UNDERSTANDING THE GENETIC MECHANISMS OF DEVELOPMENT RATE IN PETUNIA AND STEVIA

By

Prabhjot Kaur

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics, and Biotechnology – Horticulture – Doctor of Philosophy

2024

ABSTRACT

Development rate, the rate at which plants produce new nodes/leaves, is crucial for determining crop readiness for harvest. Enhancing our genetic understanding and ability to manipulate this rate can lead to faster crop cycles, increased yield, and more efficient production, particularly for plants with short lifecycles or those harvested in the vegetative stage. Despite existing knowledge from other crops like *Arabidopsis* and rice, significant gaps remain in our understanding of development rate control. This dissertation investigates the genetic mechanisms controlling development rate in *Petunia × hybrida* and *Stevia rebaudiana* to improve crop timing and yield. Petunia, a popular annual bedding plant, often relies on heated greenhouse production in cooler climates. Identifying genetic factors regulating development rate in petunia is essential for reducing production costs and accelerating development rate at sub-optimal temperatures. The first objective was to evaluate the effect of candidate genes identified in previous studies by using virus-induced gene silencing (VIGS). Despite variable silencing efficiency and phenotypic data variability, the *MEI2-like1 RNA binding protein* emerged as a promising candidate, warranting further investigation with stable transformation methods. The second study employed an F_7 recombinant inbred line (RIL) population developed from *P. axillaris* and *P. exserta* (AE population) to identify genes differentially expressed between fast- and slow-developing AE RILs. DEGs included genes related to auxin polar transport, gibberellin signaling, MATE efflux transporters, and the 2OG-Fe(II) dependent oxygenase superfamily, among others. Common DEGs between AE and a previous RIL study involved cell-wall mechanics related genes such as *PECTINACETYLESTERASE FAMILY PROTEIN* and *L-ASCORBATE OXIDASE*, providing crucial insights into genetic factors influencing development rate. Stevia, known for its zero-calorie sweetening compounds produced by leaf, benefits from breeding faster-developing

varieties to enable multiple harvests per season. Stevia leaf yield depends on morphological traits such as leaf size, leaf production rate, plant canopy width and branch production. An F1 mapping population of 200 individuals was evaluated over two years at two field locations in Michigan for leaf-yield related traits. We generated a novel high-density SNP-based linkage map with eleven linkage groups encompassing eleven chromosomes. QTLs were identified at all four environments for traits such as maximum width, secondary branching, leaf length, and plant vigor, explaining 7-15% of phenotypic variation. Overlapping QTL regions were identified for traits including secondary branching, minimum canopy width, and leaf width explaining 7-15% of phenotypic variation. Differential expression analysis of F1 lines with contrasting development rates highlighted genes related to auxin efflux carrier proteins (*PIN-like 2*), auxin biosynthesis (*YUC2*), and cell wall loosening enzymes (*EXPANSIN*). Notably, *CYP78A10*, an ortholog of the *Arabidopsis KLUH* gene known to slow development rate, was upregulated in the slow development rate lines. Results from both species suggest that development rate is a complex process regulated by multiple factors, starting in the shoot apical meristem (SAM) through auxin polar transport, cell wall mechanics, and communication signals from emerging leaf primordia to the SAM. This research lays the foundation for breeding programs aimed at accelerating crop timing and increasing yield, enhancing overall crop production efficiency.

This dissertation is dedicated to my Parents and Grandfather.
Thank you for always believing in me!!

ACKNOWLEDGEMENTS

I would like to extend my heartfelt gratitude to my supervisor, Dr. Ryan Warner, for his exceptional mentoring and unwavering support throughout my PhD journey. His guidance has been invaluable, and I am deeply appreciative of his patience and wisdom!!

I also want to express my sincere thanks to my lab members, Sue Hammar, Nate Durussel, Keivan Bahmani, and all the undergraduates I have worked with. Your support and contributions have significantly enriched my research experience! This journey would not have been possible without the support of my committee members, Dr. Corny Barry, Dr. Courtney Hollender, and Dr. Dechun Wang. Your insights and encouragement have been crucial in shaping my work! Additionally, I extend my gratitude to Dr. Joseph Hill, Joseph Coombs, and Mohit Mahey for their assistance and contributions!

To the horticulture staff, your dedication and expertise have been instrumental in the success of my journey. Thank you for your unwavering support and hard work!

To my office mates and friends, Andrea and Charity, your kindness, help, and camaraderie have made this journey so much more enjoyable! I am grateful for your presence and support!!

I am deeply thankful to my parents for giving me this life. My mom's unwavering support and sacrifices to provide us with quality education and a good life are beyond words. Thank you, Mom, for letting me live my dreams! I don't say it enough, but I appreciate you and love so much!!

Papa, I miss you every day and wish you were here to see this accomplishment! I wouldn't be here today if it weren't for you!!

To my Grandfather, thank you for raising me and my siblings with love and invaluable values!! I know you continue to be my biggest cheerleader in spirit!!

To my wonderful husband, Parminder, you are my rock and my blessing! You make my life more beautiful and meaningful!! Thank you for loving me unconditionally, especially on the days I struggled to love myself. You celebrated my achievements and comforted me during my setbacks, always reminding me of my worth and potential. I am so grateful for your unwavering support and endless encouragement!! I promise to make more time for us now that this chapter is complete. You are my partner, my confidant, and my best friend, and I am forever thankful for your presence in my life!!

To my siblings, Kamal and Navu, your love means everything to me! Kamal, you've been not just a brother but also a role model and father figure, and I deeply appreciate you. Thank you for your immense contributions to our family! Your guidance, encouragement, and unwavering belief in me have played a pivotal role in my success. Navu, my dear sister, your constant support, understanding, and love have been a constant source of comfort and strength for me!

A heartfelt thanks also goes to my brother-in-law, Aman, for his unwavering support and kindness in numerous ways, and to my sister-in-law, Rupinder, for being an integral part of our family!

To my second family, Mom, Dad, and sister-in-law (Jyoti), your love and support have not gone unnoticed. Thank you for welcoming me into the family and always being there for me!!

To my adorable nephews, Arjan and Amarbir, and my beautiful niece Rehmat, your charm and innocence have brought so much joy to my life during this journey! Daler, our

nephew, although we lost you too soon and couldn't see you growing, you will always be remembered as God's favorite child!!

To my Uncle and Aunt, Nirmal and Ginni, you have been an essential part of this journey. When I decided to move away from home for grad school, your encouragement and support made this decision so much easier. Thank you for always being there for me and treating me like your own!

To my dear friends Manjot, Nancy, and Karan, you were my second home during grad school. Your friendship and support have been invaluable. Manjot, your care, and kindness over the past four years will always stay with me!! To my other grad school friends, Harkirat, Mohit, and Sukhdeep, it was a pleasure meeting you and sharing this journey!

Finally, to my childhood friends, Aman, Bhav, Taran, Ruby, and Rajwinder and my cousin Amolak, thank you for always believing in me and offering endless support. You keep me grounded, and your encouragement has been a powerful influence in my life!

Above all, I am deeply grateful to God for His endless blessings and guidance. Your grace and strength have carried me through this journey, and I am forever thankful!

This dissertation is a testament to the love, support, and encouragement I have received from all of you.

Thank you!!

TABLE OF CONTENTS

CHAPTER 1 LITERATURE REVIEW ……....1

CHAPTER 2 EVALUATING THE EFFECT OF TRANSIENTLY SILENCED CANDIDATE GENE EXPRESSION ON PETUNIA DEVELOPMENT RATE BY USING VIRUS INDUCED GENE SILENCING (VIGS)…… 13

CHAPTER 3 TRANSCRIPTOMIC ANALYSIS OF PETUNIA RECOMBINANT INBRED LINES (RILS) WITH CONTRASTING DEVELOPMENT RATES……43

CHAPTER 4 IDENTIFICATION OF QUANTITATIVE TRAIT LOCI (QTL) RELATED TO STEVIA DEVELOPMENT RATE AND OTHER LEAF YIELD RELATED TRAITS……84

CHAPTER 5 TRANSCRIPTOMIC ANALYSIS OF STEVIA F_1 LINES WITH CONTRASTING DEVELOPMENT RATES……..139

BIBLIOGRAPHY……..189

APPENDIX…….....206

CHAPTER 1

LITERATURE REVIEW

INTRODUCTION

Crop plants grow at different speeds due to their different development rates, defined as the rate of nodes/leaves produced over time (Warner and Walworth, 2010). Development rate is an important biological phenomenon determining when crops are ready to harvest i.e. time to first crop yield. It indirectly influences seasonal biomass accumulation in forage and bioenergy crops by influencing how many leaves are available for photosynthetic carbon fixation. Since development rate varies between plant species and even within groups of plants, there is a potential to change it to improve how efficiently crops grow and produce. This could involve encouraging earlier flowering or fruiting, leading to quicker harvests and higher yields. Therefore, it is desirable to study the genetic control of development rate.

The long-term goal of studying the genetics of development rate is to use this knowledge for speeding up crop timing even in less-than-ideal conditions and boosting overall crop production. This could mean more frequent harvests and increased yield, especially in plants with short lifecycles or those harvested in vegetative stage. My dissertation will focus on studying the genetic basis of development rate in two specific crops, *Petunia* × *hybrida* and *Stevia rebaudiana*. The insights gained from this research could pave the way for breeding new varieties of crops that develop faster and produce more efficiently.

BACKGROUND

During post-embryonic development, the shoot apical meristem (SAM) orchestrates the formation of nodes in a coordinated manner, influencing plant architecture through the regulation of two fundamental aspects: temporal (plastochron) and spatial (phyllotaxy) patterns (Stuurman et al., 2002) of node production. Plastochron, defined as the inverse of developmental rate, represents the time interval between the production of successive nodes (Wang et al., 2008;

Vallejo et al., 2015). Phyllotaxy, on the other hand, denotes the spatial arrangement of leaves along the stem (Giulini et al., 2004). Historically, two models have been proposed to elucidate leaf initiation. The first model posits the presence of a diffusible substance in the SAM and existing leaf primordia, inhibiting the formation of subsequent leaf primordia (Snow and Snow, 1932). Conversely, the second model suggests that physical forces within the SAM dictate phyllotactic arrangements (Selker et al., 1992; Green et al., 1996). However, the genetic underpinnings of these models remain elusive.

Genetic and molecular investigations have provided insights into the mechanisms governing leaf initiation. Notably, polar auxin transport emerges as a critical determinant in specifying the site of leaf initiation at regular intervals (plastochron) (Reinhardt et al., 2000). Following the emergence of newly formed leaf primordia, they function as sites of auxin sink activity. Consequently, auxin is depleted from the immediate vicinity of these primordia, leading to its accumulation at a distinct distance where subsequent leaf primordia are initiated (Reinhardt et al., 2003). Furthermore, mutations in genes such as *TERMINAL EAR 1* (*TE1*) in maize (Veit et al., 1998) and *ALTERED MERISTEM PROGRAM 1* (*AMP1*) in *Arabidopsis* (Helliwell et al., 2001), encoding an MEI2-like RNA binding protein and glutamate carboxypeptidase, respectively, have been associated with alterations in both plastochron and phyllotaxy, accompanied by changes in cytokinin levels. Given the pleiotropic effects of these genes, a focused examination of the genetic basis of plastochron, distinct from spatial considerations, is critical. This approach holds promise for elucidating the molecular intricacies underlying temporal regulation of leaf initiation in plant development.

Previous research has delved into plastochron-associated genes in several species. For instance, in rice, the gene *PLASTOCHRON1* (*PLA1*), along with its *Arabidopsis* ortholog *KLUH*, which encodes the cytochrome P450 family protein CYP78A11, has been identified as a negative

regulator of development rate (Miyoshi et al., 2004). Similarly, loss-of-function mutants of the *PLA2* gene, encoding an MEI2-like RNA binding protein and serving as the ortholog of maize TE1, have been associated with an accelerated development rate (Kawakatsu et al., 2006; Mimura et al., 2012). Additionally, *PLA3*, a rice ortholog of *Arabidopsis AMP1*, and maize *VIVIPAROUS8*, encoding a homolog of glutamate carboxypeptidase, has been implicated as a negative regulator of development rate (Kawakatsu et al., 2009). Loss-of-function mutations in all three rice *PLA* genes have led to an accelerated development rate, premature leaf maturation resulting in reduced leaf size, and the conversion of reproductive branches into vegetative shoots (Miyoshi et al., 2004; Kawakatsu et al., 2006; Kawakatsu et al., 2009; Mimura et al., 2012). Despite similar phenotypes observed in individual loss-of-function mutants of these genes, the phenotypic effects were exacerbated in double mutants, indicating that these plastochron genes operate via independent pathways (Kawakatsu et al., 2006). In barley, three *MANY-NODED DWARF* genes (*MND1*, *MND4*, and *MND8*), encoding a *N*-acetyltransferase-like protein, a CYP78A family protein (ortholog of rice *PLA1*), and a MATE transporter protein, respectively, have been identified as regulators of development rate (Hibara et al., 2021). This suggests that the control of development rate in barley involves transcriptional regulation of downstream genes by histone modulation (*MND1*), synthesis or metabolism of unknown substances (*MND4*), and transportation of unidentified cell molecules (*MND8*) (Hibara et al., 2021). Interestingly, *MND4* and *MND8* are orthologs of rice *PLA1* and maize *BIG EMBRYO 1* (*BIGE1*), respectively (Miyoshi et al., 2004; Suzuki et al., 2015). Double mutant analyses reveal independent regulation of development rate by *MND1* and *MND4* from *MND8* (Hibara et al., 2021). Furthermore, these genes exhibit limited expression in the shoot apical meristem and leaf primordia, unlike *MND8*, which lacks specific expression around the shoot apex.

Moreover, negative regulation of a subset of genes within the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) plant-specific transcription factor family by miR156 leads to an augmented development rate. For instance, overexpression of *miR156* downregulates the expression of *SPL9* and *SPL15* genes, consequently accelerating development rate, promoting branching, and altering inflorescence architecture (Schwarz et al., 2008). Conversely, the miR156-resistant form of *SPL9* decreases development rate in *Arabidopsis* (Wang et al., 2008). Loss-of-function mutations in various *SPL* genes have been associated with accelerated development rates in rice and maize (Jiao et al., 2010; Chuck et al., 2014; Wang et al., 2015). Members of this gene family are also implicated in promoting vegetative and floral phase transitions (Schwarz et al., 2008). For instance, overexpression of *SPL3* triggers early adult leaf trait emergence and flowering in *Arabidopsis* (Wu and Poethig, 2006). Alternatively, overexpression of *miR156* downregulates *SPL3* expression, resulting in an abundance of juvenile trait leaves and delayed flowering (Schwab et al., 2005; Wu and Poethig, 2006). Notably, mutations in *SERRATE*, an *Arabidopsis* zinc finger protein pivotal in miRNA biogenesis and primary miRNA processing, lead to a reduction in development rate (Prigge and Wagner, 2001; Grigg et al., 2005; Lobbes et al., 2006).

Plant growth hormones, including auxins and cytokinins, are integral to the regulation of development rate in *Arabidopsis* and tobacco (Reinhardt et al., 2000; Werner et al., 2001). For instance, the overexpression of four cytokinin oxidase genes (*AtCKX*) in tobacco leads to a reduction in cytokinin concentration and a subsequent decrease in development rate (Werner et al., 2001). Similarly, the *Arabidopsis slow motion (slomo)* mutant, characterized by decreased free auxin levels, exhibits a reduced development rate (Lohmann et al., 2010). Additionally, in rice, genes *PLA1* and *PLA2* are known to operate downstream of the gibberellin signal transduction pathway (Mimura et al., 2012). Notably, *pla1* and *pla2* mutants display lower

concentrations of other phytohormones, such as cytokinin, abscisic acid, and auxin, compared to their wild-type counterparts (Kawakatsu et al., 2009).

Despite our extensive knowledge of genes and mutants, there are still significant gaps in understanding how these components interconnect within a network and how their biochemical activities are coordinated. Investigating the genetic mechanisms underlying the development rate in various plant species is crucial to determine whether the regulation of this trait involves conserved pathways or novel mechanisms.

