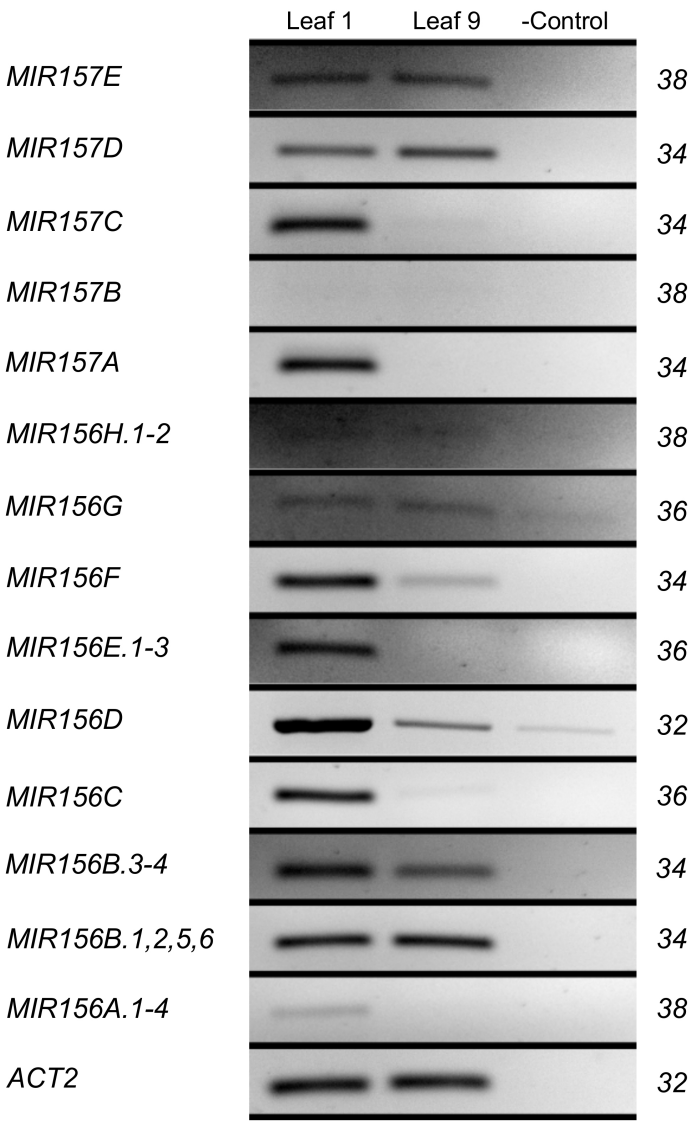


Figure 3.8. Abundance of *MIR156* and *MIR157* genes in *A. penninervis*. Semi-quantitative PCR of *MIR156/157* genes. Leaf primordia were sampled at 1-3 mm in size for nodes 1 and 9 (columns 1 and 2) and pools of 3-5 biological replicates were used for semi-quantitative PCR. Each gene was run on a single gel, allowing for direct comparisons of relative abundance between developmental stage and tissue type. Cycle numbers are on the right. Lane 3 is a no-RT control.

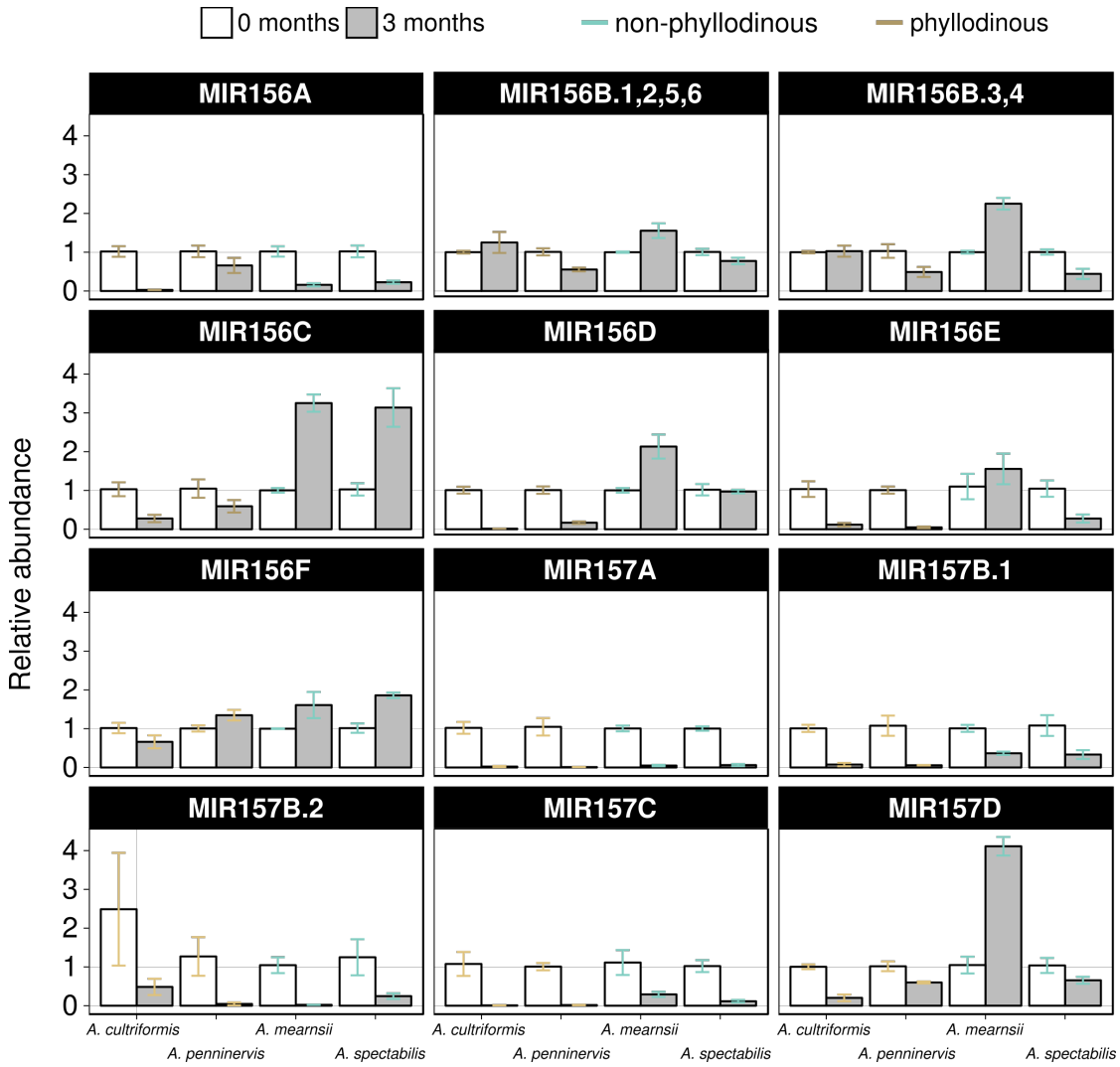


Figure 3.9. Temporal abundance of *MIR156* and *MIR157* genes in phyllodinous and non-phyllodinous species. qPCR was used to measure the abundance of genes over three months of development in fully expanded leaves.

3.4.7 *Acacia-MIR156D* is likely the functional ortholog of *MIR156C* in *Arabidopsis thaliana*

miR156 and miR157 are encoded by gene families ranging from 3 members in moss to over 30 in apple. In any given species it is unlikely that all of these genes play a significant role in vegetative phase change. In order for a gene to be important for a change in morphology during vegetative phase change it must be expressed in the appropriate tissue (leaf primordia and/or shoot apical meristem) and at levels high enough that it's eventual decline impacts SPL abundance. Because most gene family members produce an identical miRNA it is usually impossible to pinpoint the miRNA's source using qPCR. As a consequence mutant analysis is the only reliable way to implicate a gene's importance.

In *Arabidopsis thaliana* there are 8 miR156-producing loci, and 4 miR157-producing loci. Of the miR156 loci, *MIR156A* and *MIR156C* (Brassica-specific paralogs) produce nearly 90% of the miR156 at a stage when most genes have yet to decline (He et al., 2018). A similar situation is seen for miR157, where two genes, *MIR157A* and *MIR157C,* produce over 80% of the miR157. Given the disproportionate importance of these genes in vegetative phase change, it is important to investigate if these roles are conserved in other species.

I found that only a few genes in *M. truncatula* are likely orthologs to *AthMIR156A, AthMIR156C, AthMIR157A* and *AthMIR156C,* based on homology

between flanking genes in each species (Table 3.4). *AthMIR156A* shares one flanking gene with *MtrMIR156C*, *D*, *I* and *J*. *AthMIR156C* shares one flanking gene with *MtrMIR156D* and *E*. Of *AthMIR157A* and *C*, only *A* shares a flanking gene with *MtrMIR156H*. The low number of shared flanking genes illustrates the difficulty of determining gene ancestry for these microRNAs, which is a consequence of three whole genome duplication events in these lineages since their common ancestor (Ren et al., 2018). The distribution of flanking genes also suggests that the ancestor to Rosids likely had only a few *MIR156* and *MIR157* genes, and that during the process of duplication and divergence sets of flanking genes were subdivided across all resulting loci.

Table 3.4. Microsynteny analysis of *MIR156/157* genes between *A. thaliana* and *M. truncatula*. Number in boxes represent number of flanking genes showing homology between genes between species. Shading highlights the 4 most important genes for phase change in *Arabidopsis*.

		M. truncatula									
		MIR156A	*MIR156B*	*MIR156C*	*MIR156D*	*MIR156E*	*MIR156F*	*MIR156G*	*MIR156H*	*MIR156I*	*MIR156J*
A. thaliana	*MIR156A*	0*	0	**1**	**1**	0	0	0	0	**1**	**1**
	MIR156B	0	0	**1**	0	0	0	0	0	**1**	**1**
	MIR156C	0	0	0	**1**	**1**	0	0	0	0	0
	MIR156D	0	0	**1**	**2**	0	0	0	0	**1**	**1**
	MIR156E	0	**3**	**2**	0	0	0	0	0	**1**	**1**
	MIR156F	0	**1**	**1**	**1**	0	0	0	0	**1**	**1**
	MIR156G	0	0	**1**	**1**	0	0	0	0	**1**	**1**
	MIR156H	**1**	0	**1**	0	0	0	0	0	**1**	**1**
	MIR157A	0	0	0	0	0	0	0	**2**	0	0
	MIR157B	0	0	0	0	0	0	0	**2**	0	0
	MIR157C	0	0	0	0	0	0	0	0	0	0
	MIR157D	0	0	0	0	0	0	0	0	0	0

*A total of 3 up- and 3 down-stream genes were used for each locus.

Alignment and phylogenetic clustering of the hairpin sequences of *MIR156* and *MIR157* genes from *M. truncatula* with those of *A. penninervis* revealed that *MtrMIR156D* shows a high degree of sequence similarity with *ApenMIR156D,* and that these two genes cluster more closely with each other than with the Mimosoid-specific paralog of *ApenMIR156D*, *ApenMIR156C*. Thus, despite the fact that *ApenMIR156C* and *ApenMIR156D* share a common ancestry, their sequences are less similar than that of *ApenMIR156D* and a more distant homolog.

Another approach to evolutionary orthology—and—maybe more importantly—functional orthology, is to examine shared regulatory sequences. As discussed in Chapter 1, the temporal repression of *MIR156A* and *C* in *Arabidopsis* is mediated by deposition of H3K27me3 via PRC2 activity. Recently, *cis*-elements necessary for recruitment of the PRC2 complex to genes in *Arabidopsis* have been identified (Xiao et al., 2017). In particular, it has been shown that closely-spaced teleobox motifs (AAACCCTA) and GA repeats (AGAGAGAGA) are sufficient to mediate PRC2 repression. Additionally, recent work has shown that the B3 DNA-binding domain containing proteins VIVIPAROUS1/ABSCISIC ACID INSENSITIVE3-LIKE1/2/3 (VAL1/2/3) mediate the repression of *AthMIR156A* and *C* via recruitment of the PRC1/PRC2 complex to the promoters of these genes (Pico et al., 2015). VAL proteins bind to the RY-element (TGCATG) (Guerriero et al., 2009; Swaminathan et al., 2008).

I searched for these three motifs in the genomic region 1.5 Kb up- and down-stream of the miRNA sequence in each *Arabidopsis MIR156* and *MIR157* gene. Almost all of the *MIR156* and *MIR157* genes had all three of these elements in the vicinity of the miRNA sequence (8 of 12 genes). Notably, *MIR156H*, *MIR156G*, *MIR157B* and *MIR157A* had 2 or less of these elements (Table 5.5). Interestingly, *MIR156A-E*, had a cluster of all three elements within a 500 bp window (bold text Table 5.5). These clusters occur within exon 1 and intron 1 of *MIR156A* and span the TSS of *MIR156C,* two genes where the TSS has been mapped.

I was also able to find these elements in *MIR156* and *MIR157* genes of *A. penninervis*. However, in contrast to *Arabidopsis*, only 8 of 25 genes contained all three elements. Of these, only three genes contained these elements within a 500 bp window. Most strikingly was the similarity in position of the elements between *MIR156A* and *MIR156C* from *Arabidopsis* and *MIR156D* from *A. penninervis* (Figure 3.10). In fact the, absolute position of the elements in *AthMIR156C* and *ApenMIR156D* are nearly identical. This is an extraordinary degree of positional conservation given that these two species have evolved independently for over 80 million years and that there is less than 10% sequence similarity across the hairpin of these genes.

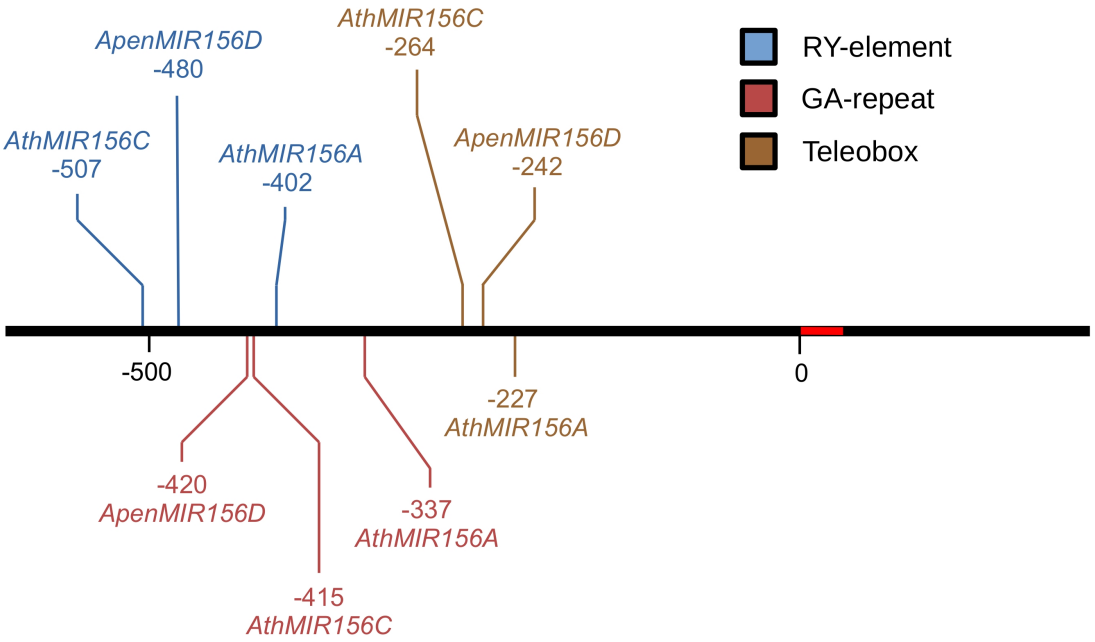


Figure 3.10. Distribution of cis-elements in vicinity of miR156. The relative and absolute positions of cis-elements are conserved between *AthMIR156A/C* and *ApenMIR156D*. Red line at 0 represents miR156 sequence. Numbers represent distance relative to start of miR156.

Additional evidence for evolutionary/functional homology is provided by structural similarities between different miRNA hairpins. Methods that utilize RNA secondary structure for phylogenetic inference have been used for ribosomal RNA (Telford et al., 2005; Keller et al., 2010) and structural information is commonly used to identify functionally-related proteins. To this end, I compared the RNA secondary structure plots of all *Arabidopsis*, *Acacia*, and *M. truncatula* *MIR156* and *MIR157* genes using a 300 bp region of genomic sequence centered on the miR156/157 sequence. Interestingly, *AthMIR156A* and *AthMIR156C* both had a branch or arm of double strandedness 6-9 bp from the nucleotide paired with the first or second nucleotide of the miR156 sequence (depending on if the processed miRNA is 20 or 21 bp in length) (Figure 3.11). A similar element was found in *AthMIR156H*, but no other genes contained this arm-element at a similar distance to the miRNA/miRNA* pairing. This same arm-element was found in *MtrMIR156D* and *MIR156D* for all 7 Mimosoid legumes that I had generated sequence data for (Figure 3.11). In Mimosoids, the arm-element was not found in other *MIR156* and *MIR157* genes, including *MIR156C* (a Mimosoid-specific paralog of *MIR156D*). The function of this element is unknown, but it appears to be required for correct processing of the primary miRNA transcript based on preliminary experiments using heterologous expression of *AcaciaMIR156D* in *Arabidopsis* (data not presented).