Economics/Uses of Petunia

Petunia × hybrida (Petunia), commonly known as garden petunia, originated from a cross between two wild species, *P. axillaris* and *P. integrifolia*, in the early 19th century (Gerats and Strommer, 2008). It has become one of the leading annual bedding crops in the United States, boasting a wholesale sales value of US$160 million recorded in 2020 (USDA, 2021). Renowned for its wide array of colorful flowers and diverse morphology, this plant belongs to the *Solanaceae* family, characterized by a base chromosome number of x=7, unlike the typical x=12 found in most other members of this family, such as tomato, potato, pepper, tobacco, and eggplant (Guo et al., 2017). The genus *Petunia* comprises 20 species native to South America (Stehmann et al., 2009; Guo et al., 2017) most of which can be readily hybridized with varying degrees of fertility (Ando et al., 2001; Anderson, 2006; Warner, 2010). Petunia is renowned for its ease of cultivation and short lifecycle, typically spanning approximately four months from seed to seed (Vandenbussche et al., 2016). It is classified as a facultative long-day plant, exhibiting accelerated flowering under extended daylight periods and relatively higher temperatures.

The development rate of petunia is intricately linked to temperature (Vandenbussche et al., 2016). Typically, plants exhibit faster development rate as temperatures rise up to a certain optimal range (18-24°C), and conversely, the rate slows as temperatures decrease towards the base temperature (Adams et al., 1998). In regions with northern latitudes across North America and Europe, greenhouse cultivation is common during colder periods of the year to ensure timely market readiness during spring (Guo et al., 2017). Consequently, greenhouse operators incur substantial energy costs to maintain optimal temperatures for petunia flowering. These expenses significantly impact profit margins within the greenhouse industry. Moreover, it is important to note that higher temperatures can compromise crop quality (Warner, 2009). Conversely, crops cultivated at lower temperatures tend to exhibit better quality due to an extended duration of exposure to harvest light (Personal communication by Erik Runkle). There exists an opportunity to mitigate these energy expenditures by developing petunia varieties capable of accelerated development rates at cooler temperatures.

Consequently, current research endeavors focus on elucidating the genetic mechanisms governing the development rate of petunias, with the aim of reducing production time. If plants could produce the same number of nodes at an increased rate and lower temperatures, it would lead to a reduction in crop timing and production costs. Moreover, increasing the development rate of petunia plants can result in higher yields of cuttings for clonal propagation, which is widely practiced in petunia horticultural production (Santos et al., 2011; Toma et al., 2011; Vandenbussche et al., 2016). Ultimately, this would benefit both growers and customers.

Genetic of development rate in Petunia

Diverse development rates observed between wild petunia species and commercial cultivars at equivalent temperatures suggest the potential for breeding varieties with accelerated development rates (Warner and Walworth, 2010). Consequently, studies employing quantitative

trait loci (QTL) mapping were initiated to identify candidate genomic regions associated with petunia development rate. Initially, interspecific F_2 populations, namely *P. integrifolia* × *P. axillaris* (the "IA" population) and between *P. axillaris* and the more recently diverged species *P. exserta* (the AE population), were utilized to assess variation in development rates at the genetic and genomic levels. These populations were chosen due to their demonstration of transgressive segregation for development rates (Warner and Walworth, 2010).

QTL analysis revealed the presence of development rate-associated loci on chromosomes 1, 2, and 5, collectively explaining 34% of the observed variation (Vallejo et al., 2015). Moreover, reference transcriptomes for *P. axillaris*, *P. exserta*, and *P. integrifolia*, comprising mRNA libraries from various tissues including shoot apex, whole 3-week-old seedlings, mixed floral development stages, trichome, and callus tissues, have been established (Guo et al., 2015). These transcriptomes were mined to identify petunia transcripts homologous to genes associated with development rate, which were subsequently converted into molecular markers and mapped to the IA F_2 population (Guo et al., 2015; Vallejo et al., 2015). However, except for one gene encoding an *MEI2*-like RNA binding protein homologous to the rice *PLA2* gene (Kawakatsu et al., 2006), none of these development rate-associated gene homologs co-localized with the identified QTL for development rate in this population. Furthermore, the low marker density in the IA F_2 population limits the efficacy of the development rate QTL in identifying candidate genes underlying this trait (Guo et al., 2017).

Subsequently, F_7 recombinant inbred lines (RILs) were established for the same species (the IA population and the AE population) (Guo et al., 2017). These RILs were phenotyped for development rate at three distinct temperatures (14, 17, and 20°C) to pinpoint QTL regions associated with the trait (Guo et al., 2017). Additionally, IA RILs exhibiting varying development rates were employed to identify 209 differentially expressed genes (DEGs) (Guo et

al., 2017). Subsequently, QTL and DEGs were mapped onto high-density single nucleotide polymorphism (SNP) bin-based linkage maps generated for both populations (Guo et al., 2017). Out of all DEGs, thirteen were found to map within 1 centimorgan (cM) of a development rate QTL, with notably large clusters of differentially expressed genes located proximate to IA development rate QTL on chromosomes 5 and 6 (Guo et al., 2017). Within these differentially expressed genes were transcripts associated with phytohormones (specifically auxin and cytokinin) synthesis or signaling pathways, as well as miRNA-mediated pathways, which have previously been implicated in the control of development rate as detailed in the background section above. This data represents a significant step forward in facilitating the identification and characterization of the genetic factors governing development rate.

Economics/Uses of Stevia

Stevia rebaudiana, commonly referred to as stevia, is a significant medicinal perennial plant within the *Asteraceae* family, with a chromosome count of 2n=22 (Goyal et al., 2010). Originating from northeast Paraguay (Shock, 1982; Ramesh et al., 2006), stevia leaves are renowned for producing a collection of zero-glycemic, low-calorie sweet-tasting compounds known as steviol glycosides (Brandle and Telmer, 2007; Ceunen and Geuns, 2013). These steviol glycosides (SGs), extracted from the leaves, can constitute up to 30% (on a dry mass basis) of these compounds, with their sweetness being 200-300 times greater than sucrose (Goyal et al., 2010; Yadav and Guleria, 2012; Ceunen and Geuns, 2013). Stevia has been utilized as a sweetener in various products in Japan since the 1970s, including seafood, soft drinks, and candies (Mizutani and Tanaka, 2001). Moreover, stevia has been explored as a weight control agent for obese individuals and as a natural diabetes control remedy in different regions worldwide (Gupta et al., 2013; Shivanna et al., 2013; Ahmad et al., 2020). Given their plant-based origin, steviol glycosides hold significant promise as alternatives to sugar and synthetic

sweeteners, appealing to those seeking natural ingredients in their diet. Consequently, there is a growing demand for high steviol glycoside-yielding cultivars, necessitating plant breeders to focus on their development.

Stevia is characterized by high heterozygosity and obligate outcrossing due to self-incompatibility (Yadav et al., 2014; Attaya, 2017). Consequently, it is commonly propagated through stem cuttings and in-vitro methods (Ramesh et al., 2006; Sairkar et al., 2009). Harvesting above-ground tissue of stevia involves stripping off leaves for the extraction of steviol glycosides. To increase steviol glycoside yield, accelerating leaf production rate over time is crucial to facilitate multiple harvests per season. Therefore, understanding the genetic basis of stevia leaf production rate is essential for breeding high-yielding cultivars. However, genetic research on stevia is hindered by factors such as limited germplasm availability, molecular markers, and a high-resolution linkage map (Basharat et al., 2021; Huber and Wehner, 2023). As stevia is a relatively new crop in the genomics era, collaborative efforts are underway to develop diverse germplasm and generate genetic and genomic resources, facilitating breeding efforts for increased stevia biomass production (Kaur et al., 2015; Bahmani, 2021; Vallejo and Warner, 2021; Xu et al., 2021; Huber and Wehner, 2023).

Potential implications of understanding the genetics of development rate

The rate at which plants generate new nodes is a fundamental determinant of crop production timing or time to first yield in agricultural crops. Given that several significant fresh market vegetable crops, such as tomato, pepper, and eggplant, belong to the same family as petunia (*Solanaceae*), the findings from this research project could be promptly applied to enhance these crops. Similarly, the insights gained from this study could be extended to other crops with restricted growing seasons, particularly in regions of the United States characterized by long winters. For instance, the outcomes of this study could benefit numerous vegetable crops

cultivated within limited seasons for fresh consumption. By enhancing the developmental pace of these crops, production efficiency could be enhanced by enabling multiple harvests within a single season. This aligns with the increasing consumer demand for locally grown produce, a trend that has surged in recent years. Moreover, comprehending the genetics of vegetative development rate may shed light on the potential coupling of genes involved in shoot growth (vegetative development) with those governing reproductive phases (flowering/fruiting), thereby creating opportunities to adjust crop timing and enhance production efficiency through early flowering/fruiting.

DISSERTATION OBJECTIVES

Considering the significance of investigating the development rate trait, my dissertation aimed to explore the genetic underpinnings of development rate in both petunia and stevia. By studying these two distinct species, we sought to provide valuable insights into the regulation of this trait in both species and laying the foundation for future breeding efforts aimed at improving crop productivity and efficiency.

The first objective in petunia was to functionally characterize potential candidate genes linked to development rate through the utilization of reverse genetics techniques. Secondly, to comprehensively examine the common and unique pathways involved in development rate, we conducted transcriptomic analyses of petunia RILs exhibiting varying development rates within a previously unexplored population.

In the case of stevia, the dissertation objectives were formulated with the aim of expanding genetic resources and investigating the genetic basis of development rate in this species. Thus, the final two objectives involved open field trials of a stevia genetic mapping population to identify genomic regions associated with morphological traits related to stevia leaf

biomass, including development rate, and employing transcriptomic analyses to identify differentially expressed genes between genotypes with contrasting development rates.

CHAPTER 2

EVALUATING THE EFFECT OF TRANSIENTLY SILENCED CANDIDATE GENE EXPRESSION ON PETUNIA DEVELOPMENT RATE BY USING VIRUS INDUCED GENE SILENCING (VIGS)

INTRODUCTION

Understanding the genetic architecture underlying development rate in crops like petunia is crucial for optimizing crop timing and improving production efficiency, especially in regions with sub-optimal temperature conditions. Development rate, characterized by the rate of vegetative node formation before floral initiation, directly impacts the timing of crop maturity and harvest (Warner and Walworth, 2010). *Petunia* × *hybrida* is a highly valuable bedding crop, with significant economic importance in the horticultural industry (USDA, 2021). In regions with cooler climates, such as the northeastern states of the U.S., petunia production often relies on heated greenhouses to maintain optimal temperatures for growth and development. However, this heating requirement increases production costs and reduces profit margins for growers (Guo et al., 2015). Therefore, there is a pressing need to identify genetic factors that regulate development rate in petunia, particularly those that could potentially accelerate development under sub-optimal temperature conditions. By studying the genetic architecture of development rate, researchers can identify key genes and pathways involved in regulating this trait. The knowledge gained from this can then be used to develop breeding strategies aimed at selecting faster-developing varieties of petunia that are better adapted to cooler temperatures. Accelerating development rate in petunia could reduce the reliance on heating in greenhouses, thereby lowering production costs and increasing profitability for growers.

The genetic control of plastochron, which represents the time interval between two successive nodes and is the inverse of development rate, provides valuable insights into the regulation of development rate in plants (Wang et al., 2008; Guo et al., 2015). Previous research has identified several key genes and pathways involved in modulating plastochron. Loss-of-function mutants of *PLASTOCHRON1* (*PLA1*) in rice and *KLUH* in *Arabidopsis*, both encoding

cytochrome P450 family proteins, have been associated with an increased development rate (Miyoshi et al., 2004, 2004; Wang et al., 2008). Similarly, loss-of-function mutants of *PLASTOCHRON2* (*PLA2*) also result in accelerated development rates (Kawakatsu et al., 2006, 2006; Mimura and Itoh, 2014). Additionally, *PLA3*, a rice ortholog of *Arabidopsis ALTERED MERISTEM PROGRAM 1* (*AMP1*) and maize *VIVIPAROUS8*, has been identified as a positive regulator of plastochron. Several genes, including *MANY-NODED DWARF* (*MND*) genes in barley (Hibara et al., 2021) and members of the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) transcription factor family, have also been implicated in regulating plastochron. Notably, *miR156*, a microRNA, negatively regulates genes belonging to the SPL transcription factor family. Overexpression of *miR156* results in a shorter plastochron and accelerated development rate in various plant species (Xie et al., 2006; Xie et al., 2012), while silencing of specific *SPL* paralogs can either reduce or increase the development rate in petunia (Preston et al., 2016). Furthermore, plant growth hormones such as auxins, cytokinins, and gibberellins also play crucial roles in modulating development rate (Reinhardt et al., 2000; Werner et al., 2001). Despite these significant advances, there are still gaps in our understanding of the control of development rate.

The identification of candidate genes for development rate control in petunia involved a multi-step approach combining genetic mapping, RNA sequencing, and literature analysis. Quantitative trait loci (QTL) associated with development rate were identified in two F_7 recombinant inbred lines (RILs) derived from crosses between different wild progenitor species of petunia, *P. integrifolia* × *P. axillaris* (IA population) and *P. axillaris* × *P. exserta* (AE population), across multiple temperatures (Guo et al., 2017). Additionally, differentially expressed genes (DEGs) between IA RILs exhibiting slow and fast development rates were identified through RNA sequencing analysis in the same study (Guo et al., 2017). Candidate

genes potentially involved in development rate control were selected using five criterion, which included parameters such as significance of QTL association, differential expression levels, and functional relevance based on previous literature studies (Table 2-1) (Reinhardt et al., 2000; Prigge and Wagner, 2001; Miyoshi et al., 2004; Kawakatsu et al., 2006; Mimura and Itoh, 2014; Preston et al., 2016). Genes meeting at least two of these criteria were prioritized as candidate genes for further investigation. In total, 24 candidate genes were identified through this selection process, representing promising targets for functional validation studies (Table 2-2).

Reverse genetics techniques, like virus-induced gene silencing (VIGS), offer a powerful means to rapidly evaluate the function of genes by suppressing their expression and observing resultant phenotypic changes. This approach bypasses the need for time-consuming plant regeneration steps, enabling quicker functional genomics studies (Benedito et al., 2004; Unver and Budak, 2009). VIGS relies on the RNA-silencing mechanism, where specific gene sequences are integrated into a viral vector, such as Tobacco rattle virus vector (TRV2), which is then introduced into the plant genome using *Agrobacterium*-mediated delivery (Reid et al., 2009; Zulfiqar et al., 2023). Once inside the plant cells, the viral vector triggers the degradation of mRNA molecules corresponding to the targeted genes, thereby silencing their expression. This process mimics the plant's natural defense mechanism against viruses. Researchers have successfully utilized VIGS to characterize phenotypes by silencing candidate genes associated with specific traits of interest. The versatility and effectiveness of VIGS have been demonstrated in various plant species, including petunia, *Arabidopsis*, tomato, tobacco, potato, barley, and more (Hein et al., 2005; Reid et al., 2009; Velásquez et al., 2009; Noor et al., 2014; Tomar et al., 2021; Singh et al., 2022).

The objective of the current study was to assess the role of candidate genes in controlling development rate in petunia through the application of VIGS. By reducing or silencing the

expression of these candidate genes, the aim was to investigate whether alterations in gene expression levels would lead to observable changes in the development rate phenotype of petunia plants.

MATERIALS AND METHODS

Plant material and growth conditions

P. axillaris (PI 667515) seeds were sown in a 72-cell tray with a cell volume of 16.4 cm^3, under short day conditions (9-h light/ 22 °C) in a growth chamber. Seedlings with 2-3 true leaves were moved to 50-cell trays with 75.4 cm^3 per cell, a few days before Agroinfiltration.

Plasmid construction and gene cloning

The VIGS vectors used in this study were derived from Tobacco rattle virus and consisted of pTRV2-LIC (Dong et al., 2007) and pTRV1 (Liu et al., 2002). Plasmid miniprep was performed with EZ-10 Spin Column Plasmid DNA Miniprep Kit (Bio-Basic, Amherst, New York) to extract pTRV2-LIC vector DNA. This plasmid vector was linearized by digesting it with the *PstI*-HF®, restriction endonuclease (New England BioLabs (NEB), Ipswich, Massachusetts) and purified. Gene-of-interest target sequences of around 300 bp were designed (Table 2-3) by using the VIGS tool in the Sol Genomics Network (SGN) (Fernandez-Pozo et al., 2015). Forward and reverse primers were designed for each gene construct ($T_m \geq 60$ °C) using the primer design tool in the Benchling program (https://www.benchling.com) (Table 2-3). Two 15-bp adapter sequences, 5'-CGACGACAAGACCCT -3'and 5'- GAGGAGAAGAGCCCT- 3', described previously (Dong et al., 2007), were included in the forward and reverse primer sequences of each gene, respectively. RNA was extracted from the leaves with the RNeasy Plant Mini kit (Qiagen, Germantown, Maryland). The gene segments were amplified with the iTaq™ Universal SYBR® Green One-Step Kit (Bio-Rad, Hercules, California) with cDNA as a

template. PCR products were purified with the EZ-10 Spin Column PCR Products Purification Kit (Bio-Basic, Amherst, New York) and run on a 1% agarose gel to visualize the ca. 300 bp bands.

Ligation independent cloning (LIC) was performed to insert the gene of interest segments into the pTRV2-LIC vector (Dong et al., 2007). Briefly, both the vector and PCR products were treated with the T4 DNA polymerase (NEB) at 22 °C for 30 min and 70 °C for 20 min on a thermocycler. Following, the TRV2-LIC vector and PCR products were mixed in a 1:1 ratio and incubated at 65 °C for 2 min and 22 °C for 10 min to facilitate the covalent bonding. Then 6 µL of the final LIC product were mixed with DH5α (*E. coli*) competent cells (NEB), incubated on ice for 30 min followed by a heat shock treatment at 42 °C for 55 seconds and then back on ice for 2 min. The cells were then mixed with 600 µL of the SOB (Super Optimal broth) medium and shaken at 200 r.p.m and 37 °C for an hour. Finally, 80 µl of the cells was spread on LB agar + Kanamycin plates and incubated overnight at 37 °C. Transformants were tested by colony PCR using Phusion® High-Fidelity DNA polymerase (NEB). Primer sequences used for colony PCR were forward or reverse primer sequences specific to each gene and a TRV2- forward or reverse primer. Positive colonies were cultured overnight in a liquid LB + Kan medium on a shaker at 200 r.p.m and 37 °C. Plasmid DNA was extracted, purified, and sent for Sanger sequencing at MSU's genomics core facility. A total 6 µL of the mixture containing 4 µL of the plasmid DNA and 2 µL of the TRV2- forward (5'-TGTTACTCAAGGAAGCACGATGAGCT -3') or reverse primer (5'-AACTTCAGGCACGGATCTACTTA -3') was used for sequencing. MEGAX was used for sequence alignment and verification of the constructs. Positive sequences (TRV2-LIC + gene of interest) were transformed into *Agrobacterium tumefaciens* GV3101 competent cells as previously described (Gelvin, 2012).

VIGS inoculation

The VIGS protocol was followed as previously described (Velásquez et al., 2009). Day 1 of the protocol includes growing *Agro*-transformed pTRV1 vector, pTRV2-gene product, pTRV2-empty vector, and pTRV2-*PDS* on LB agar plates supplemented with 50 μg/mL kanamycin and 100 μg/mL rifampicin antibiotics for two days at room temperature. On day 3, the colonies are cultured on a liquid LB with the same antibiotics by shaking at 200 r.p.m. and 30 °C for 16-18 hours. On day 4, the primary culture is diluted at 1:25 into the secondary induction medium (IM) (Velásquez et al., 2009) with the above-mentioned antibiotics plus 200 μM acetosyringone and shaken at 200 r.p.m at 30 °C for 20-24 hours. Finally, day 5 steps (Velasquez et al., 2009) are followed and pTRV1 and pTRV2 vectors are mixed 1:1. The inoculum finally becomes ready for the infiltration.

Plants with 2-4 true leaves were used for inoculations (Figure 2-1). Two fully expanded leaves on each plant were scratched with a blade to poke a hole. Agroinfiltration was carried out by injecting the inoculum into the leaves to the point of cell saturation by using a 1 mL needless syringe. At least 15 plants were used per each construct. Plants for each construct were separated by a row and gloves were changed in between constructs to avoid cross-contamination. Plants were also not watered for at least 24 hours after the inoculations. pTRV2-E vector and wild type were used as negative controls and pTRV2-*PDS* was used as a positive control in each experiment.

Silencing evaluation by qPCR analysis

The onset of photobleaching symptoms on plants inoculated with pTRV2-*PDS* were used to schedule the tissue collection from all plants (Figure 2-1). The youngest fully expanded leaf tissues (~100 mg) were collected from each plant on a 96 well-plate and flash frozen with liquid

nitrogen and either used immediately or stored at -80 °C for future use. RNA was extracted using the MagMAX™ Plant RNA Isolation Kit (Thermofisher Scientific, Waltham, Massachusetts), quantified using Nanodrop, and run on a 1% agarose gel. Gene-specific primers that amplify the region outside of the VIGS targeted region were designed with Net Primer https://www.premierbiosoft.com/netprimer/) and Benchling program tools (Table 2-4). Standard curves were made by serial dilutions of the TRV2-E RNA at five different concentrations (150, 100, 50, 25 and 12.5 ng). Primer efficiency was calculated as Efficiency (%) = ($10^{(-1/\text{slope of standard curve})}$-1) X 100 and primers with 80-110% efficiency were utilized. 10 µL qPCR reactions were performed by using iTaq™ Universal SYBR® Green One-Step Kit (Bio-Rad) on a 384-well plate. The MIQE guidelines were followed for these qPCR experiments, ensuring standardization and reproducibility (Taylor et al., 2010). The qPCR conditions were 50 °C for 15 min, 95 °C for 1 min, 95 °C for 20 sec, 55 °C for 20 sec, 72 °C for 1 min, repeat cycles 2-5 34 times and 72 °C for 3 min. *EF1α* was used as a housekeeping gene (Mallona et al., 2010). Gene sequence of *P. axillaris* EF1α (Peaxi162Scf00389g00936.1) was extracted from the Jbrowse tool of SGN and primers were designed using Net Primer (https://www.premierbiosoft.com/netprimer/). The $2^{-\Delta\Delta CT}$ method (Livak et al., 2013)was used to calculate the relative expression of genes.

Phenotyping

Plants were moved to 15.24 cm diameter pots with a cell volume of 1420.76 cm^3 in the greenhouse under short day conditions (9-hr light/ 22°C) after tissue collection. Topmost fully expanded leaves on the main stem and two side branches per plant were marked with a white paint (Figure 2-1). The number of new nodes developed beyond that point were counted for each plant. We collected data at two time points, 4- and 6-week time intervals depending on the growth of the plant. We did not count the nodes on plants that were already flowering.

Data analysis

Boxplots were generated to compare the node numbers of plants with relative gene expression thresholds of ≤ 0.4 for each gene construct separately by using R. One-way analysis of variance (ANOVA) was performed to test for significant differences between the boxplots and Tukey's HSD test was performed for pairwise comparisons between the boxplots in R ($p \leq 0.05$).

Phylogenetic analysis of MEI2-like genes

Sequence data of Arabidopsis, maize, rice and yeast MEI2-like genes was retrieved from NCBI using the accession numbers previously described (Kaur et al., 2006). To identify MEI2 genes in petunia, the gene sequences from these species were blasted against the petunia genome (*P. axillaris* v1.6.2) in the Sol Genomics Network (Bombarely et al., 2016). The mRNA sequences of five MEI2 -like genes were found by BLAST search and used in this study: Peaxi162Scf00023g00929.1 ("*MEI2-like protein 1*"), Peaxi162Scf00128g01746.1 ("*MEI2-like protein 1*"), Peaxi162Scf00035g02717.1 ("*MEI2-like protein 5*"), Peaxi162Scf00214g00068.1 ("*MEI2-like protein 5*") and Peaxi162Scf00111g00117.1 ("*MEI2-like protein 5*"). Sequences were aligned using ClustaIW in MegaX (Kumar et al., 2018; Stecher et al., 2020).

The evolutionary history was inferred by using the Maximum Likelihood method and Hasegawa-Kishino-Yano model (Hasegawa et al., 1985). The tree with the highest log likelihood (-91168.45) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 4.5615)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 18

nucleotide sequences. There were a total of 9303 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018; Stecher et al., 2020).

RESULTS AND DISCUSSION

We aimed to use VIGS as a high-throughput method to screen multiple candidate genes for their potential role in regulating development rate in plants. VIGS is a technique used to downregulate the expression of specific plant genes by introducing a virus that carries a fragment of the target gene's sequence. This leads to gene silencing and allows researchers to study the effects of gene knockdown on plant phenotypes. We selected 24 candidate genes based on their potential involvement in regulating development rate (Table 2-1). The designed gene fragments were cloned into the pTRV2 vector. Out of the 24 candidate genes, only 17 were successfully cloned into the pTRV2 vector and, for the remaining 7 genes, efficient VIGS target regions could not be designed. This was either because the genes were too small to yield suitable target regions or because there were too many off-target regions, making it difficult to design gene-specific VIGS constructs. The successful cloning of 17 genes into the VIGS vector allowed for the subsequent silencing of these genes in plants to assess their phenotypic effects.

The first step in the phenotype evaluation process involved confirming whether the plants were effectively silenced using real-time quantitative PCR (qPCR). From qPCR analysis, genes Peaxi00008 and Peaxi00014 consistently showed high Cq values on the standard curve, indicating low expression levels. The large difference (~10 cycles) between the Cq values of these genes and the reference/housekeeping gene (*EF1α*) made it challenging to measure their expression levels accurately. Such genes with low mRNA abundance are reportedly known to be less susceptible to silencing by VIGS-like methods (Hu et al., 2004). Peaxi00012 presented multiple peaks on the melt curve with different qPCR primer sets, indicating the possibility of amplifying off-target products (Figure 2-2). However, the VIGS target region of Peaxi00012 was

designed with a target score of 100%. Plants inoculated with constructs targeting Peaxi00008.2, Peaxi00015, Peaxi00016, Peaxi00125, and Peaxi01115 genes exhibited photobleaching symptoms similar to those seen in *Phytoene desaturase* (*PDS*) inoculated plants (Figure 2-3). *PDS* is involved in carotenoid biosynthesis crucial for photosynthesis, and loss-of-function mutants display photobleached leaves due to chlorophyll disruption (Wang et al., 2009). *PDS* is often used as a positive control in VIGS experiments due to its easily observable phenotype (Fu et al., 2006). The reason for non-*PDS* inoculated plants displaying albino symptoms remains unknown particularly since precautions were taken to prevent cross-contamination, including changing gloves between constructs and spatially separating *PDS*-treated plants from others.

We utilized the *PDS* phenotype as a visual indicator to guide tissue collection for qPCR analysis. Typically, PDS mutants exhibit symptoms within 10-14 days post-inoculation. However, we observed occasional delays in phenotype expression, possibly due to environmental factors such as cooler temperatures and lower humidity levels, as reported in previous studies on tomato (Fu et al., 2006). Upon evaluating relative gene expression levels in inoculated plants, we noted a wide range of expression values, consistent with observations in *PDS*-inoculated plants where the albino phenotype varied (Figure 2-4), suggesting that plants could still produce chlorophyll to some extent. To explore the correlation between *PDS* phenotype and relative expression levels, individual *PDS* leaves were subjected to qPCR analysis. Interestingly, leaves with both low (0.2) and moderate (0.5) relative expression levels exhibited weaker phenotypes, while a leaf with a relative expression level of 0.4 showed no phenotype (Figure 2-5). These findings suggest that qPCR expression levels reflect a general down-regulation of the *PDS* gene but may not precisely quantify the number of translationally active transcripts (Urso et al., 2013). This observation aligns with previous studies in petunia (Broderick and Jones, 2014),

highlighting the significance of thoroughly understanding mRNA degradation processes for effective gene suppression.

The efficiency of VIGS in plants is heavily dependent on the systematic movement of the viral construct throughout the entire plant. Effective viral movement ensures the downregulation of endogenous genes, while ineffective movement hampers gene silencing. Factors such as the inoculation method, cultivar choice, and environmental conditions, particularly temperature, influence viral movement in plants. Previous studies (Caplan and Dinesh-Kumar, 2006; Muruganantham et al., 2009; Zulfiqar et al., 2023) have highlighted the critical role of these factors in regulating viral movement and subsequent gene silencing. Optimal conditions for VIGS have been demonstrated in petunia, with the Picobella Blue cultivar identified as the most suitable for efficient gene silencing (Broderick and Jones, 2014). We conducted preliminary studies with Picobella Blue, but its compact growth habit made it difficult to evaluate the development rate phenotype, particularly in counting the number of nodes due to the dense growth. As a result, we decided to continue using the *P. axillaris* genotype for our experiments. Furthermore, factors contributing to variability in silencing efficiency are attributed to the viral method itself. VIGS operates through an RNA interference-based mechanism, where double-stranded RNAs (dsRNAs) produced in response to the virus are cleaved into short interfering RNAs (siRNAs) by the enzyme DICER. These siRNAs then guide the RNA-induced silencing complex (RISC) to degrade complementary target mRNAs (Zulfiqar et al., 2023). Various properties related to both siRNA sequences and target mRNAs influence the efficacy of siRNA-mediated gene silencing (Holen et al., 2002; Czauderna et al., 2003; Hu et al., 2004). For instance, the presence of secondary structures in the target mRNA or its specific subcellular localization makes it inaccessible to the siRNA-mediated degradation by the RISC complex (Holen et al., 2002). Additionally, structural variation and chemical properties of siRNAs also

impact their effectiveness in silencing. Moreover, a sequence complementarity of less than 11 nucleotides between siRNA and mRNA reduces the likelihood of successful gene silencing. These insights underscore the complexity of VIGS as a gene silencing tool and emphasize the importance of understanding the various factors influencing its efficiency for successful gene knockdown experiments.

We focused on analyzing genes showing at least 60% silencing or relative expression values of ≤0.4, namely Peaxi00310, Peaxi00316, Peaxi00929, Peaxi00027, Peaxi00048, Peaxi00073, Peaxi00737, and Peaxi01330. This targeted approach allowed for a more focused and efficient analysis of gene function in relation to the highly quantitative trait under investigation. Node numbers of the plants were compared to the wild type and empty vector negative controls. The use of Peaxi162Scf00069g01624.1, a *SQUAMOSA PROMOTOR BINDING-LIKE PROTEIN 12* (*SPL12*) gene, as a negative control did not yield significant results due to inefficient silencing. This gene is a homolog of *Petunia* × *hybrida SPL* gene (*PhSBP2*) known to negatively regulate development rate in petunia (Preston et al., 2016). There were no significant differences in node numbers between plants with gene expression levels ranging from 0.10 to 0.44 (equivalent to 90% to 56% silencing) and the empty vector wild type controls, both at the main stem and side branches (Figure 2-6). Similarly, comparing plants with gene expression levels of 0.1-0.4 to those with expression values of 0.5-0.9 revealed no significant differences in node numbers (Figure 2-7). Additionally, when evaluating the phenotype of inoculated plants at a relatively early developmental stage (four week-interval), no significant differences were observed in node numbers between plants with at least 80% silencing and wild type controls, both at side branches and the main stem (Figure 2-8). However, despite the lack of significant differences in phenotypic outcomes among effectively silenced plants, it is noted that the limited number of plants hindered robust conclusions, particularly

regarding Peaxi00929, which tended to produce higher node numbers compared to controls when relative expression was 0.4 or less.

The gene Peaxi00929, functionally annotated as *an MEI2-like protein 1* in petunia, is of particular interest due to evidence suggesting its involvement in development rate regulation (Guo et al., 2015; Guo et al., 2017). A previous study uncovered evidence indicating the presence of a CAPS marker co-localizing with a homolog of *MEI2-like 1* within a development rate QTL region in an interspecific *P. integrifolia x P. axillaris* F_2 population (unpublished results from a study conducted by Guo et al., 2015). Despite this QTL region likely containing other genes, this finding serves as an additional piece of evidence supporting the significance of this gene in the regulation of development rate. MEI2 like genes belong to a class of RNA binding protein genes, initially identified in the fission yeast *Schizosaccharomyces pombe* (Hirayama et al., 1997). The gene family is categorized into three functional clades based on sequence similarity, with two of these clades, *Arabidopsis* meiotic -like (AML), and terminal ear-like (TEL), identified as functional in plants (Jeffares et al., 2004). Genes belonging to these clades have been characterized in various plant species, including *Arabidopsis*, rice, maize, and barley (Hirayama et al., 1997; Jeffares et al., 2004; Mercier and Grelon, 2008; Wang et al., 2022). TEL clade genes, such as *TERMINAL EAR1* (*TE1*) in maize and *TEL1* and *TEL2* in *Arabidopsis*, are known to be expressed in the shoot apical meristem and play roles in plant architecture. Loss-of-function mutants of *TE1* in maize and its rice ortholog *PLA2*, have been shown to result in accelerated leaf initiation rates and dwarf phenotypes (Kawakatsu et al., 2006; Mimura and Itoh, 2014; Wang et al., 2022). *AML* genes, including *AML1-5*, are broadly expressed in both vegetative and reproductive tissues and are involved in meiosis and vegetative development (Anderson et al., 2004; Kaur et al., 2006). Loss-of-function mutants of *AML1* and *AML4* genes in *Arabidopsis* have been reported to lead to retarded seedling growth and seedling arrest (Kaur et al., 2006).