These three types of analyses (synteny, regulatory elements, RNA structure) strongly support the idea that *AthMIR156A/C* and *AcaciaMIR156C/D*

are orthologs. More specifically, the conservation of *cis*-elements between *AthMIR156C* and *AcaciaMIR156D* suggests that they are functional orthologs and are likely regulated similarly. I believe this functional conservation makes *AcaciaMIR156D* a strong candidate for a role in the evolution of neoteny in Acacia.

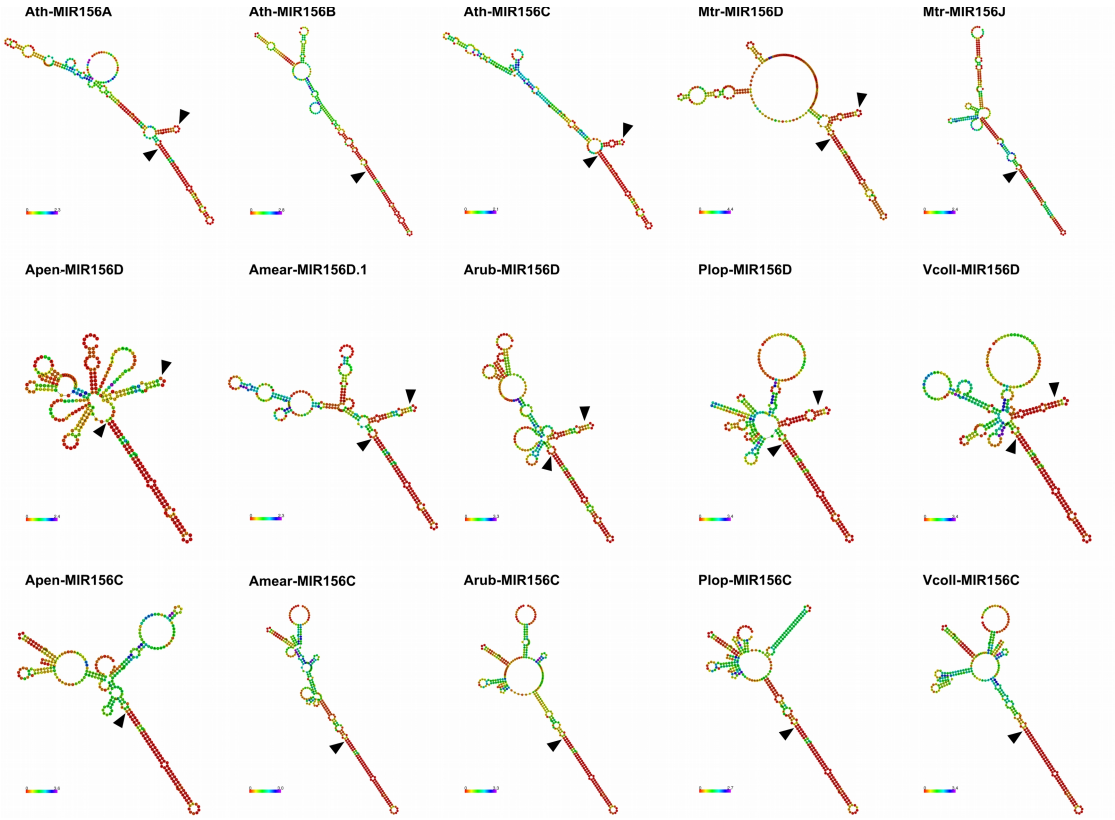


Figure 3.11. Structural comparison of hairpins from *Arabidopsis* and *Acacia*. Comparison of RNA folding in the hairpin of *MIR156* and *MIR157* genes revealed a conserved RNA element ("arm") 6-9 bp from the nucleotide paired with the first or second nucleotide in the miR156 sequence. The start of each miR156 sequence is denoted by an arrowhead on the underside of the hairpin. The arm-element is indicated with an arrowhead on the upper side of the hairpin.

3.4.8 A mutation in the *MIR156D* promoter is associated with neoteny in two independent clades

Using the our genome information and knowledge of putative *cis*-regulatory elements located in *MIR156D*, I sequenced the "promoter" region of this gene in a group of 98 species, spanning both the *A. rubida* and *A. spectabilis* neoteny events (Figure 3.3, clades H_2-H_4). I was able to sequence 3-8 copies of this region (usually obtaining both alleles for an individual) in all species/accessions listed in Figure 3.13, with the exception of *A. chrysotricha* where I was unable to amplify the region with our primer sets. The RY- and teleobox-elements were both absolutely conserved across all species (Figure 3.12, RY-element). As expected for a repeat element, the GA-repeat varied in size across the phylogeny, having no less than five GA repeats in species with the element. In addition to repeat expansion and contraction, I observed two other patterns of variation. In a small clade nested within the *A. spectabilis* event, I found that *A. parramattensis* ATSC, *A.mearnsii* ATSC1, *A.baileyana* NS, *A.dealbata* xDL all have a deletion of the GA-repeat element. The other pattern I found, was a TT-insertion into the right side of the GA-repeat element. Amazingly, an identical mutation was found in all species within the *A. neriifolia* and *A. spectabilis* clades (Figure 3.12 and 3.13).

Assuming that the TT-insertion arose once in the *Botrycephaleae*, this would mean that it has been lost no less than 4 times to account for phylogenetic

patterns. One prediction of this hypothesis is that different lineages will have different sized GA-repeat elements due to random excision of the TT-insertion. Indeed, species of the sister group to the *A. neriifolia* clade, containing *A. pravissima*, all have a GA-repeat element with 5 repeats, whereas the 3 sister clades to the *A. spectabilis* clade (containing *A. uncinata*, *A. cultriformis*, and *A. podalyriifolia*, respectively) all have elements with 6-8 repeats (Figure 3.12 and 3.13). The fact that these three clades do not differ consistent between each other could be obscurred by further contraction and expansion in each lineage following loss of the TT-insert.

The near perfect association between genotype and phenotype across this region of the phylogeny suggests that the TT-insertion is a loss-of-function mutation in the GA-repeat element. This element is predicted to mediate transcriptional repression therefore loss of this function would be predicted to reduce repression, either increasing overall levels of miR156 or interfering with the temporal decline of miR156. In *A. spectabilis MIR156D* is expressed at a constant level for 3 months after germination, whereas in *A. mearnsii* the expression of *MIR156D* increases (Figure 3.9). This difference may be due in part to the differences in the type of mutation found in each of these species. *A. spectabilis* contains the TT-insertion, whereas in *A. mearnsii* the GA-repeat element is entirely absent.

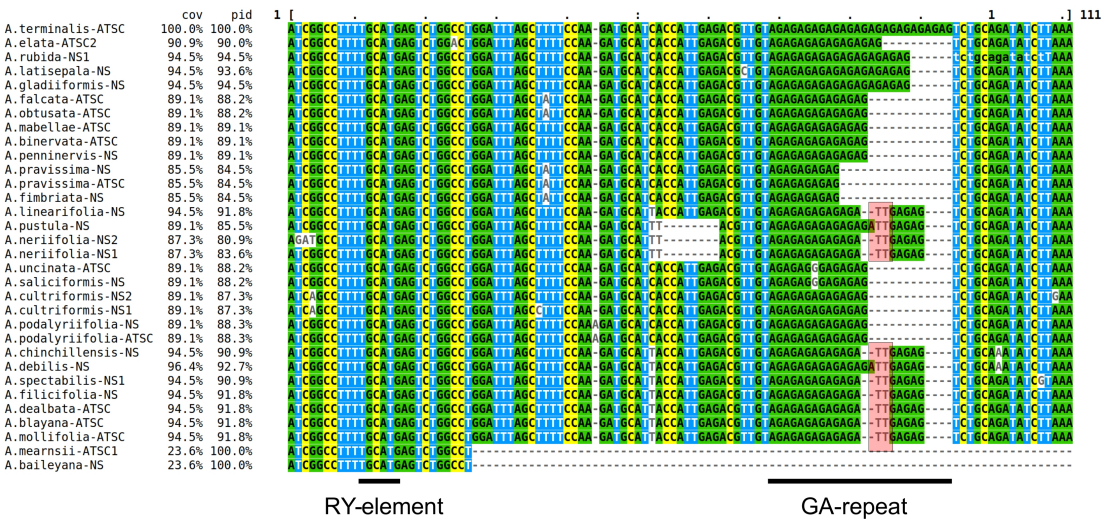


Figure 3.12. Identification of a mutation associated with neoteny in the *Botrycephaleae*. Alignment of the genomic region containing the RY- and GA-repeat elements in the promoter of *MIR156D*. Sequences are order according to the phylogeny in Figure 3.10 and were selected to represent the major clades and variants. A TT-insertion in the GA-repeat was identified in the *A. neriifolia* and *A. spectabilis* clades and shows a near perfect association with neoteny in these groups (38 of 39 species, 53 of 54 accessions).

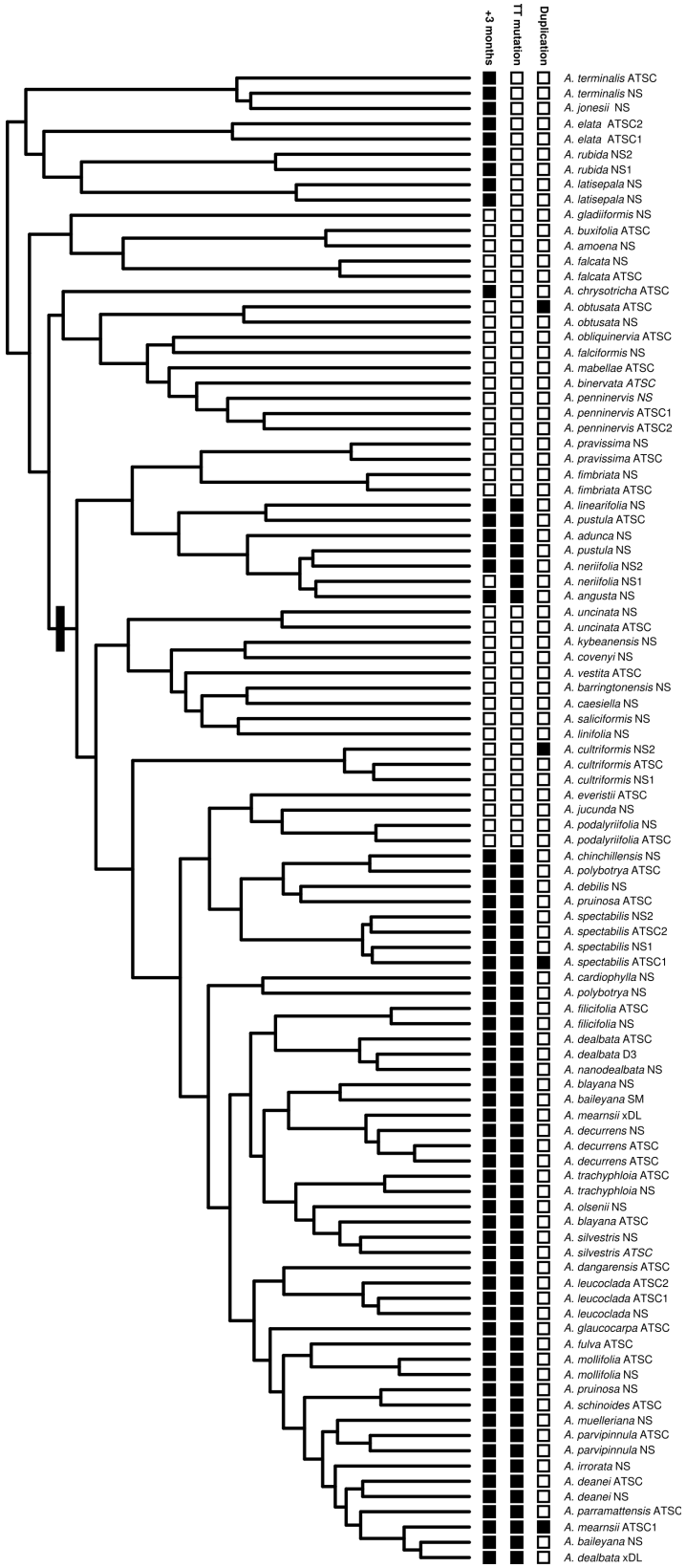

+3 months
TT mutation
Duplication
A. terminalis ATSC
A. terminalis NS
A. jonesii NS
A. elata ATSC2
A. elata ATSC1
A. rubida NS2
A. rubida NS1
A. latisepala NS
A. latisepala NS
A. gladiiformis NS
A. buxifolia ATSC
A. amoena NS
A. falcata NS
A. falcata ATSC
A. chrysotricha ATSC
A. obtusata ATSC
A. obtusata NS
A. obliquinervia ATSC
A. falciformis NS
A. mabellae ATSC
A. binervata ATSC
A. penninervis NS
A. penninervis ATSC1
A. penninervis ATSC2
A. pravissima NS
A. pravissima ATSC
A. fimbriata NS
A. fimbriata ATSC
A. linearifolia NS
A. pustula ATSC
A. adunca NS
A. pustula NS
A. neriifolia NS2
A. neriifolia NS1
A. angusta NS
A. uncinata NS
A. uncinata ATSC
A. kybeanensis NS
A. covenyi NS
A. vestita ATSC
A. barringtonensis NS
A. caesiella NS
A. saliciformis NS
A. linifolia NS
A. cultriformis NS2
A. cultriformis ATSC
A. cultriformis NS1
A. everistii ATSC
A. jucunda NS
A. podalyriifolia NS
A. podalyriifolia ATSC
A. chinchillensis NS
A. polybotrya ATSC
A. debilis NS
A. pruinosa ATSC
A. spectabilis NS2
A. spectabilis ATSC2
A. spectabilis NS1
A. spectabilis ATSC1
A. cardiophylla NS
A. polybotrya NS
A. filicifolia ATSC
A. filicifolia NS
A. dealbata ATSC
A. dealbata D3
A. nanodealbata NS
A. blayana NS
A. baileyana SM
A. mearnsii xDL
A. decurrens NS
A. decurrens ATSC
A. decurrens ATSC
A. trachyphloia ATSC
A. trachyphloia NS
A. olsenii NS
A. blayana ATSC
A. silvestris NS
A. silvestris ATSC
A. dangarensis ATSC
A. leucoclada ATSC2
A. leucoclada ATSC1
A. leucoclada NS
A. glaucocarpa ATSC
A. fulva ATSC
A. mollifolia ATSC
A. mollifolia NS
A. pruinosa NS
A. schinoides ATSC
A. muelleriana NS
A. parvipinnula ATSC
A. parvipinnula NS
A. irrorata NS
A. deanei ATSC
A. deanei NS
A. parramattensis ATSC
A. mearnsii ATSC1
A. baileyana NS
A. dealbata xDL

Figure 3.13. The origins of a mutation associated with neoteny in the *Botrycephaleae*. The phylogeny was generated from the largest ddRAD-seq matrix (minumum 30 species) and made ultrametric using the “extend” method in *phytools* (Revel, 2012). The three columns of boxes represent the length of the juvenile phase for a given species (left column, less than or greater than 3 months), the presence of the TT insertion in the GA repeat element of *MIR156D* (middle column) or the confirmed presence of a tandem duplication of *MIR156D* (right column). The black bar on the tree represents the likely origin of the TT insertion, assuming a single mutation event.

There are three likely possibilities for the origin of the TT insertion in the GA-repeat element of MIR156D. 1) Mutational convergence—whereby the same dinucleotide element was inserted at exactly the same position in independently evolving lineages. This scenario seems improbable. 2) Introgression—the mutation evolved once in one of the two lineages, and was transferred to the second lineage by introgression. Given that breeders have made hybrid varieties of different species in this group (though none reported across the scale necessary for this particular event), this might be possible. 3) Ancestral variation —the mutation evolved once in the common ancestor of both clades, and has been maintained in neotenous species while being lost in non-neotenous species.

Fortunately, these scenarios have different predictions for nucleotide variation across the locus (Manceau et al. 2010). Haplotypes of causal alleles (TT mutation) should be monophyletic if introgression was the cause, where as mutational convergence and ancestral variation predict that haplotypes should be paraphyletic (i.e. the mutations were superimposed onto a preexisting background of variation that matched each species phylogenetic history, or the mutation predates both lineages, and lineage specific sequence variation occurred after splitting from a common ancestor).

I evaluated the likelihood of these scenarios. Using a 665 bp region of *MIR156D* extending from the GA-repeat element to the miR156 sequence, I generated a phylogenetic tree of 98 species confined to the portion of the Acacia

phylogeny with most of the neotenous events (Figure 3.14A). Interestingly this tree places the *A. neriifolia* and *A. mearnsii* clades together, supporting the idea that TT-haplotype was introgressed from one clade to the other. However, if I remove the GA-repeat element from the alignment, the monophyly of these two clades is no longer supported, with multiple phyllodinous species grouping more closely with the *A. mearnsii* clade (Figure 3.14B). The discrepancy between these two trees further supports the uniqueness of the TT-haplotype and its relationship to morphology. Importantly these results are most consistent with mutational convergence or ancestral variation as the cause of this molecular pattern.

Future work could examine the frequency of SNPs and indels that show phylogenetic patterns identical to the TT-insertion. This analysis could be done using our ddRAD-seq data, and would allow us to quantify the rarity of the TT-insertion event. The bootstrap support for these clades being distantly related suggests that such nucleotide variation is rare.

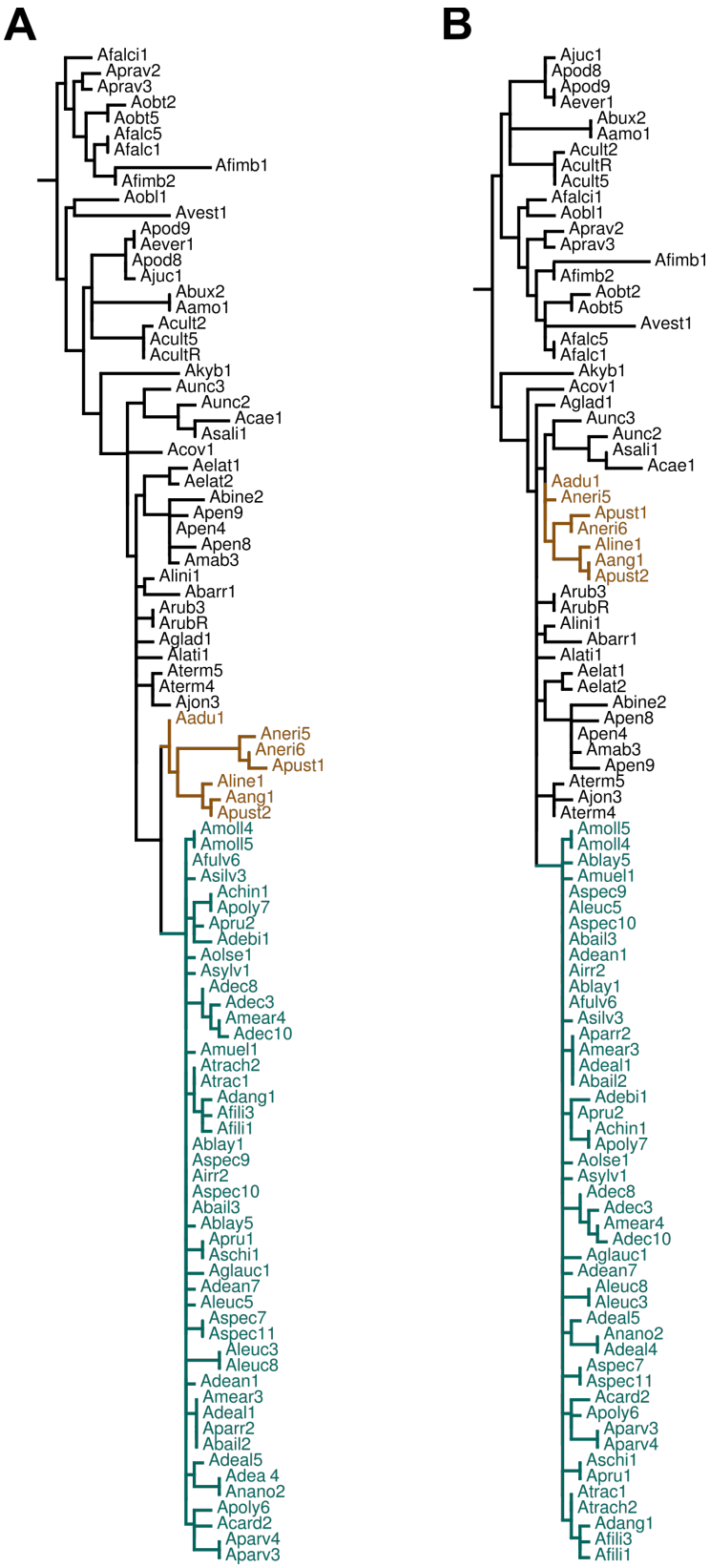

A
Afalci1
Aprav2
Aprav3
Aobt2
Aobt5
Afalc5
Afalc1
Afimb1
Afimb2
Aobl1
Avest1
Apod9
Aever1
Apod8
Ajuc1
Abux2
Aamo1
Acult2
Acult5
AcultR
Akyb1
Aunc3
Aunc2
Acae1
Asali1
Acov1
Aelat1
Aelat2
Abine2
Apen9
Apen4
Apen8
Amab3
Alini1
Abarr1
Arub3
ArubR
Aglad1
Alati1
Aterm5
Aterm4
Ajon3
Aadu1
Aneri5
Aneri6
Apust1
Aline1
Aang1
Apust2
Amoll4
Amoll5
Afulv6
Asilv3
Achin1
Apoly7
Apru2
Adebi1
Aolse1
Asylv1
Adec8
Adec3
Amear4
Adec10
Amuel1
Atrach2
Atrac1
Adang1
Afili3
Afili1
Ablay1
Aspec9
Airr2
Aspec10
Abail3
Ablay5
Apru1
Aschi1
Aglauc1
Adean7
Aleuc5
Aspec7
Aspec11
Aleuc3
Aleuc8
Adean1
Amear3
Adeal1
Aparr2
Abail2
Adeal5
Adea 4
Anano2
Apoly6
Acard2
Aparv4
Aparv3
B
Ajuc1
Apod8
Apod9
Aever1
Abux2
Aamo1
Acult2
AcultR
Acult5
Afalci1
Aobl1
Aprav2
Aprav3
Afimb1
Afimb2
Aobt2
Aobt5
Avest1
Afalc5
Afalc1
Akyb1
Acov1
Aglad1
Aunc3
Aunc2
Asali1
Acae1
Aadu1
Aneri5
Apust1
Aneri6
Aline1
Aang1
Apust2
Arub3
ArubR
Alini1
Abarr1
Alati1
Aelat1
Aelat2
Abine2
Apen8
Apen4
Amab3
Apen9
Aterm5
Ajon3
Aterm4
Amoll5
Amoll4
Ablay5
Amuel1
Aspec9
Aleuc5
Aspec10
Abail3
Adean1
Airr2
Ablay1
Afulv6
Asilv3
Aparr2
Amear3
Adeal1
Abail2
Adebi1
Apru2
Achin1
Apoly7
Aolse1
Asylv1
Adec8
Adec3
Amear4
Adec10
Aglauc1
Adean7
Aleuc8
Aleuc3
Adeal5
Anano2
Adeal4
Aspec7
Aspec11
Acard2
Apoly6
Aparv3
Aparv4
Aschi1
Apru1
Atrac1
Atrach2
Adang1
Afili3
Afili1

Figure 3.14. Comparison of trees generated from *MIR156D* haplotypes. Phylogenies were generated from promoter regions with (A) and without (B) the GA-repeat element of *MIR156D*. Branches of the *A. neriifolia* and *A. mearnsii* clades are colored brown and green, respectively. Trees were rooted with *P. lophanta*.

3.4.9 *MIR156D* exists as a tandem duplicate in some species of Acacia

The genome sequencing data revealed that *MIR156D* is tandemly duplicated in some species (Figure 3.15). This finding led us to hypothesize that duplication of the *MIR156D* locus in the *A. spectabilis* clade might explain why these species have a more severe form of neoteny than individuals in the *A. neriifolia* clade. After determining the break points of the duplication relative to *A. penninervis* (a species without the duplication) I was able to design primers that should be upstream of the break point in all species. Using such a primer in combination with a primer residing in the *MIR156D* hairpin I should be able to amplify the first tandem copy of *MIR156D*, which contains an intact copy of the hairpin sequence. Although these primers showed inconsistent results across the *Botrycephaleae*, I was able to amplify a tandem duplicate from *A. obtusata*, a species even more distant from *A. spectabilis* than *A. penninervis* (Figure 3.13). This means that the *MIR156D* duplication predates the *A. neriifolia* and *A. spectabilis* split, and has been lost in many species.

The tandem duplicates that I have been able to analyze show considerable sequence variation, with duplicates in species such as *A. mearnsii* and *A. cultriforms* showing changes that likely disrupt the function of miR156. Interestingly, I have been able to clone some of these duplicates from cDNA of leaf and leaf primordia from both *A. spectabilis* and *A. cultriformis*, meaning that they are expressed. Given the extensive variation in these tandem duplicates

and difficulties I have encountered with sequencing them, a detailed characterization is well beyond the scope of this work.

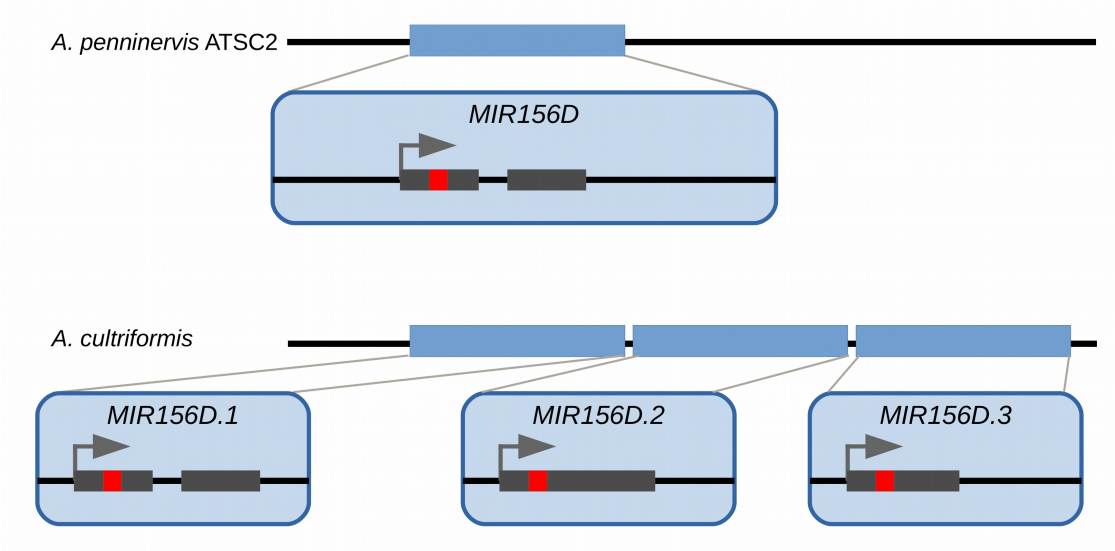


Figure 3.12. Tandem duplications of *MIR156D* exist in the *Botrycephaleae*. *MIR156D* exists as multiple tandem copies in some species within the *Botrycephaleae*. The depicted gene models are a hypothesis and have not been validated. In species with available data, the duplication starts ~800 bp upstream and ends ~5 Kb downstream of the miR156 sequence in *A. penninervis*.

3.5. Discussion

Life-history heterochrony has captivated biologists for over a century, but examples of how such changes in developmental timing are regulated at a molecular level are few. I used a candidate gene approach to examine a series of hypotheses for the origin of heterochrony in *Acacia.* Our results reveal that variation in the expression pattern and DNA sequences of genes involved in the vegetative phase change pathway have evolved in coordination with heterochronic patterns of shoot morphology in *Acacia*, and provide a molecular explanation for the evolution of neotenous species in this genus.

3.5.1 Mechanisms of neoteny

The multiple instances of neoteny in *Acacia* suggest that the molecular basis of this heterochronic event is relatively simple. Given that the temporal developmental program (miR156/SPL) that controls leaf morphology in neotenous Acacias is also present in phyllode-producing species, the simplest hypothesis for the continuous production of bipinnate leaves is the loss of the capacity to produce phyllodes. This could occur via gain-of-function (GOF) or loss-off function (LOF) mutations in genes essential for phyllode development. LOF mutations is the simplest hypothesis, given that such mutations are frequently one of the primary mechanisms of convergent evolution (e.g. *Mc1R* in animals, *FRI* and *FLC* in *Arabidopsis*). This hypothesis predicts that reversion from a non-phyllodinous to a phyllodinous state would be rare. Our phylogenetic

analyses reveal at least one such reversion event (clade H_4, Figure 3.3). Unfortunately, this single event hardly argues for one scenario over the other.

Depending on the mechanism of phyllode morphogenesis the possible routes to neoteny might be quite limited. If phyllodes evolved by utilizing genes common to juvenile leaf development, the number of “phyllode” genes that can be lost without also affecting juvenile leaf development are few. Even if neoteny occurred by this route, juvenile leaves in neotenous species would likely differ in some way from the juvenile leaves of phyllodinous species due to the pleiotropic effects of the defective gene/s. I have not observed such a difference, nor am I aware of any such reports.

An alternative hypothesis for neoteny in Acacia is that the timing of phyllode morphogenesis has been pushed back in development—in some cases so far back that phyllodes are never produced. This mechanism is supported by the finding that species like *A. rubida* and *A. latisepala*—which are known to have long juvenile phases or only produce phyllodes in certain geographical regions—are likely monophyletic with other species that never switch, like *A. elata* and *A. terminalis*. However it is naive to assume that neoteny evolves by the same mechanism in every group.

Our molecular work was almost entirely focused on the *Botrycephaleae* and much of our early analysis was confounded by the existence of independent neotenous events within this group. Future work should use a more explicit phylogenetic framework for evaluating how each event arose. Regardless, I

found that non-phyllodinous species from the *A. elata* event and from the *A. spectabilis* event have slower declines in miR156/157 during development than phyllodinous species. Not only do these miRNAs drop more slowly, but they never drop as far in non-phyllodinous species as in phyllodinous species. Unfortunately, I was unable to evaluate the *A. elata* event in detail due to limited seed stocks, and therefore focused most of our work on the *A. spectabilis* event.

Within the *A. spectabilis* event, I found a mutation within a predicted regulatory element that is also found in another clade of neotenous Acacias, the *A. neriifolia* group. The likelihood that that this mutation contributes to the delay in timing of phyllode morphogenesis in these groups seems high. However there are patterns within these two groups that need explanation. First, I identified one accession of *A. neriifolia* that only required on average 54 days to switch, relative to the 142 day average for the group. This species also has the TT-insertion, making it the one species that I identified where this mutation does not correlate with neoteny. This exception may demonstrate that the identified mutation is not causative. However, repeat expansions and contractions have a high mutation rate, and the fact I sequenced over 120 alleles in the *A. neriifolia* and *A. spectabilis* clades without observing a loss of the TT-insertion, seems to argue against this.

One explanation for this “fast-switching” *A. neriifolia* accession is that it has compensatory genetic changes that alleviate the effect of the TT-insertion effect. This could occur by changes to other components affecting the timing of

phyllode morphogenesis (e.g. another *MIR156* gene) or by changes at the same locus. Regardless, evidence that *MIR156D* has lost its temporal expression pattern in the rule-breaking *A. neriifolia* accession would bolster the idea that the TT-insertion alters this gene's regulation.

The second pattern that needs explanation is the difference in the strength of neoteny in the *A. neriifolia* and *A. spectabilis* clades. The TT-insertion may be sufficient to explain neoteny in *A. neriifolia*, but there must be additional changes that either modulate the strength of this mutation, or that act in parallel with this mutation to ablate phyllode morphogenesis entirely in the *A. spectabilis* clade. Maintenance of the TT-insertion in this clade argues against additional changes that disrupt how phyllodes are constructed. Put another way, why disrupt the timing of phyllode development if you can't make them anyway? This is probably an oversimplification, but again illustrates how knowledge of phyllode development will greatly increase our understanding of how neoteny can evolve.

3.5.2 Conservation of *Acacia-MIR156D* and *Arabidopsis-MIR156A/C*

The finding that *MIR156D* in *Acacia* is the likely functional ortholog of *MIR156A* and *MIR156C* in *Arabidopsis* raises an obvious question: are these genes functionally distinct from other members of the *MIR156/157* families, which also show temporal regulation during development? One possibility is that the conserved *cis*-elements in these genes not only control the timing of repression,

but also its location. This hypothesis might explain, why in *V. collinsii* miR156 and *pri-MIR156D* decline in leaf and stipule primordia, but not in fully expanded leaves (Chapter 2). This could mean that repression of the *MIR156D* locus occurs within the meristem and either this repression is lost or the relative abundance of *MIR156/157* genes changes during morphogenesis. Levels of both miR156 and miR157 are known to rise during leaf ontogeny (Xie et al., 2012). In *Arabidopsis* the use of *mir156a/c* and *mir157a/c* mutants demonstrated that this rise was at least in part due to these genes (He, 2017). Therefore it may be the case that the *cis*-elements responsible for the repression of these genes early in a leaf's ontogeny do not respond to repression (or can be overridden) at a later point in ontogeny.