This evidence suggests that Peaxi00929, as a *MEI2-like 1* gene, may play a significant role in regulating development rate in petunia, possibly through mechanisms involving RNA binding and meiotic functions, similar to its counterparts in other plant species.

Our phylogenetic analysis of *MEI2* genes from various species, including petunia, *Arabidopsis*, rice, maize, and yeast, provides valuable insights into the evolutionary relationships and classification of these genes into distinct clades (Figure 2-9). The phylogenetic tree reveals four main clades: two *AML* (AML 14 and AML 235) clades, a TEL clade, and a non-plant clade, consistent with previous studies (Anderson et al., 2004; Jeffares et al., 2004; Kaur et al., 2006). Interestingly, two petunia *MEI2-like 1* genes including Peaxi00929.1 belong to the AML14 clade and three petunia *MEI2-like 5* genes belong more broadly to AML235 clade, contrasting with the well-defined role of TEL clade genes in development rate regulation. This observation suggests a potentially novel role for AML14 clade genes in controlling vegetative growth traits. Therefore, the phylogenetic analysis supports the notion that *MEI2-like* genes, particularly those belonging to the AML14 clade, are promising candidate genes for the regulation of development rate-related traits. Further functional characterization of these *MEI2-like* genes in petunia could provide valuable insights into their specific roles in vegetative growth and development. This research avenue could lead to a deeper understanding of the unknown molecular mechanisms underlying vegetative development rate.

We propose to functionally analyze the *MEI2-like 1* gene using an efficient gene knockout method, particularly considering the need for consistent and widespread silencing to accurately evaluate the phenotype. Although the inefficient gene silencing is not uncommon in VIGS (Bennypaul et al., 2012), this method is more valuable for studies where the phenotype evaluation is straightforward and can be visually characterized. VIGS may not be ideal for studies where precise and consistent silencing is required, especially for quantitative traits or

complex phenotypes. In such cases, alternative gene editing techniques like CRISPR-Cas9 may offer better precision and control over gene manipulation. Clustered regularly interspaced short palindromic repeats (CRISPR) associated Cas9 has emerged as a popular gene editing tool for gene function analysis in various organisms, including plants (Noman et al., 2016). This method has been successfully employed in petunia for gene editing targeted towards various traits such as flower color, flower longevity, and disease resistance (Subburaj et al., 2016; Xu et al., 2021; Lin and Jones, 2022; Xu et al., 2023).

However, it is important to acknowledge several key considerations while proposing this research. Firstly, the phenotyping method for development rate should be well-defined. While our initial approach involved counting nodes on two branches and a main stem of each plant, it may be beneficial to study the variation in node number across the entire branches to capture a more comprehensive understanding of development rate. Additionally, determining the critical time period for phenotyping plants for development rate is crucial. The shoot apical meristem undergoes morphological and genetic changes as plants progress through different stages of the vegetative phase, so identifying the optimal time window for phenotypic assessment will be essential for accurate evaluation.

CONCLUSION

In summary, our study on evaluating development rate candidate genes using the rapid evaluation VIGS method offers valuable insights into the challenges associated with this approach when investigating complex traits such as development rate. While we have identified varying levels of gene expression levels, indicative of inefficient silencing, we have also identified a promising candidate gene that warrants further investigation. Moving forward, further work is needed to test the promising candidate gene using stable and more reliable silencing method, such as CRISPR-Cas9, to accurately assess its effect on development rate.

Combining CRISPR-Cas9-mediated gene editing with thorough and well-defined phenotyping methods will enable a robust analysis of the role of *MEI2-like 1* gene in controlling development rate in petunia. This approach holds promise for advancing our understanding of the genetic mechanisms underlying vegetative development rate and may inform breeding efforts aimed at improving crop timing in plants.

Tables & Figures

Table 2-1: Description of the five criteria used to select candidate genes for development rate. The first two criteria were based on the transcriptomics studies in IA and AE RILs (Guo et al., 2017). Third criterion is based on the *P. axillaris*.v.162 genome (Bombarely et al., 2016). Fourth and fifth criteria are based on the literature studies of development rate and expression profile of genes.

	Criteria
1.	Differentially expressed genes between fast- and slow-developing IA RILs.
2.	Differentially expressed genes mapping close (<1 cM) to a development rate QTL.
3.	Located on genomic scaffolds harboring SNP markers underlying a QTL.
4.	Related to pathways previously implicated in the control of development rate.
5.	Preferentially expressed in a shoot apex tissue.

Table 2-2: Summary of the candidate genes. GeneID and functional description are taken from *P. axillaris*. v1.6.2 genome in the Sol genomics network database. Third column lists the criteria (Table 2-1) related to each gene. Fourth column indicates the short form of the genes used in this study.

GeneID	**Functional description**	**Criteria matching**	**Short form**
Peaxi162Scf00572g00008.1	"Zinc finger CCCH domain-containing protein 32"	1, 2, 4	Peaxi00008
Peaxi162Scf00377g00012.1	"expansin B2"	1, 2, 4	Peaxi00012
Peaxi162Scf01147g00014.1	"Homeobox protein knotted-1-like 6"	3, 4, 5	Peaxi00014
Peaxi162Scf00919g00310.1	"squamosa promoter-binding protein-like 12"	3, 4	Peaxi00310
Peaxi162Scf00953g00316.1	"carotenoid cleavage dioxygenase _1"	1, 2, 4, 5	Peaxi00316
Peaxi162Scf00023g00929.1	"MEI2-like protein 1"	3, 4	Peaxi00929
Peaxi162Scf00829g00016.1	"Regulator of chromosome condensation (RCC1) family protein"	1, 4, 5	Peaxi00016
Peaxi162Scf00316g00027.1	"Argonaute family protein"	3, 4	Peaxi00027
Peaxi162Scf00141g00048.1	"Auxin transporter-like protein 2"	3, 4	Peaxi00048
Peaxi162Scf00062g00073.1	"scarecrow-like 3"	3, 4, 5	Peaxi00073
Peaxi162Scf01141g00125.1	"IAA-amino acid hydrolase ILR1-like 4"	1, 4	Peaxi00125
Peaxi162Scf00367g00737.1	"Leucine-rich repeat protein kinase family protein"	3, 4	Peaxi00737
Peaxi162Scf00023g01330.1	"zinc finger (C3HC4-type RING finger) family protein"	1, 4	Peaxi01330
Peaxi162Scf00461g00008.1	"CBS domain-containing protein"	1, 2	Peaxi00008.2
Peaxi162Scf01178g00015.1	"Gibberellin 20 oxidase 2"	1, 4	Peaxi00015
Peaxi162Scf00304g01115.1	"HAD superfamily subfamily IIIB acid phosphatase"	1, 4	Peaxi01115
Peaxi162Scf00069g01624.1	"squamosa promoter-binding protein-like 12"	Control	Peaxi01624

Table 2-3: Summary of the primers used for amplifying virus induced gene silencing (VIGS) target regions of the candidate genes. Forward (F) and reverse (R) primers are up to 30 bp in length excluding 15bp adapter sequences (refer to Methods section).

GeneID	**Forward (5' – 3')**	**Reverse (5' – 3')**
Peaxi162Scf00572g00008.1	gcgacaagagaatgctcctacta	cgagtaagtgctccagttctgaa
Peaxi162Scf00377g00012.1	caactatcccggagtatcactgg	gatcgataagtttttccgggc
Peaxi162Scf01147g00014.1	atggatcaacatgaaatgtatggtt	aatacttgaaccaccttcatcaatat
Peaxi162Scf00919g00310.1	tccatgaatggggataaaggc	agttgtactttccctgcctggc
Peaxi162Scf00953g00316.1	gaacaaatacaagaacaagccaa	taatttagccctgccgctatca
Peaxi162Scf00023g00929.1	gtgctgcctctagttcctattttaa	tgattgcccaatattaatagttgaaga
Peaxi162Scf00829g00016.1	ttacaaaattcccatccatgtgctgca	ccctgctccatcttttgacattccatc
Peaxi162Scf00316g00027.1	gacattttgaacagtttcttgtgcttttgg	gccttgaacttggcctcataaagg
Peaxi162Scf00141g00048.1	aatgcagcagagaaacctccatttttc	gactcttgaaattctgcaacacc
Peaxi162Scf00062g00073.1	aatggaaaacatgctatttggcgagg	cacctcttgcatctccaagacga
Peaxi162Scf01141g00125.1	gatgagtgtttgaatccgttattga	ctgaataggcaaagcatccatg
Peaxi162Scf00367g00737.1	agggttggtttcacttttaaagctag	tttaggtctgctatagactcaggtattc
Peaxi162Scf00023g01330.1	attgcgactttgatggcctc	acttatgttttgaggtcactgcg
Peaxi162Scf00461g00008.1	aaacagacgaccaaccggaatatatc	tcagaatcaccttcttccacaacctt
Peaxi162Scf01178g00015.1	gaaggtgtaccaagattactgcaatgc	tcctccaataccatcttgatggaggat
Peaxi162Scf00304g01115.1	tagcacgctttagtacccataaaatct	ctaattcaacttcatgatctttgtattctt
Peaxi162Scf00069g01624.1	gttccatgagctttccgaattt	gcttttacattttaggtttcactattaac

Table 2-4: Summary of the primers used for qPCR reactions. Primers are designed outside the VIGS target region to amplify the target gene.

Gene ID	**Forward (5' -3')**	**Reverse (5'-3')**
Peaxi162Scf00572g00008.1	cgctgcagttaggaggaagtgc	ggtggtggaacttgcatgatgc
Peaxi162Scf01147g00014.1	ggagctgatcctgaactcgatgagt	tgtagtcgcttcattgaaaggcct
Peaxi162Scf00919g00310.1	gcctggttgcgttgtgttaacg	tcaccagggacatccaaaagcc
Peaxi162Scf00953g00316.1	tgcaaacgattggcatgctgga	aggtggtgtctgtgagtagcca
Peaxi162Scf00023g00929.1	taagaggtgccgtccgatcctc	tgccaactcgacttttgctggt
Peaxi162Scf00316g00027.1	cgtgccaaagaccaagagatcg	tccacggatgcttgcttttcac
Peaxi162Scf00141g00048.1	tcttgttggtagctggactgca	aaggcctgctgctttccagtat
Peaxi162Scf00062g00073.1	agaaactgcgtgtgaaaacggg	catgggcaagaagagtgtgcaa
Peaxi162Scf00367g00737.1	ctgatctttatccagatccttgtggt	gggctgaattcaacatttggtgc
Peaxi162Scf00023g01330.1	ctcaatccacacctgcaagctc	cacttgtatcctggtccagcct
Peaxi162Scf00389g00936.1	tggtactgtccctgtcggtcgt	cgagctccttaccagatcgcctgt
Peaxi162Scf00038g02444.1	gccagcaatgcttggaggacaa	ctgtcaccctatctggcacacc
Peaxi162Scf00069g01624.1	cagaggttctgccaacaatgca	gccggcggcaactccttttagt

Figure 2-1: Overview of the VIGS protocol step-by-step starting from an inoculation stage to the data collection.

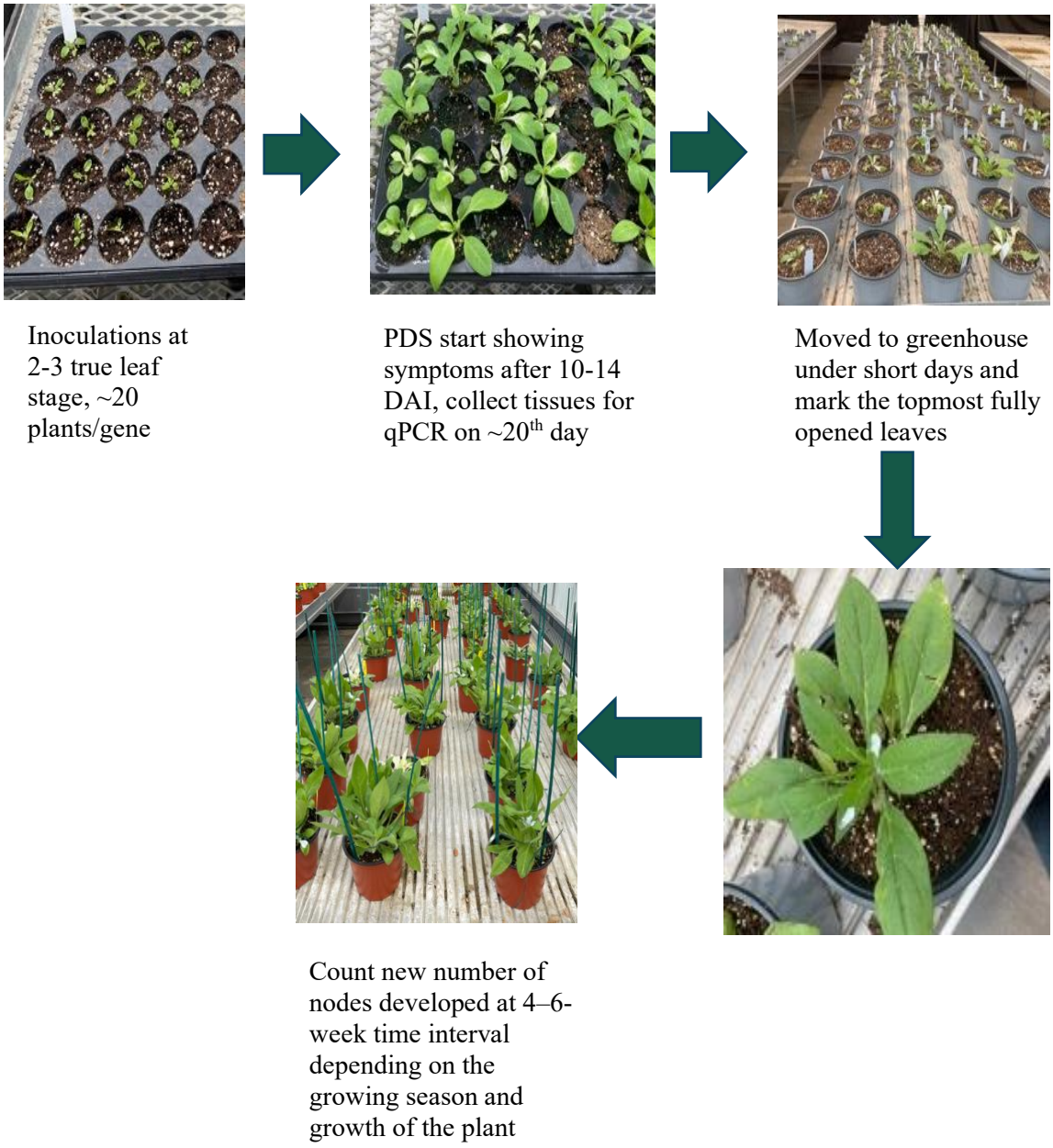


Figure 2-2: Melt curve of Peaxi162Scf00377g00012.1 with two different primer sets. Y-axis represents the negative derivative of fluorescence and x-axis represents temperature.

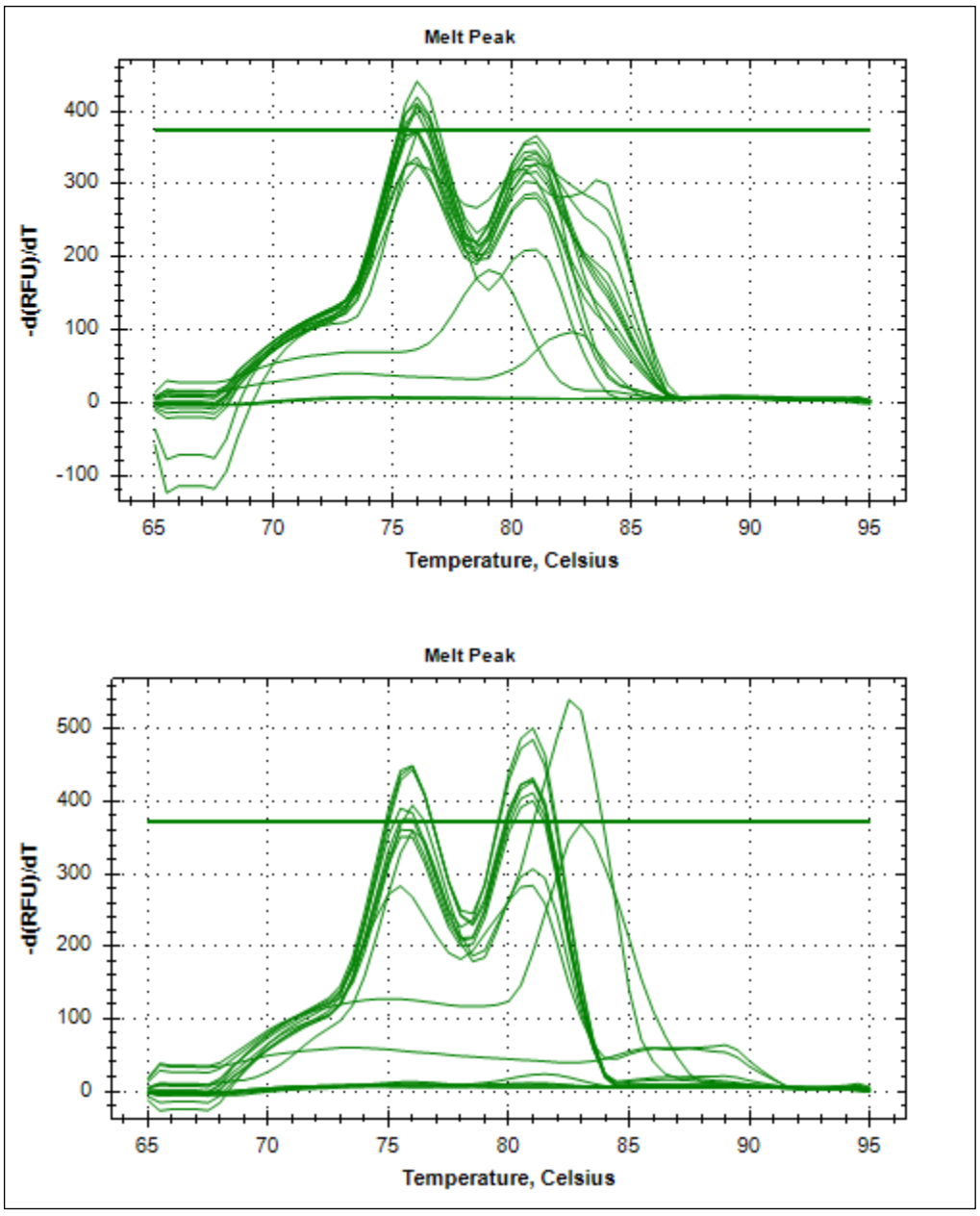


Figure 2-3: *P. axillaris* PI667515 seedlings inoculated with Peaxi162Scf00829g00016.1 (A), Peaxi162Scf01141g00125.1 (B), and *PHYTOENE DESATURASE* (*PDS*) control construct (C).

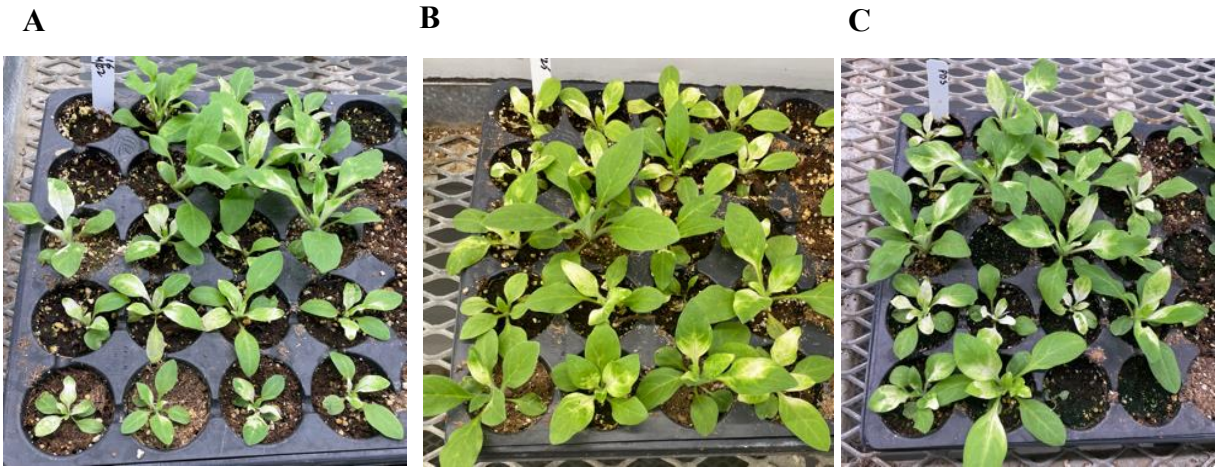


Figure 2-4: Two plants with PHYTOENE DESATURASE (PDS) gene silenced seven weeks post inoculations. PDS inoculated plants were used as positive controls for all VIGS experiments.

Figure 2-5: Leaves from three different PDS plants and their relative qPCR expression. PDS inoculated plants were used as positive controls for all VIGS experiments.

Leaf	Relative expression
PDS1	0.2
PDS2	0.52
PDS3	0.40

Figure 2-6: Boxplots comparing the node number distribution of plants inoculated with different gene sets (qPCR relative expression threshold ≤0.4). Node numbers were counted after 6-week interval for this dataset. Letters on top of boxplots separate the boxplots based on their statistical difference ($p \leq 0.05$). Y-axis represents average number of nodes on side branches (A) and number of nodes on the main stem (B). X-axis represents each gene construct. Numbers written after the dot on each gene construct indicate the number of plants for each gene construct.

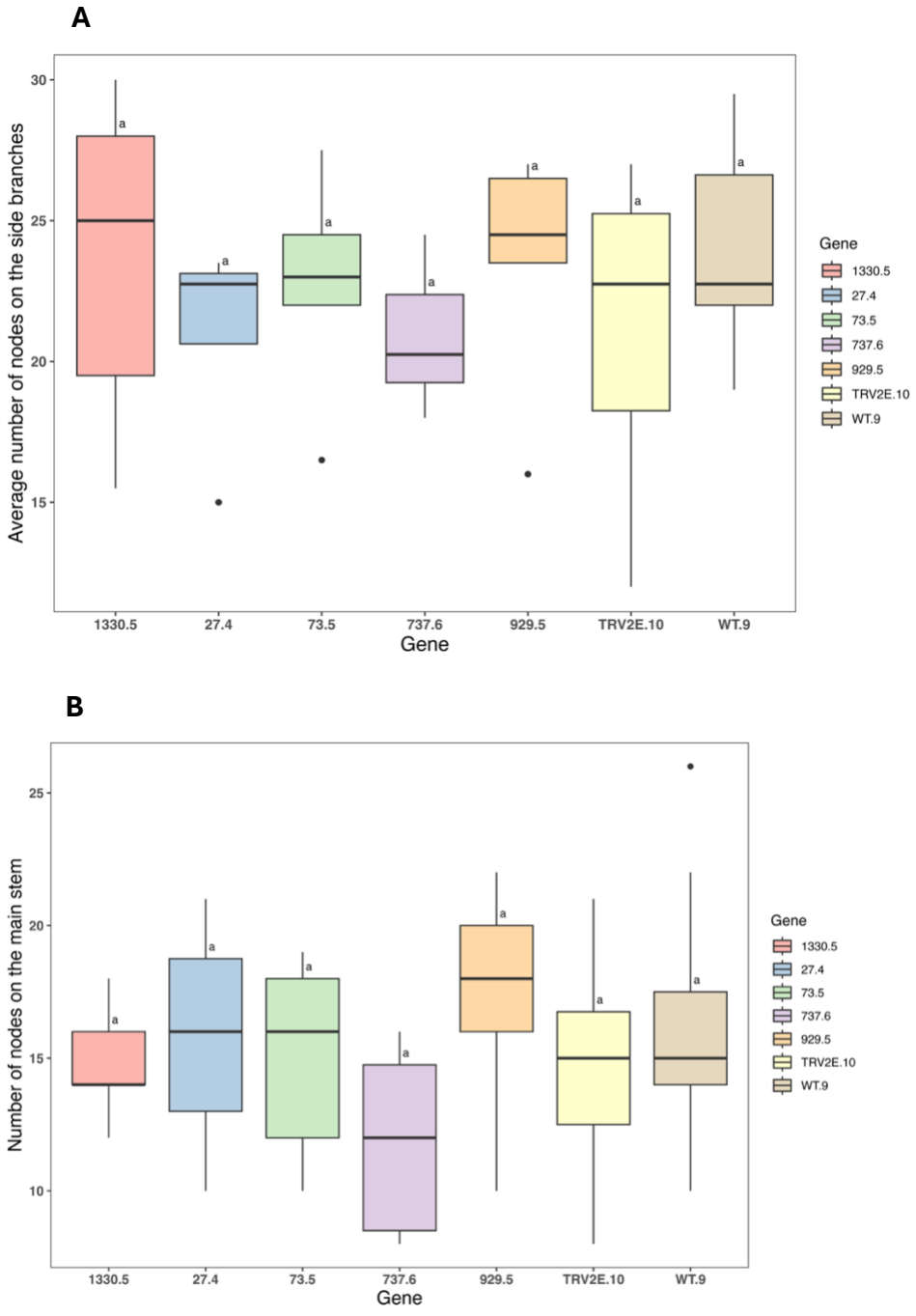
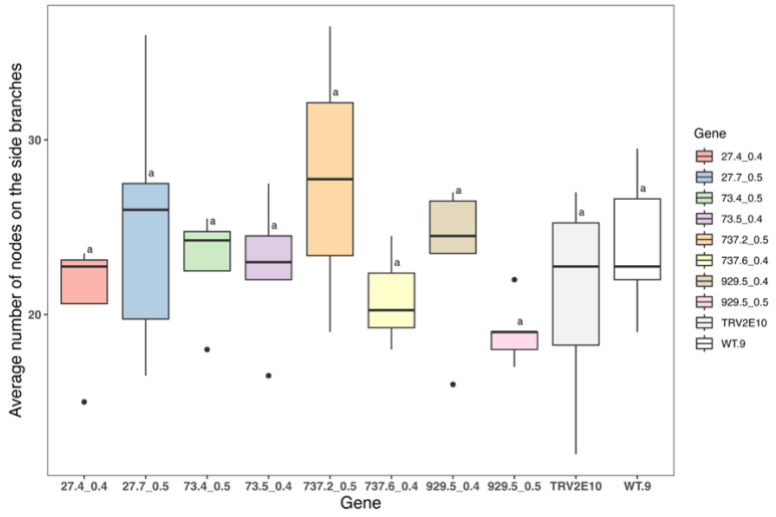
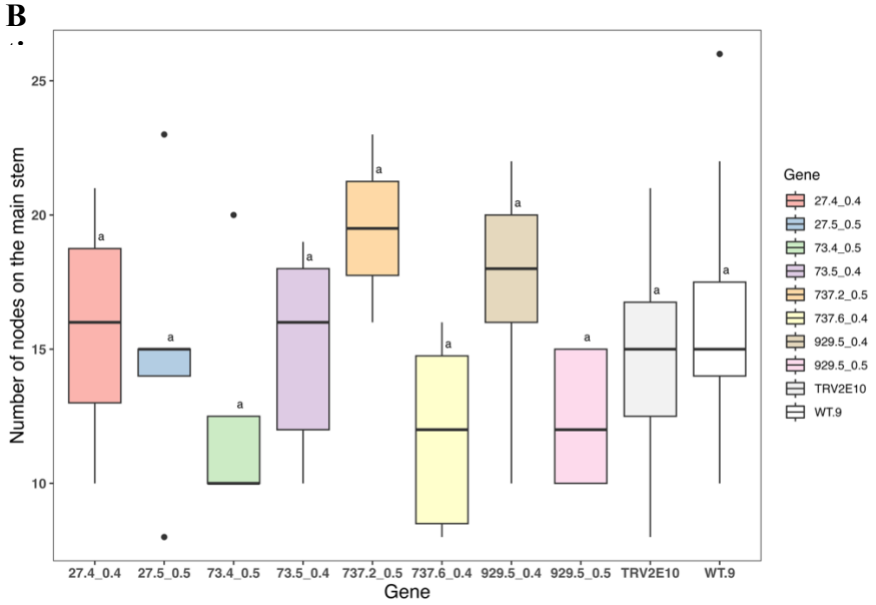


Figure 2-7: Boxplots comparing the node number distribution of plants inoculated with different gene sets with qPCR relative expression ≤0.4 and 0.5-0.95. Node numbers were counted after 6-week interval for this dataset. Letters on top of boxplots separate the boxplots based on their statistical difference ($p \leq 0.05$). Y-axis represents average number of nodes on side branches (A) and number of nodes on the main stem (B). X-axis represents each gene construct. Numbers written after the dot on each gene construct indicate the number of plants for each gene construct. Table (C) represents the range of relative expression values of plants evaluated in (A) and (B).

A



B

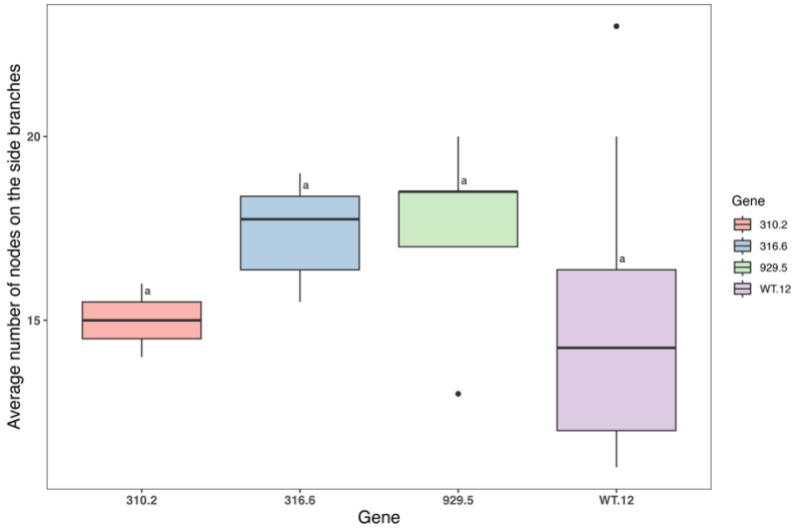


C

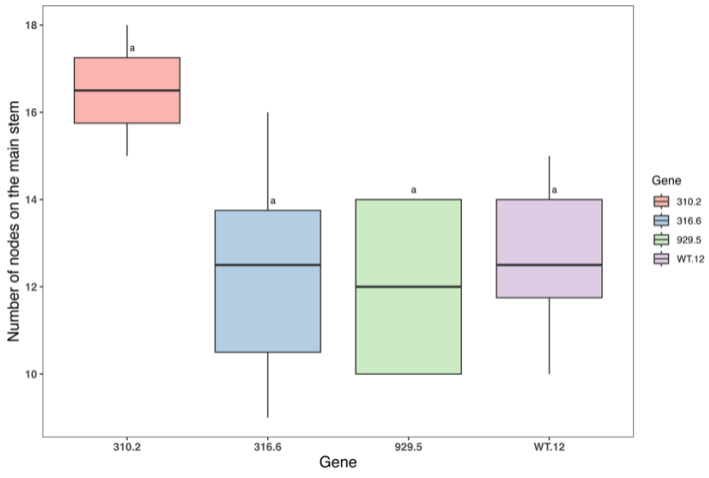
Plant	Relative expression range
27_0.5	0.47-0.93
27_0.4	0.23-0.44
73_0.5	0.48-0.68
73_0.4	0.10-0.32
737_0.5	0.53-0.60
737_0.4	0.15-0.36
929_0.5	0.49-0.83
929_0.4	0.17-0.36
1330_0.4	0.12-0.31

Figure 2-8: Boxplots comparing the node number distribution of plants inoculated with different gene sets (qPCR relative expression threshold ≤0.4). Node numbers were counted after 4- week interval for this dataset. Letters on top of boxplots separate the boxplots based on their statistical difference ($p \leq 0.02$). Y-axis represents average number of nodes on side branches (A) and number of nodes on the main stem (B). X-axis represents each gene construct. Numbers written after the dot on each gene construct indicate the number of plants for each gene construct. Table (C) represents the range of relative expression values of plants evaluated in (A) and (B).