3.5.3 How information about molecular mechanism contributes to an understand of evolution

It may seem self-evident that an understanding of the molecular mechanism of trait developmental is important for understanding how the trait evolved. However, comparative phylogenetics is only rarely accompanied by a detailed understanding of genetic mechanism. I began our study by using phylogenetics to test hypotheses concerning the frequency of neoteny in *Acacia*. From this analysis I concluded that neoteny evolved independently in the *A. neriifolia* and *A. spectabilis* clades. This result further suggests the mechanism of neoteny could be qunique to each clade. However, if one assumes that the TT-

insertion is responsible for a delay in phyllode production and that it evolved only once in the ancestor of these clades, then one comes to a radically different conclusion for the evolution of neoteny in these clades.

While most practitioners of phylogenetic comparative methods are well aware of the limits to their methods, it is important to highlight when disagreements between such methods and other approaches arise. Phylogenetic trees are a hypothesis, and ultimately (depending on the question) they must be tied to mechanistic studies to actually understand how life evolves.

4. Conclusions and future work

4.1. Conclusions

In this work I have utilized two genera of woody legumes to explore how the miR156/miR157 pathway evolves and contributes to the evolution of life history events. In chapter 2, I examined a classic model system for the study of how animal-plant mutualisms evolve, the New World Ant-Acacias (genus *Vachellia*). I found that the age-dependent decline in miR156 and miR157 is tightly correlated with the onset of the swollen thorn syndrome, the suite of plant traits necessary for attracting, housing, and feeding ants. Shade manipulations delayed the appearance of the swollen thorn syndrome, which was associated with elevated levels of miR156 and miR157, further supporting the role of this pathway in regulating the syndrome. These findings support the hypothesis that the delay in the production of syndrome traits is adaptive.

In Chapter 3, I used the genus *Acacia* as a model system for understanding how changes in the miR156/miR157 pathway might explain macroevolutionary patterns of leaf morphology. Phylogenomic analysis of over 150 species revealed that the developmental delay in the shift between juvenile and adult leaf morphology has occurred at least 7 times. I found that in two of these neotenous groups, levels of miR156, show a reduced rate and level of decline in non-phyllodinous species compared with phyllodinous relatives. Using genome sequencing and *de novo* assembly, I was able to identify *MIR156* genes

that correlated with these differences in miR156 levels. Comparative analysis of these genes implicated one gene, *MIR156D* as a functional ortholog of *MIR156C* in *Arabidopsis*. Sequence analysis of over 120 haplotypes from the *MIR156D* promoter found a mutation in a predicted *cis*-regulatory element that shows near perfect association with a delay in phyllode development in two distantly related clades. Together these studies demonstrate the value of examining non-traditional systems using a candidate gene approach to better understand the evolutionary potential of developmental regulatory pathways.

4.2. Future work

4.2.1 Remaining work on the *A. neriifolia-A. spectabilis* story

A few important experiments could resolve how neoteny evolved in the *A. neriifolia* and *A. spectabilis* clades. Primary among these are functional tests of the identified PRE elements in *Arabidopsis* and *Acacia*. Xiao and colleagues (2017) developed a reporter construct to test the ability of putative PRE elements to repress the transcription of flanking promoters in *Arabidopsis*. This construct should be used to determine if putative PREs from *AthMIR156A* and *AthMIR156C* are capable of mediating repression. However, *a priori* expectations for the pattern of repression are unclear. Given evidence that these genes increase in expression during leaf ontogeny (He, 2017) it is unlikely that these elements are capable of exerting repression in fully expanded leaves. I believe one possibility is that these elements exert repression within the meristem, and as leaves develop this repression is lost (3.5.2). This in combination with the fact

that the PRE reporter uses strong constitutive promoters such as pMAS and p35S, it is possible that repression will be not be visible in leaf primordia. Therefore tests of these putative PREs should also examine the meristem of transgenic plants to determine how these PREs function.

It is also important to test PREs from *Acacia* species with and without the TT-insertion. It is possible, given the high degree of conservation between *AcaciaMIR156D* and *AthMIR156C*, that heterologous tests of function in *Arabidopsis* would be successful, but on the chance that they are not, I have developed a method for testing function in *Acacia*. Trangenesis in *Acacia* is still a very low efficiency process. To minimize the amount of work necessary to evaluate independent transgenic events I have designed a dual reporter (Figure 4.1). Unfortunately in *Acacia*, GFP and RFP reporters work poorly and though selective agents like Basta and Hygromycin work on *Acacia* explant tissue, the vast amount of surviving tissue on these selective agents remains untransformed. GUS is a very reliable marker of successful transformation. Additionally, I have developed another visual marker system that leverages the ability of the bacterial carotene desaturase enzyme CRT1, to bypass the block in carotenoid biosynthesis mediated by the bleaching herbicide Noraflurazon. In the reporter (Figure 4.1) the putative PRE is placed next to the promoter driving GUS. Hygromycin is used to promote the proliferation of transformed tissue, and CRT1 is used to visually identify such tissue. The strength of the PRE will then be

measured as the ratio of green-blue (i.e. non-bleached and GUS positive) sectors to green only sectors.

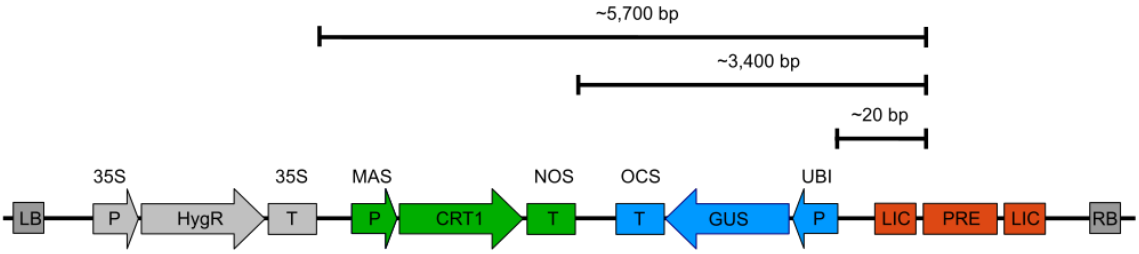


Figure 4.1. PRE reporter for *Acacia*. Putative PREs are cloned into the construct using ligation independent cloning (LIC). Selective growth of explant tissue is mediated by Hygromycin resistance (HygR), and visual identification of transgenic sectors are mediated by absence of photobleaching conferred by the CRT1 enzyme's metabolism of the bleaching herbicide Noraflurazon. GUS expression is used to quantify PRE-mediated repression.

In addition to demonstrating that the TT-insertion is capable of reducing repression, it is important to unequivocally establish that miR156 and miR157 actually control the switch from juvenile to adult leaf morphology. While the correlation between the temporal decline of these microRNAs and the switch in leaf morphology is strong, without direct evidence of this role, the entire story rests on an unvalidated assumption. I am very conscious of this fact, and have spent the past 5 years attempting to improve *Acacia* transformation for this purpose. As mentioned above, attempts to use fluorescence as a visual marker were largely unsuccessful. GUS allowed me to optimize protocols, but still meant that I had to ultimately sacrifice promising tissue. Recent development of CRT1 as a visual marker has taken me one step further. With this marker I can visually follow proliferation of positively transformed tissue. I now have a few visually positive transgenic nodules that are proliferating on Hygromycin media. The final step is to get these nodules to shift from proliferative-meristematic growth to determinant-internodal growth. This is generally a very slow process in *Acacia*.

4.2.2. How best to approach the mechanisms of neoteny in *Acacia* going forward

One major benefit of using the genus *Acacia*, is the large amount of phenotypic variation associated with neoteny. My finding that there are at least 7 instances of neoteny in the genus will be a great resource for comparative

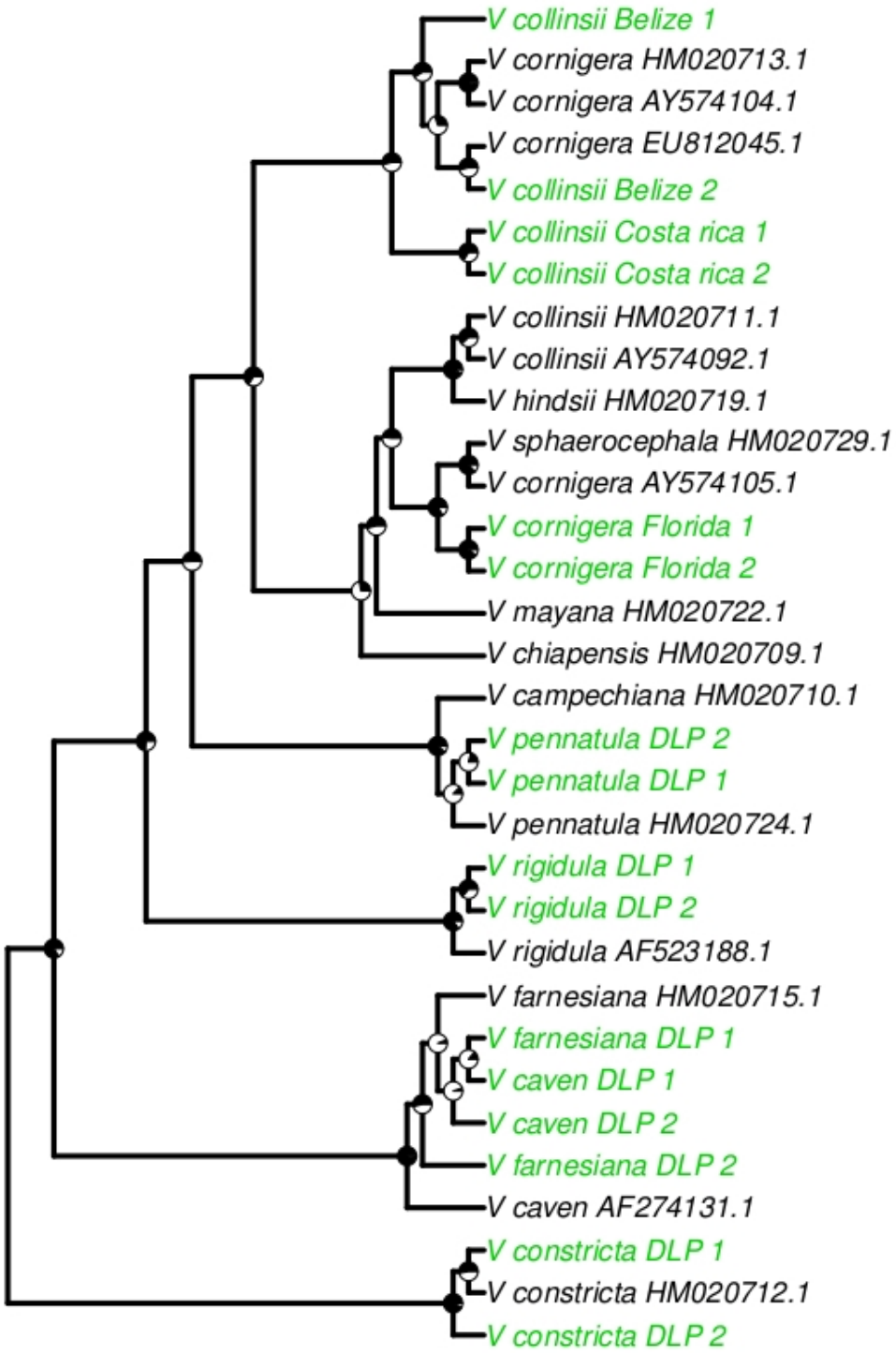
studies of convergent evolution in the future. If many of these events also end up involving the miR156/miR157 pathway, they will be an additional resource for understanding the ways in which developmental timing can be modified.

As I demonstrated in Chapter 3, not only is there extensive phenotypic variation within *Acacia*, but there is also large amounts of genetic diversity as it relates to the miR156/miR157 pathway. In some cases this degree of diversity can be overwhelming, and difficult to synthesize (e.g. the tandem duplicates of *MIR156D*), but can also be powerful in that it offers natural "mutagenesis" experiments (e.g. presence/absence of the TT-insertion). As an example, I recently sequenced an individual of the slow-switching *A. neriifolia* accession that was heterozygous for the TT-insertion. These plants are still too young to assess the length of their juvenile phase, but this one individual does appear to be transitioning to adult leaf morphology more quickly than an individual homozygous for the TT-insertion. This example suggests that there may also be extensive intraspecific variation in the timing of these traits. This is further confirmed by observations of accessions such as *A. melanoxylon* ATSC3, were some individuals resembled *A. melanoxylon* ATSC1 in the timing of phyllode production (Figure 3.2, red outliers). Future work should examine the duration of the juvenile phase from a large number of seeds of accession that show high levels of variation. Such an analysis could focus on candidate genes, but as genome assemblies improve, association mapping will become a possibility.

These observations also emphasize the likelihood that observing plants in the field might turn up new resources and hypothesis.

Another possible approach to examining neoteny, is to leverage the new phylogenetic information for genetic analyses. Phyllodinous and non-phyllodinous species can be hybridized. *Acacia* x *hanburyana* was made over 100 years ago from a cross between *A. dealbata* and *A.podalyriifolia* (Orchard, 2001). This hybrid produces transitional leaves on flowering branches, but is unfortunately sterile. Examination of my phylogeny (Figure 3.2) shows that this cross is of distant species, and that other more appropriate crosses could be made. For example, the clade of phyllodinous species containing *A. podalyriifolia*, has a sister clade of non-phyllodinous species that are more closely related than that of *A. dealbata*. Crosses between *A. neriifolia* and a fast switching sister species such as *A. pravissima* would also be worthwhile endeavor.

5. APPENDIX



5.1. *MATK* gene phylogeny of species used in Chapter 2. The coding sequence of the chloroplast marker, *MATK* was used to construct an maximum

likelihood tree using RAxML. Sequences from species/accessions in black were obtained from Genbank. Sequences from species in green were from this study. Piecharts represent bootstrap support for nodes. The root was pruned from tree. In general, the number of differences between species was very low.

5.2. Genome assembly statistics for *V. collinsii* Belize (Vcoll-BR v1). Related to Figures 2-4.	
Kmer estimated genome size	518,254,840 bp
Assembly length (percent of total)	461,064,704 (89.0%)
Number of contigs (scaffolds)	122,266 (122,260)
NG50	6,528
Longest contig/scaffold	301,273
complete BUSCOs	86.9%
missing BUSCOs	7.6%

5.3. Primers used in Chapter 2.

Primer ID	Use/Experiment	Sequence
	Phylogenetics	
trnK685F	MATK sequencing; From Wojciechowski et al., 2004	GTATCGCACTATGTATCATTTGA
trnK2R*	MATK sequencing; From Wojciechowski et al., 2004	CCCGGAACTAGTCGGATGG
trnK_3'_seq	MATK sequencing	TACTACATGAGCATTTCCTAATCCAT
	smRNA-qPCR	
SPL_degenerate_motif_F1	3'RACE	MGBTTYTGYCARCARTGYAGYAGGTTYCA
miR156_Forward	miR156 qPCR	GCGGCGTTGACAGAAGAGAGT
miR157_Forward	miR157 qPCR	GCGGCGTTGACAGAAGATAGA
miR159_Forward	miR159 qPCR	GCGGCGTTTGGATTGAAGGGA
miR168_Forward3	miR168 qPCR	AAGGCGTCGCTTGGTG
miR156/7_RT_primer	miR156/7 reverse transcription	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACGTGCTC
miR159_RT_primer	miR159 reverse transcription	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTAGAGC
miR168_RT_primer	miR168 reverse transcription	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTTCCCG
URP	universal reverse for smRNA qPCR	GTGCAGGGTCCGAGGT
	mRNA-qPCR	
legume_ACT2_F	ACT2-LIKE qPCR	TGGCTCCACCAGAGAGAAAGTA

legume_ACT2_R	ACT2-LIKE qPCR	GAAGGTGCTGAGGGATGCAAG
VcoSPL1-like1_F1	semi-qPCR	TTGTAGGGCCGATCTTGGTAATGC
VcoSPL1-like1_R1	semi-qPCR	GGATGTGTCTTCCTCCTCCTCCTG
VcoSPL1-like2_F1	semi-qPCR	TTGAATGAGGGAGCAGGCGGCC
VcoSPL1-like2_R1	semi-qPCR	CCAGCCAATCTAGTTTTCTTTCCGTTC
VcoSPL1-like3_F1	semi-qPCR	CACCGCCGACACAAAGTTTGTGAGCT
VcoSPL1-like3_R1	semi-qPCR	ATTGGTTAGCAACGGAGGCTTTAGAAG
VcoSPL3-like1_F1	semi-qPCR	AGGGTTTGCCGGAGGAGGAAGTAG
VcoSPL3-like1_R1	semi-qPCR	CATCACTGAGCTCCGCCGTGCAG
VcoSPL3-like2_F2	semi-qPCR	GAGAGGTGTGGAGCTAACTTGGC
VcoSPL3-like2_R2	semi-qPCR	TTAGCCCTGCAACGATCACAACAG
VcoSPL3-like3_F1	semi-qPCR and qPCR	ATGGCGGTGTTGCAGAAGAGCAG
VcoSPL3-like3_R1	semi-qPCR and qPCR	GACCCTCTTCTTCCATTTCCCATGA
VcoSPL3-like4_F1	semi-qPCR	TGGCTCTGGAGAAGAACATGGCC
VcoSPL3-like4_R1	semi-qPCR	CTTCCAGCTGGTTTGGGGCATCTC
VcoSPL3-like5_F1	semi-qPCR and qPCR	GATGGGAAGAAGAAGATGAGAGATCGG
VcoSPL3-like5_R1	semi-qPCR and qPCR	GTGCCATTGAGGAGGAAGAAGAAGAC
VcoSPL6-like1_F1	semi-qPCR	CAGCCATTGTTTTTAGGTTCTCTAGTTGAATC
VcoSPL6-like1_R1	semi-qPCR	GACCGAGCTCTCCTTGCAATTTTA

		TC
VcoSPL6-like2_F1	semi-qPCR	AAAGCGTTGAGTTTGTGGACTTAGGG
VcoSPL6-like2_R1	semi-qPCR	AGTTTGATGCTTCCAGTGGTTGGC
VcoSPL6-like3_F1	semi-qPCR	AGGACGAGTTGGTGATCAAAGAGATGC
VcoSPL6-like3_R1	semi-qPCR	TTTCTACGACGCTCATTATGGCCTG
VcoSPL6-like4_F1	semi-qPCR	GGATGTTCCAAGGGCCAGGTTTTG
VcoSPL6-like4_R1	semi-qPCR	GCTCAACTCTCTGAAGCTGTGAC
VcoSPL7-like_F1	semi-qPCR and qPCR	CGAAGCTGTAGAAGAAAATTAGAGCGC
VcoSPL7-like_R1	semi-qPCR and qPCR	GGCAGTTTCTACGTCGTTTATCTCAC
VcoSPL8-like1_F1	semi-qPCR	ATGGCCTCGTCGAACTCGCCTC
VcoSPL8-like1_R1	semi-qPCR	GTGGTAGTGCTTGGCTTGTGATAG
VcoSPL8-like2_F1	semi-qPCR	GGGCTCTACCATGTCGTTGAACAAC
VcoSPL8-like2_R1	semi-qPCR	GATGATAGTGCTTCGCGTGGGAG
VcoSPL9-like1_F1	semi-qPCR	CTATTCCAGGCATAAAGTCTGTGGC
VcoSPL9-like1_R1	semi-qPCR	AGACCAGCAACAATGACTTTAGGGC
VcoSPL9-like2_F1	semi-qPCR	TGTCAGGTGGAGGGCTGCAAAGTG
VcoSPL9-like2_R1	semi-qPCR	CTCTTGACCTGCGACAACGACCCT
VcoSPL10-like1_F1	semi-qPCR	AAACGACTGTACTTTGAAGATGTC

		TGTG
VcoSPL10-like1_R1	semi-qPCR	GATAGGTCAAGTTTACAGCCTTCAACC
VcoSPL10-like2_F1	semi-qPCR	TTGTACGGAGGAGGCGGTGTGAAG
VcoSPL10-like2_R1	semi-qPCR	CAACCCTAACACCACAACCTTCGGT
VcoSPL13-like1_F1	semi-qPCR and qPCR	GGAGAACAACCAGAGGAACAATGACC
VcoSPL13-like1_R1	semi-qPCR and qPCR	TCGACGTGAAGTGTAGCTGCCGG
VcoSPL13-like2_F1	semi-qPCR	TGGGGTCTTCTGAAAGCGCCAGAA
VcoSPL13-like2_R1	semi-qPCR	CAGAGCTTCAACACATGGCTTCTCAAG
VcoSPL13-like3_F2	semi-qPCR and qPCR	TTGGTGGATGGATGCAGCTTTGAC
VcoSPL13-like3_R2	semi-qPCR and qPCR	AGGTTTCCTTCTGCGCTTATTATGG
VcoSPL13-like4_F1	semi-qPCR	ATCGGCGACATAGAGTGTGTGAGAAG
VcoSPL13-like4_R1	semi-qPCR	CCCCTACCAACACAATGGGCGTC
VcoMIR156a_F2	semi-qPCR	GATGAATCGAGATCAGTTTTAGCTGC
VcoMIR156a_R2	semi-qPCR	CTTCCCTTACCCTTCAAAGTCTCCATT
VcoMIR156b_F2	semi-qPCR	CTGCCTTCTTGTGGTTGATTTTTCTTC
VcoMIR156b_R1	semi-qPCR	GCACCGCGTATGTATGCCTCTC
VcoMIR156c_F1	semi-qPCR	CTCGACTCCTATAGCTTCTTTCATATG

VcoMIR156c_R1	semi-qPCR	GAACCAGAAATAGCAGTGAGAGACTC
VcoMIR156d_F1	semi-qPCR	CACTCATGATGTTGCTTCTTTCTTGGG
VcoMIR156d_R1	semi-qPCR	AAAAGGGAACGAAGCAGAGCCAGA
VcoMIR156e_F2	semi-qPCR	TCCGATTATTGCAGATTGCAGAGTAG
VcoMIR156e_R2	semi-qPCR	CTGTATAAGTAGTAGCAGAAAGTAAAAGG
VcoMIR156f_F1	semi-qPCR	ACGTGAACAAACAAAGTAGAGATTAATGAC
VcoMIR156f_R1	semi-qPCR	CACCCAACTGCAAATGTATGAACG
VcoMIR156g_F1	semi-qPCR	AGAAGAAGAAGAAGATAAGTTATGTGTTG
VcoMIR156g_R1	semi-qPCR	CACAGCTTTTTAGTCATTCCCGGG
VcoMIR156h_F2	semi-qPCR	TATTGTGTGAYCGTGTGTGTGTGTAGAG
VcoMIR156h_R2	semi-qPCR	CACCACCCTCCCTCTTTTCCTTC
VcoMIR157a_F1	semi-qPCR	GATAAAGAGGAATGGCTATGGCGATG
VcoMIR157a_R1	semi-qPCR	GATGCAATATTGATTTCCTCGGTTCAG
VcoMIR157b_F2	semi-qPCR	CCTTTTTTGCTAATTGTTAAGAGGAGCTG
VcoMIR157b_R2	semi-qPCR	CATCACCAAATTAAAACACGCCTCCCTT
VcoMIR157c_F2	semi-qPCR	GAAAGAGATACATCAGGCAGGAGGC
VcoMIR157c_R2	semi-qPCR	AATGAGGAGATAGGGACTGGAGGC

VcoMIR157d_F1	semi-qPCR	GAGGGAGATCGAGGAGGTTTCAG
VcoMIR157d_R1	semi-qPCR	AGATGATGATTTGCATGAAACTCACTGC
VcoMIR157e_F2	semi-qPCR	CAGGTTTGATTGTTTCCGTGATGCAC
VcoMIR157e_R2	semi-qPCR	TGCCGCCTTCTTCTTCTTCAGC

5.4. Summary of libraries used for genome assembly in Chapter 3.

				Post Filtering (trimming, quality filtering, error correction)		
Species	Insert Size (PE) or Read length (SE) (bp)**	Library Type	Read Format	Total Data (Gb)	Sequence Depth*	Reads (M)
A. cultriformis (AcultR)	X**	Truseq PCR-free; Nextra Matepair	merged or singletons	34.1	35.9	166.7
	308	Nextra Matepair (from filtering)	150PE	13.1	13.8	54.6
	422	Truseq PCR-free	250PE	1.2	1.3	2.6
	1,861	Nextra Matepair (gel-free)	150PE	16.3	17.2	136.4
total				64.7	68.1	360.3
A. mearnsii	142**	Nextera	merged or singleton	8.6	9.6	60.9
	500	Nextera	100PE/250PE	23.5	26.1	256.4
total				32.1	35.7	317.3

A. penninervis (Apen4)	186		merged or singleton	81.2	95.5	436.0
	181	Nextra Matepair (from filtering)	125PE	31.8	37.4	163.6
	276	Truseq PCR-free	150PE/ 250PE	54.4	64.0	208.4
	2743	Nextra Matepair (gel-selected)	125PE	12.5	14.7	62.3
	4042	Nextra Matepair (gel-selected)	125PE	9.8	11.5	49.0
	6146	Nextra Matepair (gel-selected) (gel-selected)	125PE	9.5	11.2	46.9
	7479	Nextra Matepair (gel-selected)	125PE	9.8	11.5	48.7
	9199	Nextra Matepair (gel-selected)	125PE	9.7	11.4	48.4
	15,171	Nextra Matepair (gel-	125PE	11.5	13.5	57.1

		selected)				
	5,062	PacBio P5-C3	500-43,893	10.4	12.1	2.0
total				240.6	282.9	1122.4
A. rubida (ArubR)	115	Truseq PCR-free	merged or singleton	13.2	10.3	115.2
	750	Truseq PCR-free	125PE/150PE	17.6	13.8	177.2
total				30.8	24.1	292.4
A. spectabilis (AspecR)	341	Truseq PCR-free and Nextera Matepair	merged or singleton	35.8	37.7	146.2
	308	Nextra Matepair (from filtering)	150PE	6.6	6.9	27.7
	491	Truseq PCR-free	250PE	11.9	12.5	24.3
	646	Truseq PCR-free	125PE	51.7	54.4	220.9
	2,050	Nextra Matepair (gel-	150PE	17.5	18.4	74.0

		free)				
total				123.5	129.9	493.1
P. lophanta (PlopR)	406**	Truseq PCR-free	merged or singleton	20.6	41.4	50.7
	450	Truseq PCR-free	250PE	3.1	6.2	12.2
total				23.7	47.6	62.9
*Using estimated genome sizes **Observed average read length or insert size						

5.5. Occurrence of cis-elements in vicinity of miRNA of *MIR156* and *MIR157* genes of *Arabidopsis* and *A. penninervis*.

Gene	Teleobox (distance relative to miRNA)	GA-repeat	RY-element
A. thaliana			
MIR156a	-1075, **-227**	**-337**, +109	**-402**, +229, +611
MIR156b	**-109**	**+84**	**0**, +1057, +1420
MIR156c	**-264**	**-530**, **-415**	**-507**, +11, +230, +630
MIR156d	**-481**, -**274**	**-439**	**-406**, +1137
MIR156e	**+29**, +847	**+62**, **+80**	-952, **+13**, +1258
MIR156f	-71, +587	-296, +89	-1003
MIR156g			-1227,-428, -189, -56, +165, +882
MIR156h			
MIR157a	-1256		+1199, +1444
MIR157b		-1019	
MIR157c	-672	+93, +1426	+8, +766
MIR157d	-914, -311	-1130	+526
A. penninervis			
MIR156a.1		+47, +76	+13
MIR156a.2		+47, +73	+13, +1224
MIR156a.3			+13, +397, +749, +841
MIR156a.4		+47, +73, +83	-55, +13

MIR156b.1		-383, +87, +698	+307, +872
MIR156b.2		-355, +77, +87, +600	
MIR156b.3			
MIR156b.4			
MIR156b.5		+85	+1002
MIR156b.6		+79	
MIR156c	-255	-797	-499
MIR156d	**-242**	**-420**	**-480**, +579, +957, +1240, +1352
MIR156e.1	-650	-1068, -173, +73, +83	+13
MIR156e.2	-667	-1089, -189, -179, +73	+13
MIR156e.3	-631	-1049, -149, +101	+41
MIR156f	-709		+808
MIR156g	**+299**	**-124**, **0**	**-106**, **-108**, +931
MIR156h.1		-88, -100, +14, +108	+38
MIR156h.2	**-231**	**-83**	**+38**, +741
MIR157a	+1125	-222, +94	+611
MIR157b.1		-62	+384, +1207, +1227, +1395
MIR157b.2		-67	+389, +1287, +1459
MIR157c			-808

MIR157d		+155	-316, +248, +1439
MIR157e			-830, +330, +961

5.6 Primers used in Chapter 3.

Primer ID	Use/Experiment	Sequence
		smRNA-qPCR
miR156_RT-F1	miR156 RT	GTTGGCTCTGGTGCAGGGTCCGAGGTATTC GCACCAGAGCCAACGTGCTCA
miR156_F1	miR156 qPCR	GCGGCGTTGACAGAAGAGAGT
miR156_R1	miR156 qPCR	TGGTGCAGGGTCCGAGGT
miR157_RT-F1	miR157 RT	GTTGGCTCTGGTGCACCCTCCGACCTATTC GCACCAGAGCCAACGTGCTCT
miR157_F1	miR157 qPCR	GCGGCGTTGACAGAAGATAGA
miR157_R1	miR157 qPCR	TGGTGCACCCTCCGACCT
miR159_RT_primer	miR159 RT	GTTGGCTCTGGTGCAGGGTCCGAGGTATTC GCACCAGAGCCAACTAGAGC
miR159_Forward	miR159 qPCR	GCGGCGTTTGGATTGAAGGGA
miR168_RT_primer	miR168 RT	GTTGGCTCTGGTGCAGGGTCCGAGGTATTC GCACCAGAGCCAACTTCCCG
miR168_Forward3	miR168 qPCR	AAGGCGTCGCTTGGTG
URP	miR168 and miR159 qPCR	GTGCAGGGTCCGAGGT
Aspec_miR156v3_RT-F1	miR156v3 RT	GTTGGCTCTGGTGCAGCCTTTGACCTTCGC GCACCAGAGCCAACATAATCAC
Aspec_miR156v2/3_F1	miR156v3 qPCR	ATGACAAGCGGCGTTGACAGAAT
Aspec_miR156v3_R1	miR156v3 qPCR	TGGTGCAGCCTTTGACCTTC
miR156_RNA	standard curves	rUrGrArCrArGrArArGrArGrArGrUrGrArGrCrArC
miR157_RNA	standard curves	rUrGrArCrArGrArArGrArUrArGrArGrArGrCrArC

Aspec_miR156v3_RNA	standard curves	rUrGrArCrArGrArArUrArGrArGrUrGrArUrUrArU
	mRNA-qPCR	
Apen-MIR156a_F1	used with A. penninervis and some other Acacias	GACTTTGAAGGGTAAGGGAA
Apen-MIR156a_R1	used with A. penninervis and some other Acacias	ATAGTTTCRAGCATAAACCATGC
Apen-MIR156b_F1	used with A. penninervis and some other Acacias	TCTRGGACAAGAGAGGGAGAG
Apen-MIR156b.1_R1	used with A. penninervis and some other Acacias	GAAAGAGTAGTGAGCACGCACC
Apen-MIR156b.2_R1	used with A. penninervis and some other Acacias	CCAAAAAGAAAGKCATTAGTATCG
Apen-MIR156c_F1	used with A. penninervis and some other Acacias	GACTCCTAAAGCTTGTTTCATGT
Apen-MIR156c_R1	used with A. penninervis and some other Acacias	GTGAGACTCGTTTCCAGCACT
Apen-MIR156d_F1	used with A. penninervis and some other Acacias	CTCTATGTTGCTTCTTTCTTGG
Apen-MIR156d_R1	used with A. penninervis and some other Acacias	GAGCACGCAAAAGCAATAGTATA
Apen-MIR156e_F1	used with A. penninervis and some other Acacias	GGGATGAAATTAGTAGGGAGATT
Apen-MIR156e_R1	used with A. penninervis and some other Acacias	CATAGCTTCAAGCACGAAACATGA