A



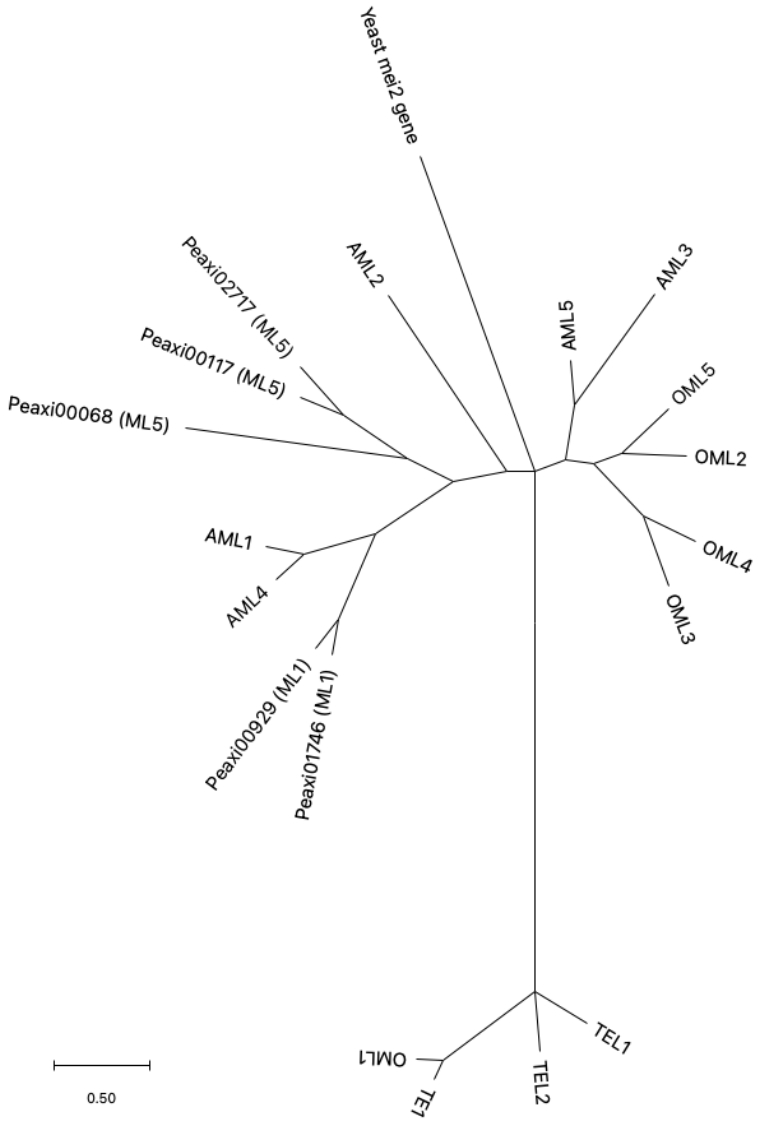
B



C

Plant	Relative expression range
310	0.08-0.13
316	0.04-0.21
929	0.07-0.21

Figure 2-9: Phylogenetic tree of MEI2-like genes from *Arabidoposis thaliana* (AML1 -AML5, TEL1 and TEL2), *Oryza sativa* (OML1-OML5), *Zea mays* (TE1), *Petunia axillaris* (three *MEI2-like protein 5* and two *MEI2-like protein 1* genes) and a MEI2-like gene from *Schizosaccharomyces pombe* (Yeast Mei2).



CHAPTER 3

TRANSCRIPTOMIC ANALYSIS OF PETUNIA RECOMBINANT INBRED LINES (RILS) WITH CONTRASTING DEVELOPMENT RATES

INTRODUCTION

Vegetative development rate, defined as the rate of nodes or leaves produced over time (expressed as nodes day^{-1}) (Warner and Walworth, 2010), is a crucial biological phenomenon in crop plants. This rate directly affects crop timing, such as the time to first yield, and indirectly influences seasonal biomass accumulation in forage and bioenergy crops by determining the number of leaves available for photosynthetic carbon fixation. Leaf production is initiated by the shoot apical meristem during post-embryonic vegetative shoot development (Stuurman et al., 2002), and it varies across species and within germplasm pools. This variability suggests that it is possible to alter the development rate to enhance production efficiency (e.g., early flowering or fruiting), yield, and biomass accumulation. Therefore, studying the genetic control of development rate is highly desirable.

Petunia (*Petunia × hybrida*) ranks among the top annual bedding crops in the United States, with a wholesale sales value of approximately $160 million in 2020 (USDA, 2021). It is renowned for its diverse flower colors and is produced in greenhouses during the colder months in northern latitudes of North America and Europe (Guo et al., 2017). To meet spring production demands, greenhouses must be heated to ensure sufficient temperatures for flowering, resulting in significant energy costs and narrow profit margins for producers (Guo et al., 2015). Developing varieties with faster development rates at cooler temperatures could reduce these energy costs. Therefore, understanding the genetic mechanisms that control development rate is crucial. This knowledge can inform strategies to enhance development rates under sub-optimal temperatures, thereby optimizing crop timing and increasing the production efficiency of seasonal crops.

To investigate the genetics of development rate in petunia, F_7 recombinant inbred lines (RILs) derived from the wild progenitor species *P. integrifolia* and *P. axillaris* (the IA population) were previously utilized (Guo et al., 2017). This study identified 209 differentially expressed genes (DEGs) between RILs exhibiting slow and fast development rates, as well as quantitative trait loci (QTL) associated with development rate. Some of these DEGs co-localized with QTL for development rate and/or had functions related to vegetative development as indicated in the literature, while many others had not previously been implicated in the control of development rate. Thus, these studies generated evidence for the genetic regulation of development rate in petunia, complementing the existing knowledge in the literature. However, despite these advancements, gaps in our understanding of this trait still exist.

The current study aimed to understand the genetics of development rate by employing a second RIL population developed from a cross between *P. axillaris* and *P. exserta* (the AE population) (Guo et al., 2017). As *P. axillaris* and *P. exserta* diverged more recently than *P. axillaris* and *P. integrifolia* (Chen et al., 2007), this narrower cross was expected to identify novel genetic components. In 2014 and 2015, 171 AE RILs were grown under long-day conditions in the greenhouse (Guo et al., 2017). These RILs were phenotyped for development rate and other crop timing-related traits at three temperatures (14, 17, and 20°C). From this work, fast- and slow-developing RILs were carefully selected for the current study and further evaluated under greenhouse conditions to ensure robust identification of fast- and slow-development RILs. The primary objective was to explore the transcriptomics of these RILs, aiming to pinpoint genes that are consistently differentially expressed between fast- and slow-developing IA and AE RILs, and to leverage the narrower AE cross to uncover novel DEGs potentially involved in development rate control. Additionally, the study sought to identify a

subset of genes that co-expressed, shedding light on the pathways associated with development rate.

MATERIALS AND METHODS

Plant materials

From the previous study by Guo et al. (2017), 12 and 13 RILs were consistently identified at the fast and slow ends of development rate, respectively, of the entire AE population across both years and three temperatures. To conduct the current study, seeds were sown in a 72-cell tray with each cell having a volume of 16.4 cm^3 under short-day conditions (9 hours of light/22°C) in a growth chamber. Once the seedlings reached 4-6 nodes, they were transplanted into 15.24 cm round pots with a volume of 1420.76 cm^3 and were grown in two replications following a randomized complete block design (RCBD) in a greenhouse under short-day conditions (9 hours of light/22°C). Each RIL was represented by at least five and maximum twenty-five plants. To track the number of new leaves produced over time, the edges of the two topmost fully expanded leaves on the main stem and two side shoots of each plant were marked with white paint. After four weeks from the date of marking, the number of new nodes/leaves produced were counted.

Data analysis

One-way ANOVA and Tukey's HSD post-hoc tests were employed to identify the RILs with significant mean node number differences on both the main stem and side shoots (α=0.05). These statistical analyses were conducted using SPSS version 27 (IBM; Chicago, IL).

RNA extraction and sequencing

Shoot apex tissue samples of ca. 2 mm length from the tip were harvested and leaf primordia were removed as much as possible with forceps. These samples were collected from

all plants of the selected 13 RILs from both the replications, pooled for each RIL and stored at -80°C. Total RNA was extracted by using MagMAX™ Plant RNA Isolation Kit (Catalog #A33784, Thermo Fisher Scientific). RNA samples were evaluated for Quality Control (QC) after running the samples on the TapeStation Analysis Software 3.2 at MSU's Research Technology Support Facility (RTSF) Genomics core facility. RNA samples from ten RILs and their biological replicates (five each of fast and slow node production) with an RNA integrity (RIN) score ≥ 5.5 were selected for sequencing, resulting in a total of 20 samples. Libraries were prepared using the Illumina TruSeq Stranded mRNA Library Prep Kit with IDT for Illumina Unique Dual Index adapters following the manufacturer's recommendations. Completed libraries were QC'd and quantified using a combination of Qubit dsDNA HS and Agilent 4200 TapeStation HS DNA1000 assays. All 20 libraries were pooled in equimolar amounts and the pool was quantified using the Kapa Biosystems Illumina Library Quantification qPCR kit. This pool was loaded onto two lanes of an Illumina HiSeq 4000 Single Read flow cell. Sequencing was performed in a 1x50bp single end read format using HiSeq 4000 SBS reagents. Base calling was done by Illumina Real Time Analysis (RTA) v2.7.7 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.20.0.

Quantification of transcripts

The raw reads were trimmed for adaptor sequences and low quality using Trimmomatic version 0.39 (Bolger et al., 2014), and the quality of these reads was assessed using FastQC (Andrews, 2017). As sequencing for each sample was conducted on two separate lanes, reads from both lanes were merged into a single read. The reads were aligned to the *P. axillaris* genome v.1.6.2 (Bombarely et al., 2016) using STAR (Dobin et al., 2013). HTSeq (Putri et al.,

2022) was then employed to count the number of reads for each gene, encompassing a total of 35,851 genes across all samples.

Identification of differentially expressed genes (DEGs)

DEGs analysis was conducted on the raw counts using the DESeq2 package in R with default parameters (Love et al., 2014). Before proceeding with the DEGs analysis, clustering of samples was performed using a Principal Component Analysis (PCA) plot and a heatmap of sample distances based on regularized logarithm (rlog) transformed counts. Outlier samples were removed from further analysis, resulting in three samples each of fast and slow categories. Subsequently, DEGs analysis was performed on three RILs each from the fast and slow node-producing categories. DEGs were identified using two approaches.

In the first approach, the three RILs (and their biological replicates) from each fast and slow node-producing category were pooled into two conditions: "Fast" and "Slow", and DEGs between these two conditions were identified. In the second approach, each fast and slow genotype was considered as a separate condition, resulting in a total of six conditions: "Fast1", "Fast2", "Fast3", "Slow1", "Slow2", and "Slow3". Biological replicates were pooled for each condition. This approach aimed to identify DEGs that are robustly differentially expressed between RILs of fast and slow node-producing categories, resulting in a total of nine comparisons. DEGs were filtered using a p-value threshold of 0.05 and a log fold change threshold of |1.5|.

Weighted gene co-expression network analysis (WGCNA)

The WGCNA R package was employed to construct the coexpression network (Pei et al., 2017). Initially, samples were clustered using the 'hclust' function to detect and eliminate any outliers from further analysis. The R function 'pickSoftThreshold' was then utilized to calculate

the soft threshold power, employing "signed" networks and the "bicor" correlation function to build the adjacency matrix. A soft power of 22 was chosen based on an R2 fit of greater than or equal to 0.85. The Topological Overlap Matrix (TOM) was computed using the adjacency matrix, and gene dendrograms were plotted based on their dissimilarity. Hierarchical clustering and the dynamic tree cut function were subsequently employed to detect modules, with a tree cut height threshold of 0.25 used to cluster the module eigengenes. Gene significance (GS) and module membership (MM) were calculated to establish the relationship between modules and the development rate trait. Hub genes were identified from each module using MM >= 0.8 and GS >= 0.8 as thresholds. The corresponding module gene information was then extracted for further analysis.

Gene ontology (GO) enrichment analysis

Gene ontology (GO) enrichment analysis was conducted on both up- and down-regulated genes, as well as genes within significant modules, utilizing the GO enrichment tool in the PlantRegMap program (Tian et al., 2020). The analysis utilized topGO and Fisher's exact tests to identify significantly over-represented GO terms (p-value < 0.05) within the input gene set, with all genes in P. axillaris serving as the background. Venn diagrams were generated using Venny 2.1.0 (Oliveros, 2016)to visualize the overlap of enriched GO terms between different gene sets.

RESULTS

Selection of slow and fast lines

Selecting plants at the extremes of slow and fast development rates proved challenging due to the significant phenotypic variability in development rate. We focused on selecting RILs that were falling under one category (slow or fast) for at least three data points. Of the twenty-five evaluated RILs, thirteen RILs (highlighted in bold) falling under contrasting development

rate groups were selected for shoot apex tissue collection (Table 3-1). These thirteen RILs consisted of six fast (denoted with f; 45f, 85f, 110f, 173f, 193f and 319f) and seven slow developing RILs (denoted with s; 208s, 219s, 216s, 252s, 279s, 298s and 318s). Subsequently, RNA extraction was performed on these thirteen lines, each with two biological replicates. Specifically, five lines were selected from both the fast (45f, 85f, 110f, 173f and 319f) and slow (208s, 219s, 216s, 252s and 318s) development groups, ensuring that the chosen RNA samples had an RNA integrity (RIN) score of ≥ 5.5 (Table 3-2).

Processing of reads

RNA sequencing was performed to generate at least 25 million single-end 50 bp raw reads per sample. Quality control processes ensured that at least 99.7% of the reads survived filtering for adaptor sequences and low-quality reads. The percentage of uniquely mapped reads to the *P. axillaris* genome ranged from 84% to 91%, while 74% to 82% of the reads mapped to the exon regions of genes (Table 3-2). The number of reads for each gene in each sample was counted to identify differentially expressed genes.

Differential gene expression analysis

Pooled comparison

PCA analysis revealed that the samples did not clearly group into distinct clusters based on slow and fast development rates, showing significant overlap instead (Fig. 3-1). Consequently, obvious outlier samples were removed, and samples from each category that formed close clusters were selected for further analysis. Specifically, three samples from the fast development group (AE319f, AE110f, and AE45f) and three from the slow development group (AE252s, AE219s, and AE216s) were chosen, with two biological replicates each, except for AE216s, which had one replicate. These samples were used for differential expression analysis (Fig. 3-1).

In the pooled comparison, 1834 genes were downregulated and 260 genes were upregulated in the slow development lines compared to the fast development lines (Fig. 3-2). Among the upregulated genes were those encoding *GIBBERELLIN 2-OXIDASE* (Peaxi162Scf00111g00920.1), an *AUXIN EFFLUX CARRIER FAMILY PROTEIN* (Peaxi162Scf00033g00711.1) and *CAROTENOID CLEAVAGE DIOXYGENASE 8* (Peaxi162Scf00227g00714.1). Additionally, upregulated gene families included UDP-Glycosyltransferase superfamily protein, cytochrome P450 families 71, 76, 718, MATE efflux family protein, 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (2OG-Fe(II)).

Predominant gene families among the downregulated genes included the 2-oxoglutarate 2OG-Fe(II)-dependent oxygenase superfamily protein family (2OG-Fe(II)), basic helix-loop-helix (bHLH) DNA-binding superfamily protein, Pentatricopeptide repeat-containing protein (PPR), F-box/FBD/LRR-repeat protein, mitogen-activated protein kinase 17, 19, 20, 21 and MATE efflux family protein.

Pairwise comparison

Each individual slow and fast line was compared to identify robustly differentially expressed genes (Table 3-3 and Fig. 3-3). Generally, more genes were downregulated than upregulated in the slow development rate lines. Core DEGs were defined as genes that were commonly differentially expressed in at least six pairwise comparisons (Fig. 3-3). Only 14 core upregulated genes were identified, which were commonly upregulated in two slow lines (219s and 216s) (Fig. 3-3B).