Apen-MIR156f_F1	used with A. penninervis and some other Acacias	GCGAACACACACACAGAGATTG
Apen-MIR156f_R1	used with A. penninervis and some other Acacias	GTGAGCACTCACCTGCAAATGTAC
Apen-MIR156g_F1	used with A. penninervis and some other Acacias	GGAAATTTGAAGAGATAAGCATTATG
Apen-MIR156g_R1	used with A. penninervis and some other Acacias	CCCAGCAACACTAAACACATG
Apen-MIR156h_F1	used with A. penninervis and some other Acacias	TAAGTTTCACAAGTGGTAAGTAGTGG
Apen-MIR156h_R1	used with A. penninervis and some other Acacias	AGAAGAGAGAAAGCACACAGCTT
Apen-MIR157a_F1	used with A. penninervis and some other Acacias	AAAGGCGATGATGGCGATGTT
Apen-MIR157a_R1	used with A. penninervis and some other Acacias	GTGAGATGCAATATTTATTGTGG
Apen-MIR157b_F1	used with A. penninervis and some other Acacias	AAGATGGGGGTGAAGACACAGAG
Apen-MIR157b_R1	used with A. penninervis and some other Acacias	CAGAGGTTGAAATGCACAAAGG
Apen-MIR157c_F1	used with A. penninervis and some other Acacias	TAGATAGATAGGAGGCACTTTGG
Apen-MIR157c_R1	used with A. penninervis and some other Acacias	TGAGATGTGATTGTATATACGCAAG

Apen-MIR157d_F1	used with A. penninervis and some other Acacias	AAGAGGATGAAGAAGAAAGGAG
Apen-MIR157d_R1	used with A. penninervis and some other Acacias	GAAGAGGGAGAAGGAGGCCACT
Apen-MIR157e_F1	used with A. penninervis and some other Acacias	CTGATTAGTGATGAATAGGTCTGTTTC
Apen-MIR157e_R1	used with A. penninervis and some other Acacias	ACTGCACAAACACAAACAAGGAT
Apen-SPL1-like1_F1	used with A. penninervis and some other Acacias	TCAATGGAACGAACATAGCCTC
Apen-SPL1-like1_R1	used with A. penninervis and some other Acacias	CATTATCTGTCTTGATGGAAGAAC
Apen-SPL1-like2_F1	used with A. penninervis and some other Acacias	ATGATTCCCCAATTGTTGATAG
Apen-SPL1-like2_R1	used with A. penninervis and some other Acacias	GCCTATGATAATCCTTAGCACTG
Apen-SPL3-like1_F1	used with A. penninervis and some other Acacias	GATGCGAAGCCGTATCATAAG
Apen-SPL3-like1_R1	used with A. penninervis and some other Acacias	CAAACTCACCTAACTCATGGAATC
Apen-SPL3-like2_F1	used with A. penninervis and some other Acacias	CATAAAGTGTGTGAGTTTCATTCC
Apen-SPL3-like2_R1	used with A. penninervis and some other Acacias	ATCGTTGTGAGATTCAGCATTG

Apen-SPL3-like3_F1	used with A. penninervis and some other Acacias	AGCACTGGCGAAATTTACAATG
Apen-SPL3-like3_R1	used with A. penninervis and some other Acacias	AGGAACGCTCATTTGAACTCTTC
Apen-SPL3-like4_F1	used with A. penninervis and some other Acacias	AAGCAGCGGTCGTACTCATTTC
Apen-SPL3-like4_R1	used with A. penninervis and some other Acacias	CTTGGTAATCAGATGAAACCTTTC
Apen-SPL3-like5_F1	used with A. penninervis and some other Acacias	TCAGGTGGAAAGCTGTGAGAG
Apen-SPL3-like5_R1	used with A. penninervis and some other Acacias	AGCCAGAAATGAGCACAGAAGC
Apen-SPL6-like1_F1	used with A. penninervis and some other Acacias	TCATAGTAAAAGAGTGGCTTCCTC
Apen-SPL6-like1_R1	used with A. penninervis and some other Acacias	GATAAATTGGGACTTTTAGATTGAAC
Apen-SPL6-like2_F1	used with A. penninervis and some other Acacias	GTTTGTAGATTTGGGGTTTCCAG
Apen-SPL6-like2_R1	used with A. penninervis and some other Acacias	TTAAGAGCAATAACATGAGTAGAGG
Apen-SPL6-like3_F1	used with A. penninervis and some other Acacias	TTCAACCTATGCATGAGTCGTCA
Apen-SPL6-like3_R1	used with A. penninervis and some other Acacias	ACCAGCAAGCAGAGATGTTCCT

Apen-SPL6-like4_F1	used with A. penninervis and some other Acacias	GCTATTCCTCTTTTGTGTTCTTCAG
Apen-SPL6-like4_R1	used with A. penninervis and some other Acacias	GACTTGGCAGTAAGAAGTCTGTG
Apen-SPL6-like5_F1	used with A. penninervis and some other Acacias	CTTTCTAACTCCGTGGTTGAATC
Apen-SPL6-like5_R1	used with A. penninervis and some other Acacias	CGATGATTCTGAAGAAGACAAAATAG
Apen-SPL7-like1_F1	used with A. penninervis and some other Acacias	AGCAGTGAGATAAATGACCTAGAAAC
Apen-SPL7-like1_R1	used with A. penninervis and some other Acacias	CAACAGAGGACACAAAACTATCATC
Apen-SPL7-like2_F1	used with A. penninervis and some other Acacias	CAGTGAGATAAACGACCTAGAAAT
Apen-SPL7-like2_R1	used with A. penninervis and some other Acacias	CAACAGAGGACACCAAACTATCATT
Apen-SPL8-like1_F1	used with A. penninervis and some other Acacias	TTCTAATGGTACCCAAGTCCGAAG
Apen-SPL8-like1_R1	used with A. penninervis and some other Acacias	TAGAGACGGTTAACGAGGTCGTC
Apen-SPL8-like2_F1	used with A. penninervis and some other Acacias	ACACTTCGCATTCCTTATTGC
Apen-SPL8-like2_R1	used with A. penninervis and some other Acacias	GGTCAAAGAAGTAGTGCTGGTG

Apen-SPL9-like1_F1	used with A. penninervis and some other Acacias	TATCTACTTTGAGGATGTGACTGTC
Apen-SPL9-like1_R1	used with A. penninervis and some other Acacias	AGAATGAAAGCCACAGACTTTATGC
Apen-SPL9-like2_F1	used with A. penninervis and some other Acacias	GATATACTTTGAGGATGTGGGAGTG
Apen-SPL9-like2_R1	used with A. penninervis and some other Acacias	GGAGTGCATACTACAGACTTTATGT
Apen-SPL10-like1_F1	used with A. penninervis and some other Acacias	AGAAGAGCCTGTTGGAAATTCTC
Apen-SPL10-like1_R1	used with A. penninervis and some other Acacias	GAATCCCAGAGTGAGATGTGTTC
Apen-SPL10-like2_F1	used with A. penninervis and some other Acacias	CATTATTTTGAGGACGCTAATGCTC
Apen-SPL10-like2_R1	used with A. penninervis and some other Acacias	CTAACACCACAACCTTCGCTG
Apen-SPL13-like1_F1	used with A. penninervis and some other Acacias	AGGGGAATTTTCTGTGGATTTG
Apen-SPL13-like1_R1	used with A. penninervis and some other Acacias	CTTTAGAGACTCCTCCAACAACATC
Apen-SPL13-like2_F1	used with A. penninervis and some other Acacias	TTGGGATTTGAGTGAATTAGATG
Apen-SPL13-like2_R1	used with A. penninervis and some other Acacias	ACTAGCAACCTGACCTAGCTTC

Apen-SPL13-like3_F1	used with A. penninervis and some other Acacias	AACTTTGACCTGGGAACTTTC
Apen-SPL13-like3_R1	used with A. penninervis and some other Acacias	TGGCAGAATCTCTGTTCCTGTC
Apen-SPL13-like4_F1	used with A. penninervis and some other Acacias	CGTTACTATTTAGTGGCGGTTTAG
Apen-SPL13-like4_R1	used with A. penninervis and some other Acacias	TCCAACAACATCTATCTTTCCAC
Apen-SPL13-like5_F1	used with A. penninervis and some other Acacias	GGATTTGAGTCTTGGAGAGATTTC
Apen-SPL13-like5_R1	used with A. penninervis and some other Acacias	CCTTAGGTGTATTTGAGGTTGATTT
Apen-SPL13-like6_F1	used with A. penninervis and some other Acacias	GTTGATGGATGCAATTCTGAC
Apen-SPL13-like6_R1	used with A. penninervis and some other Acacias	TTCATCAAACTCCTCCAATGAATG
Apen-SPL13-like7A_F1	used with A. penninervis and some other Acacias	AGTGTGTCGTGCTTGGTGTATG
Apen-SPL13-like7A_R1	used with A. penninervis and some other Acacias	AACTCCTCTTAACTCCATCGAAGTC
Apen-SPL13-like7B_F1	used with A. penninervis and some other Acacias	AGAGTCTGTTGTGCTCGGTGG
Apen-SPL13-like7B_R1	used with A. penninervis and some other Acacias	ACTCCTGTTAACTCCATCGAAC

	***Acacia-MIR156D* sequencing**	
Acacia_MIR156D_assay_F1	promoter sequencing, most species	TAAGTGCGCAATTCCCGAGAAGACAAAG
Acacia_MIR156D_assay_F3	promoter sequencing, problematic species (e.g. *A. neriifolia* clade)	GCCTTTTGCATGAGTCTGGCCTGGATT
Acacia_MIR156D_assay_R1	promoter sequencing; use with F1 or F3	GGGAACAAAGCAGAGCCAGAGAGCAACGAAATTG
Acacia_MIR156D_dup_F1	use with Acacia_MIR156D_assay_R1 to amplify 1st tandem duplicate	ATTGGGTGTATTTCTYCTATGAKAAGG
Acacia_MIR156D_dup_F2	use with Acacia_MIR156D_assay_R1 to amplify 1st tandem duplicate	GGAGGTYACGCTGAGCTTTGGAATGGAG
Acacia_MIR156D_dup_F3	use with Acacia_MIR156D_assay_R1 to amplify 1st tandem duplicate	AGCACTSAGCACACCTGGACTGAGGAGTTGG
Acacia_MIR156D_dup_F4	use with Acacia_MIR156D_assay_R1 to amplify 1st tandem duplicate	AATGATGATTTGGTTAKTGTGACTTGATGAATG
	ddRAD-seq	
GCATG_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATG
AACCA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCA
TCGAT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
TGCAT_MfeI_P	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATC

1.1		TTGCAT
GGTTG_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTG
AAGGA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGA
AGCTA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTA
ACACA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACACA
AATTA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTA
ACGGT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGGT
ACTGG_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGG
ATACG_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATACG
ATGAG_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAG
CATAT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATAT
CGAAT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAAT
CGGCT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGCT
CGGTA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTA
GCCGT_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCGT
GCTGA_MfeI_P1.1	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA
GGATA_MfeI_P	5’ adapter with P1.2	ACACTCTTTCCCTACACGACGCTCTTCCGATC

1.1		TGGATA
GCATG_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTCATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGCAT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTATGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGTTG_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTCAACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AGCTA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTAGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACACA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AATTA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTAATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACGGT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTACCGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTGG_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTCCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATACG_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTCGTATAGATCGGAAGAGCGTCGT

		GTAGGGAAAGAGTGT
ATGAG_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTCTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CATAT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTATATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGAAT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTATTCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGCT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTAGCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGTA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCCGT_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTACGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCTGA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGATA_MfeI_P1.2	5’ adapter with P1.1	/5Phos/AATTTATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
HindIII_P2.1	3’ adapter with P2.2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
HindIII_P2.2	3’ adapter with P2.1	/5Phos/AGCTAGATCGGAAGAGCGAGAACAA
PCR1	post adapter-ligation PCR	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG
PCR2_Idx_1_ATCACG	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_2_CGATGT	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC

PCR2_Idx_3_TTAGGC	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_4_TGACCA	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_5_ACAGTG	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_6_GCCAAT	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_7_CAGATC	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_8_ACTTGA	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_9_GATCAG	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_10_TAGCTT	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_11_GGCTAC	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_12_CTTGTA	post adapter-ligation PCR	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGC

5.7. Full list of samples used for ddRAD-seq and targeted sequencing, Chapter 3.

Genus	Species	Seed Source	Stock number	Phylogeny ID	DNA sample ID
Acacia	acuminata	Nindethana Seed Service	NS-31988	A. acuminata NS	Aacu1
Acacia	adunca	Nindethana Seed Service	NS-40978	A. adunca NS	Aadu1
Acacia	aestivalis	Nindethana Seed Service	NS-1494	A. aestivalis NS	Aaes1
Acacia	alata	Nindethana Seed Service	NS-22285	A. alata NS	Aala1
Acacia	alcockii	Nindethana Seed Service	NS-17116	A. alcockii NS	Aalc1
Acacia	amoena	Nindethana Seed Service	NS-24426	A. amoena NS	Aamo1
Acacia	anceps	Nindethana Seed Service	NS-23332	A. anceps NS	Aance1
Acacia	aneura	Australian Tree Seed Center	ATSC-19943	A. aneura ATSC	Aane1
Acacia	aneura	Desert Legume Program	xDL 96-0052	A. aneura xDL	Aane2
Acacia	angusta	Nindethana Seed Service	NS-3712	A. angusta NS	Aang1
Acacia	applanata	Nindethana Seed Service	NS-39663	A. applanata NS	Aapp1
Acacia	aphylla	Nindethana Seed Service	NS-24123	A. aphylla NS	Aaph1
Acacia	baileyana	Nindethana Seed Service	NS-100	A. baileyana NS	Abail2
Acacia	baileyana	Seedman	D1164	A. baileyana SM	Abail3
Acacia	bancroftii	Nindethana Seed Service	unknown	A. bancroftii NS	Aban2

Acacia	barringtonensis	Nindethana Seed Service	NS-17615	A. barringtonensis NS	Abarr1
Acacia	biflora	Nindethana Seed Service	NS-27074	A. biflora NS	Abif1
Acacia	binata	Nindethana Seed Service	unknown	A. binata NS	Abin1
Acacia	binervata	Australian Tree Seed Center	ATSC-16245	A. binervata ATSC	Abine2
Acacia	blakelyi	Australian Tree Seed Center	ATSC-17639	A. blakelyi ATSC	Ablak9
Acacia	blakelyi	Nindethana Seed Service	NS-43408	A. blakelyi NS	Ablak5
Acacia	blayana	Australian Tree Seed Center	ATSC-16256	A. blayana ATSC	Ablay1
Acacia	blayana	Nindethana Seed Service	NS-37823	A. blayana NS	Ablay5
Acacia	browniana	Nindethana Seed Service	NS-12556	A. browniana NS	Abrow1
Acacia	brumalis	Nindethana Seed Service	NS-34503	A. brumalis NS2	Abrum3
Acacia	brumalis	Nindethana Seed Service	NS-28794	A. brumalis NS1	Abrum1
Acacia	buxifolia	Australian Tree Seed Center	ATSC-14412	A. buxifolia ATSC	Abux2
Acacia	caesiella	Nindethana Seed Service	NS-11913	A. caesiella NS	Acae1
Acacia	calamifolia	Nindethana Seed Service	NS-33685	A. calamifolia NS	Acala1
Acacia	cardiophylla	Nindethana Seed Service	NS-22391	A. cardiophylla NS	Acard2
Acacia	celastrifolia	Nindethana Seed Service	NS-12415	A. celastrifolia NS	Acela2

Acacia	chamaeleon	Nindethana Seed Service	NS-22053	A. chamaeleon NS	Acham1
Acacia	chinchillensis	Nindethana Seed Service	NS-43999	A. chinchillensis NS	Achin1
Acacia	chrysotricha	Australian Tree Seed Center	ATSC-18620	A. chrysotricha ATSC	Achry1
Acacia	complanata	Nindethana Seed Service	NS-43289	A. complanata NS	Acomp1
Acacia	confusa	Asklepios-seeds		A. confusa Ask	Acon1
Acacia	covenyi	Nindethana Seed Service	NS-42389	A. covenyi NS	Acov1
Acacia	crassicarpa	Australian Tree Seed Center	ATSC-19726	A. crassicarpa ATSC1	Acra2
Acacia	crassicarpa	Australian Tree Seed Center	ATSC-19724	A. crassicarpa ATSC2	Acra3.RX
Acacia	chrysella	Nindethana Seed Service	NS-502	A. chrysella NS	Achry8
Acacia	cultriformis	Nindethana Seed Service	NS-36528	A. cultriformis NS2	AcultR
Acacia	cultriformis	Nindethana Seed Service	NS-44314	A. cultriformis NS1	Acult5
Acacia	cultriformis	Australian Tree Seed Center	ATSC-13625	A. cultriformis ATSC	Acult2
Acacia	dangarensis	Australian Tree Seed Center	ATSC-18608	A. dangarensis ATSC	Adang1
Acacia	daphnifolia	Nindethana Seed Service	NS-42493	A. daphnifolia NS	Adaph1
Acacia	dawsonii	Nindethana Seed Service	NS-3725	A. dawsonii NS	Adaw1
Acacia	dealbata	Desert Legume Program	xDL 89-0388	A. dealbata xDL	Adeal1
Acacia	dealbata ssp. dealbata	Australian Tree Seed Center	ATSC-19076	A. dealbata ATSC	Adeal5