A total of 271 downregulated core DEGs were identified in all slow lines compared to the fast development rate lines (Fig. 3-3A). These genes included *AUXIN RESPONSIVE FAMILY*

PROTEIN (Peaxi162Scf00745g00810.1, Peaxi162Scf00006g00111.1, Peaxi162Scf00352g00718.1, Peaxi162Scf00945g00013.1), *PPR CONTAINING PROTEIN* (Peaxi162Scf00037g01030.1, Peaxi162Scf00604g00016.1), cytochrome P450 families 71 (Peaxi162Scf00032g01117.1, Peaxi162Scf00109g00064.1) and 718 (Peaxi162Scf00322g01519.1, Peaxi162Scf00474g00416.1), *2OG-Fe(II)-DEPENDENT OXYGENASE SUPERFAMILY PROTEIN* (Peaxi162Scf00045g01928.1), *DNA BINDING PROTEIN* (Peaxi162Scf00064g00423.1, Peaxi162Scf00931g00117.1), *RNA BINDING PROTEIN* (Peaxi162Scf00036g00222.1), *CELLULASE SYNTHASE LIKE C5* (Peaxi162Scf00160g00715.1), *bHLH DNA-BINDING SUPERFAMILY PROTEIN* (Peaxi162Scf00075g01418.1), *MATE EFFLUX FAMILY PROTEIN* (Peaxi162Scf00684g00553.1) and the meristem identity gene *WUSCHEL* (Peaxi162Scf00083g00516.1).

WGCNA analysis

Sample clustering dendrogram revealed three distinct clusters (Fig. 3-4A): one cluster containing the three fast samples (AE319f, AE110f, and AE45f), another with two slow samples (AE252s and AE219s), and a third cluster with sample AE216s-2, consistent with previous clustering in DEGs analysis. The first two clusters were selected for further analysis (Fig. 3-4B). After filtering genes with low counts or many missing values, 18,842 genes across 11 samples remained. Using a soft power threshold of 22, 48 modules were identified. Among these, six modules each were positively and negatively correlated with the development rate phenotype ($p < 0.05$ and $R^2 > 0.6$) (Fig. 3-5). The number of total genes per module ranged from 54 to 2,429 (Table 3-4), with hub genes ranging from 2 to 310 per module (Table APP-3-1).

Robust modules included bisque4, deepskyblue4, and orange1, which were negatively correlated with all slow line samples (Fig. 3-6A), and black, mistyrose2, sienna2, and tan2, which were positively correlated with all slow line samples (Fig. 3-6B). Other modules showed varying eigengene values within samples of the same development rate categories. Eigengene values represent a common gene expression matrix of the modules (Langfelder and Horvath, 2008).

Genes in each module were compared to the core genes identified in pairwise comparisons. Module bisque4 had 54 genes, and orange1 had only 3 genes in common with the downregulated core genes, whereas module sienna2 had 6 genes in common with the core upregulated genes (Table 3-6). The common genes included *bHLH DNA-BINDING SUPERFAMILY PROTEIN* (Peaxi162Scf00075g01418.1), *MADS-BOX TRANSCRIPTION FACTOR 3* (Peaxi162Scf00022g00098.1), *WUSCHEL* (Peaxi162Scf00083g00516.1), and *AUXIN-RESPONSIVE GH3 FAMILY PROTEIN* (Peaxi162Scf00945g00013.1), among others.

Gene Ontology (GO) analysis

Gene ontology (GO) enrichment analysis was performed on the differentially expressed genes and significant module genes to understand their roles in biological processes, cellular components, and molecular functions. This analysis identified 40 GO enriched terms for pooled upregulated genes and 110 for downregulated genes (Table APP-3-2). In the pairwise comparison, 69 GO terms were enriched for the downregulated core DEGs. Further GO analysis of the modules revealed 181 GO enriched terms for bisque4, 108 for deepskyblue4, 70 for orange1, 321 for black, 50 for mistyrose2, 348 for sienna2, and 35 for tan2 (Table APP-3-2).

DISCUSSION

Petunia wild species demonstrate faster development rates compared to commercial cultivars at similar temperatures, implying the possibility of breeding varieties with accelerated development rates (Warner and Walworth, 2010). Given the significance of wild species as a rich resource for investigating the genetics of development rate and the potential for narrower crosses to introduce novel genetic variability, we conducted transcriptomic analysis. Specifically, we analyzed six AE RILs exhibiting divergent development rates (three fast and three slow), meticulously phenotyped for development rate prior to analysis.

Nevertheless, the distinction between selected fast and slow lines based on their normalized read counts was not clear on the PCA plot. The indistinct grouping of slow and fast lines could stem from various factors. Firstly, the trait itself displays substantial natural variation, complicating the separation of genotypes. Moreover, interference from other inherent traits might interfere with the explicit clustering of genotypes with contrasting development rates.

The identification of more downregulated genes suggests that the slow development rate phenotype results from the decreased expression of a large set of genes and the increased expression of a small set of genes.

Previous studies have provided insights into the development rate, indicating that plastochron—the inverse of development rate, defined as the time interval between two successive nodes (Guo et al., 2015), is controlled by multiple independent pathways rather than a single unified model. For instance, in rice, genes such as *PLA1*, *PLA2*, and *PLA3*, which encode the cytochrome P450 family protein *CYP78A11*, *MEI2-LIKE RNA BINDING PROTEIN*, and *GLUTAMATE CARBOXYPEPTIDASE*, respectively, negatively regulate the development rate (Miyoshi et al., 2004; Kawakatsu et al., 2009). These genes exhibit pleiotropic effects, including

the regulation of plant height, changes in shoot apical meristem size, and the transition from vegetative to reproductive phases, underscoring the complex regulation of this trait.

In our current study, while we observed differential expression of multiple genes related to various cytochrome P450 families, this does not directly imply their role in regulating development rate. However, we identified that a cytochrome P450 family protein (Peaxi162Scf01514g00021.1) is associated with gibberellin biosynthetic and metabolic processes, and another cytochrome P450 protein (Peaxi162Scf00814g00018.1) is linked to cytokinin biosynthetic processes, as indicated by GO terms. Given that the cytochrome P450 family is large and involved in diverse functions, with the possibility that members of the same families and sub-families participate in different pathways (Bak et al., 2011), it is crucial to understand the role of this gene family in development rate. This includes exploring their involvement in phytohormone synthesis and conducting in-depth studies on different sub-families to elucidate their specific functions.

The differential expression of genes related to the MATE efflux family suggests that transport molecules may play a role in regulating development rate, as previously proposed. The Multidrug and Toxic Compound Extrusion (MATE) family, a large family involved in various pathways, including phytohormone transport and the movement of other substrates within a cell (Suzuki et al., 2015; Upadhyay et al., 2019), might have significant transport activity that impacts leaf initiation rate. For instance, the *MND8* gene in barley and its *Arabidopsis* ortholog *BIGE1,* which encodes a MATE transporter, have been shown to negatively regulate development rate (Suzuki et al., 2015; Hibara et al., 2021). Additionally, *BIGE1* is involved in the feedback regulation of the CYP78A pathway, indicating its role in development rate regulation through cell transport activities and/or CYP78A regulation (Wang et al., 2008). This

potential involvement necessitates further investigation to fully understand the role of MATE transporters in regulating development rate.

The notable alteration in the expression of numerous genes associated with the 2OG-Fe(II) dependent oxygenase superfamily protein is intriguing. In a prior study, another gene from this family was identified as differentially expressed and closely mapped to a development rate QTL (Guo et al., 2017) in the petunia IA population. Members of the 2OG-Fe(II) dependent dioxygenase superfamily, including genes responsible for catalyzing gibberellin (GA) biosynthesis and inactivation reactions (such as *GA2oxs*, *GA3oxs*, and *GA20oxs*), play a pivotal role in maintaining the endogenous GA balance (Li et al., 2019; Kaur and Das, 2023). Gibberellin serves as a crucial hormone in regulating various aspects of plant development, encompassing stem elongation, meristem maintenance, phase transitions, flowering, seed maturation and germination (Peng and Harberd, 2002; Ogawa et al., 2003; Jasinski et al., 2005; Schwarz et al., 2008; Zhang et al., 2009; Bao et al., 2020). Additionally, its specific role in relation to development rate has also been explored. In rice, GA signaling positively influences *PLA1* and *PLA2* genes, thereby prolonging plastochron and reducing the development rate (Mimura et al., 2012).

In our study, a gene encoding *GIBBERELLIN OXIDASE 2* (*GA2ox*) was found to be upregulated in the slow lines, alongside other genes belonging to the 2OG-Fe(II) dependent oxygenase superfamily. It is conceivable that an endogenous gibberellin signal upregulates the expression of *PLA1*-like genes in petunia, thereby slowing down the development rate. Consequently, *GA2ox* is upregulated as a feedback mechanism to maintain the endogenous GA balance by inactivating the biologically active GAs, as observed in previous studies (Mimura et al., 2012). While this explanation appears plausible, the differential expression of GA-related

genes may or may not be directly associated with development rate in our study due to its pleiotropic effects on shoot apical meristem (SAM)-related functions. A future investigation specifically focusing on the impact of gibberellin signaling on regulating development rate while closely examining SAM would provide a more comprehensive understanding into the regulation of development rate.

Two genes related to auxin polar transport, namely *AUXIN EFFLUX CARRIER PROTEIN* and *CAROTENOID CLEAVAGE DIOXYGENASE 8*, were upregulated in the slow lines. Polar auxin transport (PAT) is critical for leaf initiation at SAM, mediated by auxin efflux carrier proteins known as PIN proteins (Forestan and Varotto, 2012). In the shoot apical meristem (SAM), an auxin gradient is established as newly formed leaf primordia act as auxin sinks, depleting auxin in neighboring cells and creating an auxin maximum at distant sites where new leaf primordia can form (Reinhardt et al., 2000; Reinhardt et al., 2003). Polar auxin transport also regulates a class of cell wall-loosening enzymes called expansins, which facilitate the formation of bulges at sites in the SAM where new leaf primordia develop (Fleming et al., 1997; Reinhardt et al., 1998). In a previous petunia IA RILs study, *CAROTENOID CLEAVAGE DIOXYGENASE 1* and *EXPANSIN B2* (Peaxi162Scf00953g00316.1 and Peaxi162Scf00377g00012.1), were differentially expressed and mapped close to a development rate QTL (Guo et al., 2017).

It is plausible that these proteins are overexpressed as part of a feedback mechanism in response to the slower development rate, signaling the SAM to maintain the integrity of leaf initiation events. The observation of fewer development rate-specific genes and a greater number of genes related to broader developmental pathways suggests that the SAM might be undergoing

structural and molecular changes, thereby regulating genes associated with phase transitions. Additionally, these differences could be influenced by both genotype and environmental factors.

GO terms associated with basic helix-loop-helix (bHLH) proteins include regulation of transcription DNA-templated, transcription factor activity, red or far-red light signaling pathway, cellular response to red or far-red light, regulation of circadian rhythm, positive regulation of circadian rhythm, and entrainment of the circadian clock. The bHLH DNA-binding superfamily protein is a large family of transcription factors (TFs) characterized by a N-terminal basic DNA binding domain and a C-terminal protein interaction domain (Anderson et al., 1997). These proteins play pleiotropic regulatory roles in plant growth and development, including the regulation of phytohormone cross-talk, flowering time, and clock-derived signaling pathways (Anderson et al., 1997; Hao et al., 2021). Among the bHLH family, phytochrome-interacting factors (PIFs) are key transcription factors involved in light signaling pathways, including both phyA and phyB signaling in *Arabidopsis* (Huq and Quail, 2002; Jing and Lin, 2020). Beyond flowering time regulation, bHLH family members also participate in flower organ development and floral morphogenesis (Heisler et al., 2001; Groszmann et al., 2010). Although the GO terms and existing evidence point to their roles in reproductive phase regulation, it is plausible that bHLH proteins might also regulate other downstream genes related to development rate through their involvement in phytohormonal signaling. The overlap of genes involved in development rate and phase change remains an area of interest. Currently, we cannot definitively explain the role of bHLH transcription factors in regulating development rate based on the available evidence, but this family represents a critical area for further study.

Similarly, several MADS-box transcription factors identified were associated with GO terms such as specification of organ identity, specification of floral organ identity, post-

embryonic organ morphogenesis, and floral organ formation, suggesting that SAM signals may activate genes related to the reproductive phase.

The differential expression of genes related to mitogen-activated protein kinase (MAPK) elucidates the role of post-transcriptional modifications through phosphorylation of downstream signaling targets or transcription factors, leading to altered gene expression (Cristina et al., 2010; Zhang and Zhang, 2022). Specifically, the interplay between MAPK pathways and transcription factors, such as bHLH and MADS-box, highlights how MAPK cascades modulate the expression of genes regulated by these transcription factors through phosphorylation (Wei et al., 2018). Therefore, signal transduction pathways and transcription factors are critical for understanding the regulation of development rate and warrant further investigation.

Similarly, several genes from the pentatricopeptide repeat (PPR) containing protein family, which are known to regulate genes involved in reproductive processes such as embryogenesis, gametogenesis, and seed development (Liu et al., 2013; Li et al., 2018), were differentially expressed in this study. The PPR family is involved in the post-transcriptional modification of organellar genes, relying on its RNA binding activity (Lurin et al., 2004; Barkan and Small, 2014). This finding aligns with previous research in IA petunia RILs, where a PPR family gene (Peaxi162Scf01021g00215.1) was found near genomic scaffolds harboring SNP markers associated with a development rate QTL (Guo et al., 2017). Although it is not definitively known whether this gene is directly related to development rate due to the presence of multiple genes within the QTL region, the evidence suggests that the PPR family is an important candidate for future studies on development rate.

Based on the discussion above, a substantial proportion of differentially expressed genes are associated with reproductive phase functions, such as morphogenesis, pollen development,

and floral organ development. This finding is further supported by GO analysis of highly significant modules, which revealed functions related to meiosis, pollen development, and gametophyte development. Additionally, significant modules exhibited functions relevant to both vegetative and reproductive phases, including leaf morphogenesis, meristem development, post-embryonic development, meiotic chromosome segregation, shoot system development, regulation of flower development, and floral whorl development. These functional terms suggest a potential overlap between genes involved in vegetative phase processes, reproductive phase processes, and/or phase transition.

For instance, MEI2 (meiotic inducer 2) gene family (Jeffares et al., 2004), consisting of three functional clades play role in both meiosis and development rate. AML (*Arabidopsis*-meiotic like) clade genes participate in both vegetative growth and meiosis and are expressed in both vegetative and reproductive tissues (Kaur et al., 2006). Similarly, loss-of-function mutants of the TEL (terminal ear-like clade) gene *TE1* exhibit an accelerated development rate and dwarf architecture (Veit et al., 1998). Furthermore, the *MND1* gene, which regulates plastochron, also influences phase transition (Hibara et al., 2021). Reports indicate that some transcription factors function in both vegetative and reproductive phases by interacting with different elements to regulate a set of target genes (Gregis et al., 2013). This evidence suggests that genes involved in development rate exhibit pleiotropic effects related to both vegetative and reproductive phases. Therefore, it is crucial to study in detail the gene families involved in the reproductive phase that were found to be differentially expressed in our study.

To functionally characterize the roles of these genes, if any, in the development rate phenomenon, it is crucial to ensure that the shoot apical meristem (SAM) is structurally and molecularly in the vegetative phase. In our study, meristematic tissues, along with remnants of

surrounding leaf tissue, were macroscopically collected during the vegetative stage. However, to validate that the meristematic tissue indeed corresponds to the vegetative stage, visualization through sectioning under confocal microscopy could be employed (Lian et al., 2021). Understanding the structural and molecular changes occurring in the SAM will enhance our comprehension of the genetics underlying vegetative development rate, explaining the diverse set of gene families identified when comparing fast and slow developing plants.