Acacia	dealbata	D3 autogamy		A. dealbata D3	Adeal4
Acacia	deanei ssp. deanei	Australian Tree Seed Center	ATSC-16922	A. deanei ATSC	Adean1
Acacia	deanei	Nindethana Seed Service	NS-44135	A. deanei NS	Adean7
Acacia	debilis	Nindethana Seed Service	NS-15296	A. debilis NS	Adebi1
Acacia	decora	Nindethana Seed Service	NS-34045	A. decora NS	Adeco2
Acacia	decurrens	Australian Tree Seed Center	ATSC-19764	A. decurrens ATSC	Adec3, Adec8
Acacia	decurrens	Nindethana Seed Service	NS-42143	A. decurrens NS	Adec10
Acacia	delphina	Nindethana Seed Service	unknown	A. delphina NS	Adelph1
Acacia	difformis	Nindethana Seed Service	NS-24515	A. difformis NS	Adiff1
Acacia	drummondii (affinis)	Nindethana Seed Service	NS-32967	A. drummondii NS	Adrum3
Acacia	elata	Australian Tree Seed Center	ATSC-18243	A. elata ATSC2	Aelat2
Acacia	elata	Australian Tree Seed Center	ATSC-15831	A. elata ATSC1	Aelat1
Acacia	elongata	Australian Tree Seed Center	ATSC-09689	A. elongata ATSC	Aelon1
Acacia	empelioclada	Nindethana Seed Service	NS-22309	A. empelioclada NS	Aemp1
Acacia	eriopoda	Australian Tree Seed Center	ATSC-19051	A. eriopoda ATSC	Aeri1
Acacia	euthycarpa	Nindethana Seed Service	NS-26250	A. euthycarpa NS	Aeuth2
Acacia	everistii	Australian Tree Seed Center	ATSC-16654	A. everistii ATSC	Aever1

Acacia	falcata	Australian Tree Seed Center	ATSC-15554	A. falcata ATSC	Afalc1
Acacia	falcata	Nindethana Seed Service	NS-24875	A. falcata NS	Afalc5
Acacia	falciformis	Nindethana Seed Service	NS-43291	A. falciformis NS	Afalci1
Acacia	filicifolia	Nindethana Seed Service	NS-36451	A. filicifolia NS	Afili1
Acacia	filicifolia	Australian Tree Seed Center	ATSC-17893	A. filicifolia ATSC	Afili3
Acacia	fimbriata	Australian Tree Seed Center	ATSC-15472	A. fimbriata ATSC	Afimb2
Acacia	fimbriata	Nindethana Seed Service	NS-40977	A. fimbriata NS	Afimb1
Acacia	fulva	Australian Tree Seed Center	ATSC-18863	A. fulva ATSC	Afulv6
Acacia	gilbertii	Nindethana Seed Service	NS-40728	A. gilbertii NS	Agilb1
Acacia	gillii	Nindethana Seed Service	NS-7165	A. gillii NS	Agill1
Acacia	gittinsii	Nindethana Seed Service	NS-3707	A. gittinsii NS	Agitt1
Acacia	gladiiformis	Nindethana Seed Service	NS-17616	A. gladiiformis NS	Aglad1
Acacia	glaucocarpa	Australian Tree Seed Center	ATSC-17892	A. glaucocarpa ATSC	Aglauc1
Acacia	glaucoptera	Nindethana Seed Service	unknown	A. glaucoptera NS	Aglauc2
Acacia	gonophylla	Nindethana Seed Service	NS-32310	A. gonophylla NS	Agono1
Acacia	guinetii	Nindethana Seed Service	NS-12565	A. guinetii NS	Aguin1

Acacia	hakeoides	Australian Tree Seed Center	ATSC-16243	A. hakeoides ATSC	Ahake1
Acacia	harveyi	Nindethana Seed Service	NS	A. harveyi NS	Aharv1
Acacia	hemiteles	Desert Legume Program	xDL 96-0054	A. hemiteles xDL	Ahem2
Acacia	heterochroa ssp. Heterochroa	Nindethana Seed Service	NS-836	A. heterochroa NS	Ahet1
Acacia	hexaneura	Nindethana Seed Service	NS-17117	A. hexaneura NS	Ahex4
Acacia	implexa	Australian Tree Seed Center	ATSC-15832	A. implexa ATSC	Aimp1
Acacia	inaequilatera	Australian Tree Seed Center	ATSC-19193	A. inaequilatera ATSC	Aineq5
Acacia	incurva	Nindethana Seed Service	NS-38979	A. incurva NS	Ainc1
Acacia	intricata	Nindethana Seed Service	NS-17023	A. intricata NS	Aint1
Acacia	irrorata	Nindethana Seed Service	NS-40979	A. irrorata NS	Airr2
Acacia	jennerae	Nindethana Seed Service	NS-33644	A. jennerae NS	Ajenn1
Acacia	jucunda	Nindethana Seed Service	NS-15445	A. jucunda NS	Ajuc1
Acacia	jonesii	Nindethana Seed Service	NS-974	A. jonesii NS	Ajon3
Acacia	koa	unknown	unknown	A. koa	Akoa1
Acacia	kybeanensis	Nindethana Seed Service	NS-33091	A. kybeanensis NS	Akyb1
Acacia	latisepala	Nindethana Seed Service	NS-16844	A. latisepala NS	Alati1, Alati3
Acacia	leichhardtii	Nindethana Seed Service	NS-3702	A. leichhardtii NS	Aleich1

Acacia	leioderma	Nindethana Seed Service	NS-7566	A. leioderma NS	Aleio1
Acacia	leiophylla	Nindethana Seed Service	NS-22406	A. leiophylla NS	Aleiop2
Acacia	leucoclada ssp. leucoclada	Australian Tree Seed Center	ATSC-18621	A. leucoclada ATSC1	Aleuc3
Acacia	leucoclada	Nindethana Seed Service	NS-19620	A. leucoclada NS	Aleuc5
Acacia	leucoclada ssp. Argentifolia	Australian Tree Seed Center	ATSC-20802	A. leucoclada ATSC2	Aleuc8
Acacia	linearifolia	Nindethana Seed Service	NS-1551	A. linearifolia NS	Aline1
Acacia	lineolata	Nindethana Seed Service	NS-28796	A. lineolata NS	Alineo1
Acacia	linifolia	Nindethana Seed Service	NS-31598	A. linifolia NS	Alini1
Acacia	longifolia var. sophorae	Australian Tree Seed Center	ATSC-15849	A. longifolia ATSC	Along1
Acacia	mabellae	Australian Tree Seed Center	ATSC-14400	A. mabellae ATSC	Amab3
Acacia	macradenia	Australian Tree Seed Center	ATSC-16662	A. macradenia ATSC	Amac1
Acacia	macradenia	Nindethana Seed Service	NS-37071	A. macradenia NS	Amac2
Acacia	mangium	Australian Tree Seed Center	ATSC-21072	A. mangium ATSC	Aman1
Acacia	mearnsii	Australian Tree Seed Center	ATSC-20446	A. mearnsii ATSC1	Amear3
Acacia	mearnsii	Desert Legume Program	xDL 89-0389	A. mearnsii xDL	Amear4
Acacia	meisneri	Nindethana Seed Service	NS-32829	A. meisneri NS	Ameis1

Acacia	melanoxylon	Australian Tree Seed Center	ATSC-14176	A. melanoxylon ATSC1	Amel2
Acacia	melanoxylon	Australian Tree Seed Center	ATSC-21112	A. melanoxylon ATSC2	Amel1
Acacia	melanoxylon	Australian Tree Seed Center	ATSC-19505	A. melanoxylon ATSC3	Amel3
Acacia	microbotrya	Nindethana Seed Service	NS-22432	A. microbotrya NS	Amic1
Acacia	midgleyi	Australian Tree Seed Center	ATSC-18358	A. midgleyi ATSC	Amidg1
Acacia	mollifolia	Australian Tree Seed Center	ATSC-16250	A. mollifolia ATSC	Amoll5
Acacia	mollifolia	Nindethana Seed Service	NS-22444	A. mollifolia NS	Amoll4
Acacia	muelleriana	Australian Tree Seed Center	ATSC-17028	A. muelleriana ATSC	Amuel5
Acacia	muelleriana	Nindethana Seed Service	NS-4760	A. muelleriana NS	Amuel1
Acacia	murrayana	Australian Tree Seed Center	ATSC-20378	A. murrayana ATSC	Amurr2
Acacia	myrtifolia	Nindethana Seed Service	NS-33945	A. myrtifolia NS2	Amyrt4
Acacia	myrtifolia	Nindethana Seed Service	NS-2682	A. myrtifolia NS1	Amyrt5
Acacia	myrtifolia	Nindethana Seed Service	NS-28053	A. myrtifolia NS3	Amyrt3
Acacia	nanodealbata	Nindethana Seed Service	NS-29412	A. nanodealbata NS	Anano2
Acacia	nematophylla	Nindethana Seed Service	NS-10656	A. nematophylla NS	Anema1
Acacia	neriifolia	Nindethana Seed Service	NS-945	A. neriifolia NS2	Aneri5
Acacia	neriifolia	Nindethana Seed Service	NS-39421	A. neriifolia NS1	Aneri6

Acacia	nigricans	Nindethana Seed Service	NS-2315	A. nigricans NS	Anig1
Acacia	notabilis	Nindethana Seed Service	NS-16575	A. notabilis NS	Anota1
Acacia	obliquinervia	Australian Tree Seed Center	ATSC-16273	A. obliquinervia ATSC	Aobl1
Acacia	obtusata	Australian Tree Seed Center	ATSC-6867	A. obtusata ATSC	Aobt5
Acacia	obtusata	Nindethana Seed Service	NS-24433	A. obtusata NS	Aobt2
Acacia	olsenii	Nindethana Seed Service	NS-11750	A. olsenii NS	Aolse1
Acacia	parramattensis	Australian Tree Seed Center	ATSC-18259	A. parramattensis ATSC	Aparr2
Acacia	parvipinnula	Australian Tree Seed Center	ATSC-15844	A. parvipinnula ATSC	Aparv4
Acacia	parvipinnula	Nindethana Seed Service	NS-41387	A. parvipinnula NS	Aparv3
Acacia	pendula	Longwood Gardens	2006-0336*A	A. pendula LG	Apend1
Acacia	penninervis	Australian Tree Seed Center	ATSC-15592	A. penninervis ATSC1	Apen9
Acacia	penninervis	Australian Tree Seed Center	ATSC-15595	A. penninervis ATSC2	Apen4
Acacia	penninervis	Nindethana Seed Service	NS-25185	A. penninervis NS	Apen8
Acacia	pentadenia	Nindethana Seed Service	NS-14849	A. pentadenia NS	Apenta1
Acacia	phlebopetala	Nindethana Seed Service	NS-11843	A. phlebopetala NS	Aphl1
Acacia	podalyriifolia	Australian Tree Seed Center	ATSC-14737	A. podalyriifolia ATSC	Apod8
Acacia	podalyriifolia	Nindethana Seed Service	NS-35621	A. podalyriifolia NS	Apod9

Acacia	polybotrya	Australian Tree Seed Center	ATSC-14741	A. polybotrya ATSC	Apoly7
Acacia	polybotrya	Nindethana Seed Service	NS-37771	A. polybotrya NS	Apoly6
Acacia	pravissima	Nindethana Seed Service	NS-34166	A. pravissima NS	Aprav3
Acacia	pravissima	Australian Tree Seed Center	ATSC-16874	A. pravissima ATSC	Aprav2
Acacia	pruinosa	Australian Tree Seed Center	ATSC-18066	A. pruinosa ATSC	Apru2
Acacia	pruinosa	Nindethana Seed Service	NS-674	A. pruinosa NS	Apru1
Acacia	pulchella	Desert Legume Program	xDL 97-0012	A. pulchella xDL	Apul3
Acacia	pulchella var. glaberrima	Nindethana Seed Service	NS-14279	A. pulchella NS1	Apul1
Acacia	pulchella var. pulchella	Nindethana Seed Service	NS-40399	A. pulchella NS2	Apul5
Acacia	pustula	Australian Tree Seed Center	ATSC-15474	A. pustula ATSC	Apust2
Acacia	pustula	Nindethana Seed Service	NS-22489	A. pustula NS	Apust1
Acacia	pycnantha	Australian Tree Seed Center	ATSC-19346	A. pycnantha ATSC	Apycn2
Acacia	pyrifolia	Desert Legume Program	XDL 96-0040	A. pyrifolia xDL	Apyr1
Acacia	pyrifolia	Australian Tree Seed Center	ATSC-19203	A. pyrifolia ATSC	Apyr4
Acacia	redolens	Desert Legume Program	xDL 99-0100	A. redolens xDL	Ared1
Acacia	retinodes	Nindethana Seed Service	NS-26698	A. retinodes NS2	Areti3
Acacia	retinodes	Nindethana Seed Service	ns-31578	A. retinodes NS1	Areti1

Acacia	rivalis	Nindethana Seed Service	NS-10457	A. rivalis NS	Ariv1
Acacia	rubida	Nindethana Seed Service	NS-43048	A. rubida NS2	Arub3
Acacia	rubida	Nindethana Seed Service	NS-33844	A. rubida NS1	ArubR
Acacia	sabulosa	Australian Tree Seed Center	ATSC-19186	A. sabulosa ATSC	Asab1
Acacia	sabulosa	Nindethana Seed Service	NS-22515	A. sabulosa NS	Asab2
Acacia	saliciformis	Nindethana Seed Service	NS-35091	A. saliciformis NS	Asali1
Acacia	saligna	Australian Tree Seed Center	ATSC-15795	A. saligna ATSC	Asal1
Acacia	saligna	Nindethana Seed Service	NS-30512	A. saligna NS	Asal2
Acacia	schinoides	Australian Tree Seed Center	ATSC-18258	A. schinoides ATSC	Aschi1
Acacia	semirigida	Nindethana Seed Service	NS	A. semirigida NS	Asemi1
Acacia	silvestris	Australian Tree Seed Center	ATSC-15852	A. silvestris ATSC	Asylv1
Acacia	silvestris	Nindethana Seed Service	NS-7734	A. silvestris NS	Asilv3
Acacia	spectabilis	Australian Tree Seed Center	ATSC-15555	A. spectabilis ATSC1	AspecR
Acacia	spectabilis	Nindethana Seed Service	NS-30777	A. spectabilis NS1	Aspec10
Acacia	spectabilis	Nindethana Seed Service	NS-40980	A. spectabilis NS2	Aspec7
Acacia	spectabilis	Australian Tree Seed Center	ATSC-17032	A. spectabilis ATSC2	Aspec11
Acacia	steedmanii	Nindethana Seed Service	NS	A. steedmanii NS	Asteed2

Acacia	subulata	Nindethana Seed Service	NS-41388	A. subulata NS	Asubu1
Acacia	terminalis	Nindethana Seed Service	NS-28726	A. terminalis NS	Aterm4
Acacia	terminalis	Australian Tree Seed Center	ATSC-16248	A. terminalis ATSC	Aterm5
Acacia	tetragonophylla	Nindethana Seed Service	NS-17655	A. tetragonophylla NS	Atet1
Acacia	tindaleae	Nindethana Seed Service	NS-22528	A. tindaleae NS	Atind1
Acacia	trachyphloia	Australian Tree Seed Center	ATSC-16249	A. trachyphloia ATSC	Atrac1
Acacia	trachyphloia	Nindethana Seed Service	NS-4770	A. trachyphloia NS	Atrach2
Acacia	tumida var. tumida	Australian Tree Seed Center	ATSC-20302	A. tumida ATSC	Atum2
Acacia	uncinata	Nindethana Seed Service	NS-22567	A. uncinata NS	Aunc2
Acacia	uncinata	Australian Tree Seed Center	ATSC-15968	A. uncinata ATSC	Aunc3
Acacia	validinervia	Nindethana Seed Service	NS-9446	A. validinervia NS	Aval2
Acacia	verticillata	Australian Tree Seed Center	ATSC-9969	A. verticillata ATSC	Avert1
Acacia	vestita	Australian Tree Seed Center	ATSC-16251	A. vestita ATSC	Avest1
Acacia	victoriae	Australian Tree Seed Center	ATSC-19325	A. victoriae ATSC1	Avic9
Acacia	victoriae	Australian Tree Seed Center	ATSC-18654	A. victoriae ATSC2	Avic3
Acacia	wattsiana	Nindethana Seed Service	N/C-03539	A. wattsiana NS	Awatt1

Faidherbia	albida	Desert Legume Program	xDL 91-0047	F. albida xDL1	Falb6
Paraserianthes	falcataria	Australian Tree Seed Center	ATSC-19109	P. falcataria ATSC	Pfal3
Paraserianthes	lophantha ssp. lophantha	Australian Tree Seed Center	ATSC-17678	P. lophantha ATSC	PlopR

6. BIBLIOGRAPHY

Acacia rubida (2018, September 4) Retrieved from http://worldwidewattle.com/speciesgallery/rubida.php.

Allsopp, A. (1953) Experimental and analytical studies of pteridophytes XXI. Investigations on Marsilea 3. The effect of various sugars on development and morphology. *Annuals of Botany* 17, 447-463.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology 215*, 403-410.

Atkinson, I.A.E., and Greenwood, R.M. (1989) Relationships between moas and plants. *New Zealand Journal of Ecology* 12, 67-96.

Barton, K.E. and Koricheva, J. (2010) The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. *American Naturalist* 175, 481-493.

Barton, K.E., and Boege, K. (2017). Future directions in the ontogeny of plant defence: understanding the evolutionary causes and consequences. *Ecology Letters* 20, 403-411.

Belt, T. (1874). The Naturalist in Nicaragua. First Edition (London: J.M. Dent & Sons).

Beydler, B., Osadchuk, K., Cheng, C., Manak, J.R., Irish, E.E. (2016). The juvenile phase of maize sees upregulation of stress-response genes and is extended by exogenous jasmonic acid. *Plant Physiology* 171, 2648-2658.

Boege, K., and Marquis, R.J. (2005) Facing herbivory as you grow up: the ontogeny of resistance in plants. *Trends in Ecology and Evolution* 20, 441-448.

Box, M.S., and Glover, B.J. (2010) A plant developmentalist's guide to paedomorphosis: reintroducing a classic concept to a new generation. *Trends in Plant Science* 15, 241-246.

Brown, G.K., Ariati, R.A., Murphy, D.J., Miller, J.T., Ladiges, P.Y. (2006) Bipinnate acacia (Acacia subg. Phyllodineae sect. Botrycephalae) of eastern Australia are polyphyletic based on DNA sequence data. *Australian Systematic Botany* 19, 315-326.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.

Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S., Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research* 18, 188-196.

Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research* 33:e179.

Chen, Y.T., Shen, C.H., Lin, W.D., Chu, H.A., Huang, B.L., Kuo, C.I., Yeh, K.W., Huang, L.C., and Chang, I.F. (2013) Small RNAs of *Sequoia sempervirens* during rejuvenation and phase change. *Plant Biology* 15, 27-36.

Cho, S.H., Coruh, C., and Axtell, M.J. (2012) miR156 and miR390 regulate tasiRNA accumulation and developmental timing in *Physcomitrella patens*. *The Plant Cell* 24, 4837-4849.

Choi, K., Kim, J., Müller, S. Y., Oh, M., Underwood, C., Henderson, I., et al. (2016). Regulation of microRNA-mediated developmental changes by the SWR1 chromatin remodeling complex in *Arabidopsis thaliana*. *Plant Physiology* 171, 1128-1143.

Chuck, G., Cigan, A.M., Saeteurn, K., Hake, S. (2007). The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nature Genetics* 39, 544-549.

Clark, L.L., and Burns, K.C. (2015) The ontogeny of leaf spines: progressive versus retrogressive heteroblasty in two New Zealand plant species. *New Zealand Journal of Botany* 53, 15-23.

Crawley, M.J. (1983) Herbivory: the dynamics of animal-plant interactions. Blackwell Scientific Publications, Oxford.

Cuperus, J.T., Fahlgren, N., and Carrington, J.C. (2011) Evolution and functional diversification of *MIRNA* genes. *The Plant Cell* 23, 431-442.

De Beer, G.R. (1951) Embryos and Ancestors. Oxford: Clarendon Press.

Eaton, D.A.R. (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844-1849.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A., Paz-Ares, J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics* 39, 1033-1037.

Farrell, T.P. and Ashton, D.H. (1978) Australian Journal of Botany 26, 365-379.

Goebel, K. (1889) Ueber die Jungendzustände der Pflanzen. *Flora* 72, 1-45.

Goebel, K. (1900) Organography of plants Part I. General organography. (English translation by I. B. Balfour). Clarendon Press, Oxford.

Goebel, K. (1908) Einleitung in die experimentelle morphologie der Pflanzen. B.G. Teubner, Leipzig.

Gomez-Acevedo, S., Rico-Arce, L., Delgado-Salinas, A., Magallon, S., Eguiarte, L.E. (2010). Neotropical mutualism between Acacia and Pseudomyrmex: phylogeny and divergence times. *Molecular Phylogenetics and Evolution* 56, 393-408.

Gonzalez-Teuber, M., Kaltenpoth, M., Boland, W. (2014). Mutualistic ants as an indirect defence against leaf pathogens. *New Phytologist* 202, 640-650.

Gould, S.J. (1977) Ontogeny and phylogeny. Harvard University Press.

Gould, S.J. (1992) Ontogeny and phylogeny revisted and reunited. BioEssays 14, 274-279.

Griffiths-Jones, S., Grocock, R.J., van Dogen, S., Bateman, A., Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* 34, D140-144.

Guerriero, G., Martin, N., Golovko, A., Sundstrom, J.F., Rask, L., Ezcurra, I. (2009) The RY/Sph element mediates transcriptional respression of maturation genes from late maturation to early seedling growth. *New Phytologist* 184, 552-565.

He, J. (2017) Developmental Functions Of Mir156 And Mir157 In Arabidopsis Publicly Accessible Penn Dissertations. 2338. https://repository.upenn.edu/edissertations/2338.

He, J., Xu, M., Willmann, M.R., McCormick, K., Hu, T., Yang, L., Starker, C.G., Voytas, D.F., Meyers, B.C., Poethig, R.S. (2018). Threshold-dependent repression of *SPL* gene expression by miR156/miR157 controls vegetative phase change in *Arabidopsis thaliana. PLoS Genetics* 14, e1007337.

Heil, M., and McKey, D. (2003) Protective ant-plant interactions as model systems in ecological and evolutionary research. *Annual Review of Ecology, Evolution, and Systematics* 34, 425-453.

Heil, M., Greiner, S., Meimberg, H., Kruger, R., Noyer, J., Heubl, G., Linsenmair, K.E., Boland, W. (2004). Evolutionary change from induced to constitutive expression of an indirect plant resistance. *Nature* 430, 205-208.

Hibara, K., Isono, M., Mimura, M., Sentoku, N., Kojima, M., Sakakibara, H., Kitomi, Y., Yoshikawa, T., Itoh, J., Nagato, Y. (2016). Jasmonate regulates juvenile-to-adult phase transition in rice. *Development* 143, 3407-3416.

Hildebrand, F. (1875) Ueber die Jungendzustände solcher Pflanzen, welche im Alter vom vegetativen Charakter ihrer Verwandten abweichen. *Flora* 21, 321-330.

Hocking, B. (1970). Insect associations with the swollen thorn acacias. *Transactions of the Royal Entomological Society of London* 122, 211-255.

Holeski, L.M., Kearsley, M.J.C., and Whitham, T.G. (2009) Separating ontogenetic and environmental determination of resistance to herbivory in cottonwood. *Ecology* 90, 2969-2973.

James, S.A. and Bell, D.T. (2000) Leaf orientation, light interception and stomatal conductance of *Eucalyptus globulus* ssp. *globulus* leaves. *Tree Physiology* 20, 815-823.

Janzen, D.H. (1966). Coevolution of mutualism between ants and acacias in central America. *Evolution* 20, 249-275.

Janzen, D.H. (1967). Interaction of the bull's-horn acacia (Acacia cornigera L.) with an ant inhabitant (*Pseudomyrmex ferruginea* F. Smith) in eastern Mexico. *The University of Kansas Science Bulletin* 47, 315-558.

Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T (2014) Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Research* 24:1384-95.

Kaplan, D.R. (1980) Heteroblastic leaf development in *Acacia*. Morphological and morphogenetic implications. *Cellule* 73, 137-203.

Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution 30*, 772–780.

Keller, A., Forster, F., Muller, T., Dandekar, T., Schultz, J., Wolf, M. (2010) Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. *Biology Direct* 5, 4.

Keyte, A.L., and Smith, K.K. (2014) Heterochrony and developmental timing mechanisms: changing ontogenies in evolution. *Seminars in Cell and Developmental Biology* 0, 99-107.

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics* 5, 59.

Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research* 20:265–72.

Livak, K.J., Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25, 402-408.

Loney, P.E., McArthur, C., Potts, B.M., Jordan, G.J. (2006) How does ontogeny in a *Eucalyptus* species affect patterns of herbivory by Brushtail Possums? *Functional Ecology* 20, 982-988.

Lorenz, R., Bernhart, S.H., Hoener zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., Hofacker, R.F. (2011). ViennaRNA Package 2.0. *Algorithms for Molecular Biology* 6, 26.

Magoc, T. and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957-2963.

Manceau, M., Domingues, V.S., Linnen, C.R., Rosenblum, E.B., Hoekstra, H.E. (2010) Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B* 365, 2439-2450.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10-12.

Massad, T.J. (2013) Ontogenetic differences of herbivory on woody and herbaceous plants: a meta-analysis demonstrating unique effects of herbivory on the young and the old, the slow and the fast. *Oecologia* 172, 1-10.

Mayer, V.E., Frederickson, M.E., McKey, D., and Blatrix, R. (2014) Current issues in the evolutionary ecology of ant-plant symbioses. *New Phytologist* 202, 749-764.

Miller, J.T., Andrew, R., and Bayer, R.T. (2003) Molecular phylogenetics of the Australian acacias of subg. Phyllodineae (Fabaceae:Mimosoideae) based on the trnK intron. *Australian Journal of Botany* 51, 167-177.

Mitchell-Olds, T. (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology and Evolution* 16, 693-700.

Murphy, D.J., Miller, J.T., Bayer, R.J., and Ladiges, P.Y. (2003) Molecular phylogeny of Acacia subgenus Phyllodineae (Mimosoideae: Leguminosae) based on DNA sequences of the internal transcribed spacer region. Australian Systematic Botany 16, 19-26.

Murphy, D.J. (2008) A review of the classification of Acacia (Leguminosae, Mimosoideae). *Muelleria* 26, 10-26.

Obeso, J.R. (1997) The induction of spinescence in european holly leaves by browsing ungulates. *Plant Ecology* 129, 149-156.

O'Connell J, O Schulz-Trieglaff, E Carlson, MM Hims, NA Gormley, AJ Cox (2015) NxTrim: optimized trimming of Illumina mate pair reads. *Bioinformatics* 31:2035-2037.

Orchard, A.E. (2001) Flora of Australia. Volume 11A: Mimosaceae, Acacia Part 1. CSIRO Publishing, Australia.

Paradis, E., Claude, J., Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289-290.

Parra G, Bradnam K, Ning Z, Keane T, and Korf I (2009) Assessing the gene space in draft genomes. *Nucleic Acids Research* 37:298-297.

Pedley, L. (1986) Derivation and dispersal of Acacia (Leguminosae), with particular reference to Australia, and the recognition of *Senegalia* and *Racosperma*. *Botanical Journal of the Linnean Society* 92, 219-254.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One* 7, e37135.

Pico, S., Ortiz-Marchena, M.I., Merini, W., and Calonje, M. (2015) Deciphering the role of POLYCOMB REPRESSIVE COMPLEX1 variants in regulating the acquisition of flowering competence in Arabidopsis. *Plant Physiology* 168, 1286-1297.

Prud'homme, B., Gompel, N., and Carroll, S.B. (2007) Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences USA* 104, 8605-8612.

Poethig, R.S. (2013). Vegetative phase change and shoot maturation in plants. *Current Topics in Developmental Biology* 105, 125-152.

Quintero, C., Barton, K.E., and Boege, K. (2013) The ontogeny of plant indirect defenses. *Perspectives in Plant Ecology and Evolution* 15, 245-254.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raff, R.A., and Wray, G.A. (1989) Heterochrony: developmental mechanisms and evolutionary results. *Journal of Evolutionary Biology* 2, 409-434.

Razo-Belman, R., Molina-Torres, J., Martinez, O., Heil, M. (2017). Plant-ants use resistance-related plant odours to assess host quality before colony founding. *Journal of Ecology* 106, 379-390.

Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., Ma, H., and Qi, J. (2018) Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Molecular Plant* 11, 414-428.

Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3, 217-223.

Revell, L.J. (2013) Two new graphical methods for mapping trait evolution on phylogenies. *Methods in Ecology and Evolution* 4, 754-759.

Schmieder R and Edwards R (2003) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863-864.

Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental Cell* 8, 517-527.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, and Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210-3212.

Smith, K.K. (2011) Time's arrow: heterochrony and the evolution of development. International Journal of Developmental Biology 47, 613-621.

Stamatakis, A. (2014). RaxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312-1313.

Supnick, M. (1983) On the function of leaf spines in *Ilex opaca*. *Bulletin Torrey Botany Club* 110, 228-230.

Swaminathan, K., Peterson, K., Jack, T. (2008) The plant B3 superfamily. *Trends in Plant Science* 13, 647-655.

Tang F, Hajkova P, Barton SC, Lao K, Surani MA (2006) MicroRNA expression profiling of single whole embryonic stem cells. *Nucleic Acids Research* 34:e9.

Telford, M.J., Wise, M.J., Gowri-Shankar, V.G. (2005) Consideration of RNA secondary structure significantly improves likelihood-based estimates of phylogeny: examples from the bilateria. *Molecular Biology and Evolution* 22, 1129-1136.

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E.F., Hellens, R.P. (2007). Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* 3, 12.

Voss, S.R., and Smith, J.J. (2005) Evolution of salamander life cycles: a major-effect quantitative trait locus contributes to discrete and continuous variation for metamorphic timing. *Genetics* 170, 275-281.

Wang, J., Park, M.Y., Wang, L., Koo, Y., Chen, X., Weigel, D., Poethig, R.S. (2011). MiRNA control of vegetative phase change in trees. *PloS Genetics* 7, e1002012.

Wheeler, W.M. (1942). Studies of neotropical ant-plants and their ants. *Bulletin of the Museum of Comparative Zoology at Harvard College* 90, 3-262.

Wickham, H (2016). ggplot2: Elegant Graphics for Data Analysis, First Edition (New York: Springer-Verlag).

Wojciechowski, M.F., Lavin, M., Sanderson, M.J. (2004). A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gen resolves many well-supported subclades within the family. *American Journal of Botany* 91, 1846-1862.

Wu, G., and Poethig, R.S. (2006). Temporal regulation of shoot development in *Arabidopsis thaliana* by *miR156* and its target *SPL3*. *Development* 133, 3539-3547.

Wu, G., Park, M.Y., Conway, S.R., Wang, J., Weigel, D., Poethig, R.S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138, 750-759.

Xie, K., Shen, J., Hou, X., Yao, J., Li, X., Xiao, J., and Xiong, L. (2012) Gradual increase of miR156 regulates temporal expression changes of numerous genes during leaf development in rice. *Plant Physiology* 158, 1382-1394.

Xu, M., Hu, T., Smith, M. R. and R. S. Poethig (2016a). Epigenetic regulation of vegetative phase change in *Arabidopsis*. *The Plant Cell* 28, 28-41.

Xu, M., Hu, T., Zhao, J., Park, M., Earley, K.W., Wu, G., Yang, L., Poethig, R.S. (2016b). Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes in *Arabidopsis thaliana*. *PLoS Genetics* 12, e1006263.

Xu, Y., Guo, C., Zhou, B., Li, C., Wang, H., Zheng, B., et al. (2016c). Regulation of vegetative phase change by SWI2/SNF2 chromatin remodeling ATPase BRAHMA. *Plant Physiology* 172, 2416–2428.

Xu, M., Leichty, A.R., Hu, T., and Poethig, R.S. (2018). H2A.Z promotes the transcription of *MIR156A* and *MIR156C* in *Arabidopsis* by facilitatingthe deposition of H3K4me3. *Development* 145:dev152868.

Yang, L., Conway, S.R., and Poethig, R.S. (2011) Vegetative phase change is mediated by a leaf-derived signal that represses the transcription of miR156. *Development* 138, 245-249.

Yang, L., Xu, M., Koo, Y., He, J., Poethig, R.S. (2013) Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of *MIR156A* and *MIR156C*. *eLIFE* 2, e00260.

Yongchao L, Schroeder J, and Schmidt B (2013) Musket: a multistage k-mer spectrum based error corrector for Illumina sequence data. *Bioinformatics* 29:308-315.

Yu, S., Cao, L., Zhou, T., Lian, H., Sun, Y., Wu, J., Huang, J., Wang, G., Wang, J. (2003) Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. *eLIFE* 2, e00269.

Zimin, A.V., Marcais, G., Puiu, D., Roberts, M., Salzber, S.L., and Yorke, J.A. (2013). the MaSuRCA genome assembler. *Bioinformatics* 29, 2669-2677.