Several genes common between IA and AE RILs (Table 3-6) are associated with plant cell wall mechanics such as *PECTINACETYLESTERASE FAMILY PROTEIN (PAE)*, *PECTIN METHYLESTERASE INIHIBITOR SUPERFAMILY PROTEIN (PMEI)*, which play essential role during various stages of the plant life cycle, including cell division, elongation, and differentiation (Cosgrove, 2016; Houston et al., 2016). Pectinacetylesterases (PAEs) and pectin methylesterases (PMEs), including DUF proteins, play significant roles in cell wall pectin dynamics by modulating pectin acetylation and methyl esterification, respectively (de Souza et al., 2014; Salazar-Iribe et al., 2016; Coculo and Lionetti, 2022). This modulation regulates cell growth and shape by affecting the remodeling and physicochemical properties of cell wall polysaccharides, thereby influencing cell extensibility (Gholizadeh, 2020).

Genes in these families are crucial for plant growth and development. For example, loss-of-function mutants of *pae* exhibit a significant increase in total cell wall acetate levels in *Arabidopsis* leaves and a decrease in inflorescence stem height. Furthermore, pectin deacetylation impairs cell elongation of floral organs and the germination and growth of pollen tubes in tobacco (Gou et al., 2012; de Souza et al., 2014; Houston et al., 2016). In addition to regulating growth, plant cell wall remodeling is integral to the heat response network (Wu et al., 2018; Ezquer et al., 2020). The differential expression of *HEAT SHOCK PROTEIN 21* and cell

wall remodeling enzymes in both IA and AE studies is consistent with previous observations where PME-related genes were upregulated along with heat shock proteins in response to heat stress (Pineda-Hernández et al., 2022). This prompts an intriguing question regarding potential shared mechanisms connecting the development rate and the response to heat stress.

Similarly, another gene related to the cell wall, *L-ASCORBATE OXIDASE* (*AO*), was commonly differentially expressed in both studies. *AO* is involved in rapid cell wall loosening and cell expansion mechanisms (Smirnoff and Wheeler, 2000). These mechanisms have been shown to accelerate plant development, as demonstrated by the overexpression of *AO* in tomato, leading to earlier flowering (Stevens et al., 2017). The differential expression of genes related to cell wall remodeling and loosening, coupled with existing evidence linking cell wall mechanisms to development rates, underscores the importance of further investigating these genes.

Another common gene between both studies, *LAG1*, belongs to a gene family involved in the synthesis of ceramides, which are lipid second messengers crucial for various cellular processes, including the determination of cell polarity (Venkataraman and Futerman, 2002). Studies have shown that ceramide depletion leads to defective targeting of auxin polar carriers, *AUX1* and *PIN1*, resulting in auxin-dependent inhibition of lateral root emergence (Markham et al., 2011). Given the role of *PIN1* proteins in regulating development rates, *LAG1* is a significant candidate for further study, particularly regarding its auxin-related activities and their impact on plant developmental processes.

Furthermore, the *ALPHA/BETA-HYDROLASE SUPERFAMILY PROTEIN (ABH),* identified as commonly differentially expressed in both studies, is a member of a family that constitutes the core structure of phytohormone receptors in the gibberellin and other phytohormone pathways (Mindrebo et al., 2016). Specifically, in rice, the ABH protein

GIBBERELLIN INSENSITIVE DWARF 1 (*GID1*) serves as the gibberellin receptor, orchestrating GA signaling through GID1-mediated degradation of DELLA proteins (Ueguchi-Tanaka et al., 2005). As discussed above, GA signaling holds potential significance in governing development rate.

CONCLUSION

In summary, our discussion has shed light on the potential roles of auxin polar transport, gibberellin signaling, and MATE efflux transporters in regulating development rate in petunia. Furthermore, we delved into the pleiotropic effects of multiple gene families, indicating the necessity for additional studies to functionally characterize these gene families involved in both vegetative and reproductive phases. Additionally, we examined the shared genetic factors between IA and AE differentially expressed genes (DEGs), emphasizing their significance for further investigation. Such endeavors are crucial for achieving a comprehensive understanding of the genetic factors that influence development rate.

Tables & Figures

Table 3-1: Mean and standard deviation of leaf number on the side branches over a four-week interval of the total twenty-five AE RILs phenotyped for development rate. N represents number of plants on which data was collected for each genotype. Genotypes marked bold were selected for shoot apex tissue collection and RNA extraction under slow and fast development rate categories.

Genotype	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem
	Rep 1		Rep 2	
17	10.67 ± 3.1 (3)	7.33 ± 0.6 (3)	14 ± 0 (1)	9.00 ± 1.4 (2)
26	12.14 ± 2.4 (7)	8.13 ± 0.8 (8)	12.30 ± 2.8 (10)	6.92 ± 2.3 (12)
39	11.30 ± 3.0 (10)	8.92 ± 2.4 (13)	13.50 ± 2.1 (8)	9.45 ± 2.0 (11)
45	13.58 ± 4.2 (12)	11.80 ± 2.4 (10)	14.17 ± 2.9 (12)	12.46 ± 1.7 (13)
49	12.91 ± 1.9 (11)	10.56 ± 2.6 (9)	13.78 ± 4.2 (9)	10.27 ± 2.0 (11)
61	12.00 ± 1.4 (4)	8.20 ± 1.8 (5)	13.75 ± 1.7 (8)	7.67 ± 1.5 (6)
85	16.14 ± 3.6 (7)	11.50 ± 1.0 (6)	14.25 ± 2.0 (8)	9.75 ± 1.3 (8)
87	13.44 ± 3.3 (9)	11.22 ± 4.7 (9)	10.82 ± 4.0 (11)	9.00 ± 2.7 (11)
110	13.29 ± 2.9 (14)	11.68 ± 2.1 (19)	12.80 ± 3.3 (10)	12.18 ± 2.1 (17)
116	12.00 ± 0.0 (1)	8.00 ± 0.0 (1)	12.75 ± 3.6 (4)	10.67 ± 2.3 (3)
126	6.00 ± 0.0 (1)	10.00 ± 2.8 (2)	11.67 ± 2.1 (3)	10.60 ± 2.6 (5)
157	13.00 ± 1.5 (6)	7.00 ± 2.6 (4)	12.29 ± 2.4 (7)	9.33 ± 2.1 (6)
173	17.33 ± 3.1 (3)	8.33 ± 0.6 (3)	8.00 ± 0.0 (1)	10.38 ± 2.7 (8)
193	12.38 ± 1.8 (8)	9.43 ± 1.1 (7)	13.67 ± 2.1 (9)	7.45 ± 1.8 (11)
199	14.33 ± 3.2 (3)	9.50 ± 2.1 (2)	16.00 ± 0 (2)	9.67 ± 2.5 (3)

Table 3-1 (cont'd)

Genotype	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem
	Rep 1		Rep2	
208	11.13 ± 2.3 (8)	9.00 ± 1.0 (4)	9.75 ± 2.0 (8)	7.43 ± 1.5 (7)
216	12.71 ± 1.9 (7)	9.75 ± 1.5 (8)	10.63 ± 1.6 (8)	8.67 ± 2.0 (9)
219	11.66 ± 3.5 (3)	9.00 ± 1.4 (2)	11.50 ± 1.0 (4)	8.80 ± 1.1 (5)
252	11.44 ± 0.9 (9)	10.56 ± 1.6 (9)	10.00 ± 1.4 (5)	10.89± 2.4 (9)
259	14.00 ± 2.3 (8)	9.00 ± 2.0 (4)	11.00 ± 3.0 (12)	8.00 ± 2.0 (12)
279	9.10 ± 2.7 (10)	9.18 ± 1.6 (11)	11.17 ± 2.3 (12)	9.93 ± 1.9 (14)
298	13.75 ± 0.5 (4)	9.50 ± 1.9 (4)	11.00 ± 1.4 (2)	9.83 ± 2.0 (6)
318	11.60 ± 2.2 (10)	8.83 ± 2.1 (12)	11.11 ± 1.1 (9)	9.09 ± 1.0 (11)
319	13.14 ± 3.9 (7)	13.00 ± 4.9 (9)	14.25 ± 1.3 (4)	13.00 ± 2.6 (8)
321	12.11 ± 1.4 (9)	9.22 ± 1.5 (9)	12.00 ± 2.2 (9)	10.50 ± 1.9 (12)

Table 3-2: Summary of twenty RNA samples and their biological replicates including the RNA integrity number (RIN), lane information, number of raw reads generated, number of reads after merging the two lanes, final number of reads that survived trimming, percentage of reads uniquely mapped to the *P. axillaris* genome and percentage of reads mapped to the exon regions.

RIL	Biological Replicate	RIN score	Lane	Number of raw reads	Final number of reads after trimming	Number of merged reads	Percent of uniquely mapped reads to the genome	Percent of reads mapped to the exon
AE318s	1	7.1	1	17,342,626	17296141	34853295	89.7	79
			2	17,600,922	17557154			
	2	6.9	1	15,737,006	15693363	31781844	87.7	80
			2	16,129,461	16088481			
AE252s	1	6.6	1	19,555,233	19503720	39394483	86.0	80
			2	19,940,218	19890763			
	2	5.6	1	12,554,115	12520592	25358743	83.9	81
			2	12,870,603	12838151			
AE219s	1	7.5	1	15,501,065	15456973	31177925	90.5	80
			2	15,763,098	15720952			
	2	7.2	1	16,114,232	16069049	32542079	88.2	80
			2	16,516,377	16473030			
AE216s	1	7.7	1	20,645,638	20589479	41610734	91.2	74
			2	21,074,545	21021255			
	2	7.1	1	15,870,623	15827345	32004906	89.1	78
			2	16,219,001	16177561			
AE208s	1	6.8	1	19,501,537	19448910	39194271	88.9	79
			2	19,794,733	19745361			
	2	7.1	1	19,967,546	19914789	40398108	85.5	80

Table 3-2 (cont'd)

RIL	Biological Replicate	RIN score	Lane	Number of raw reads	Final number of reads after trimming	Number of merged reads	Percent of uniquely mapped reads to the genome	Percent of reads mapped to the exon
			2	20,533,515	20483319	40398108	85.5	80
AE319f	1	6.9	1	16,204,948	16159045	32762137	89.6	78
			2	16,647,313	16603092			
	2	7.3	1	16,627,647	16582737	33549819	87.7	79
			2	17,009,885	16967082			
			2	14,954,491	14914420			
AE173f	1	6.9	1	14,628,426	14586530	29500950	90.3	80
			2	14,954,491	14914420			
	2	6.5	1	17,150,853	17104292	34641052	88.5	80
			2	17,581,035	17536760			
AE110f	1	6.8	1	16,568,767	16520738	33394712	86.6	82
			2	16,919,750	16873974			
	2	7.0	1	19,528,317	19474938	39497021	87.4	81
			2	20,072,463	20022083			
AE85f	1	7.3	1	19,496,034	19440662	39260523	89.6	81
			2	19,872,300	19819861			
	2	6.5	1	18,986,587	18933799	38265044	90.1	80
			2	19,381,081	19331245			

Table 3-2 (cont'd)

RIL	Biological Replicate	RIN score	Lane	Number of raw reads	Final number of reads after trimming	Number of merged reads	Percent of uniquely mapped reads to the genome	Percent of reads mapped to the exon
AE45f	1	7.4	1	23,295,629	23230884	46860056	89.4	80
			2	23,689,962	23629172			
	2	7.3	1	19,409,346	19354918	39194168	90.3	80
			2	19,891,452	19839250			

Figure 3-1: Principal component analysis conducted based on normalized gene expression count of samples. Round circle represents three red- colored samples (two biological replicates of each sample) from the fast and oval circle represents three blue-colored samples (at least one biological replicate per sample) from slow development rate categories, respectively. The x-axis represents the PC1 and the percentage of variance explained and y-axis represents PC2 and the percentage of variance explained.

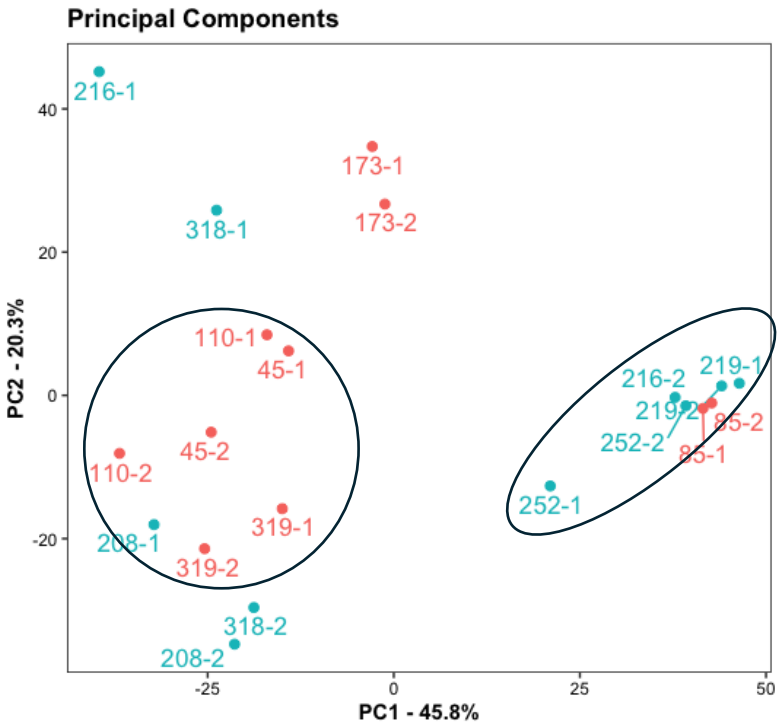
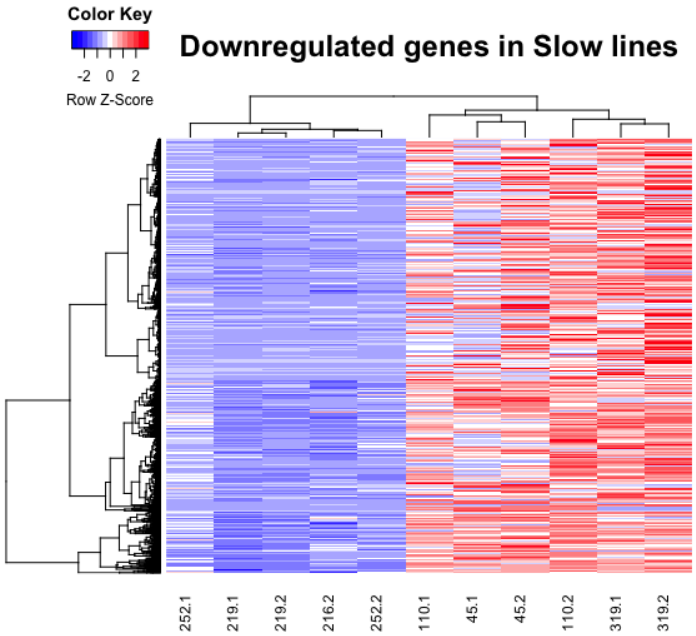


Figure 3-2: Heatmaps of differentially expressed genes in slow lines as compared to the fast lines. 1834 downregulated genes in the pooled comparisons (A), 210 upregulated genes in the pooled comparisons (B), and 271 core downregulated genes in the pairwise comparisons (C).

A)



B)

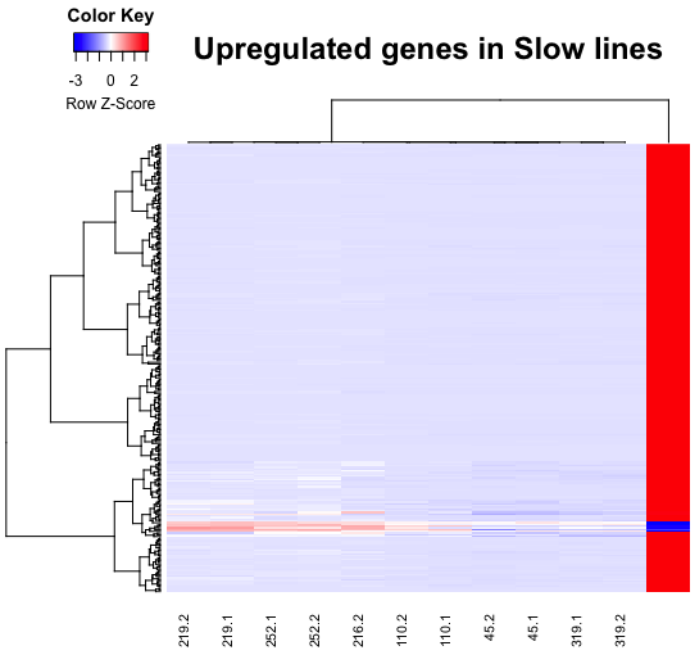


Figure 3-2 (cont'd)

C)

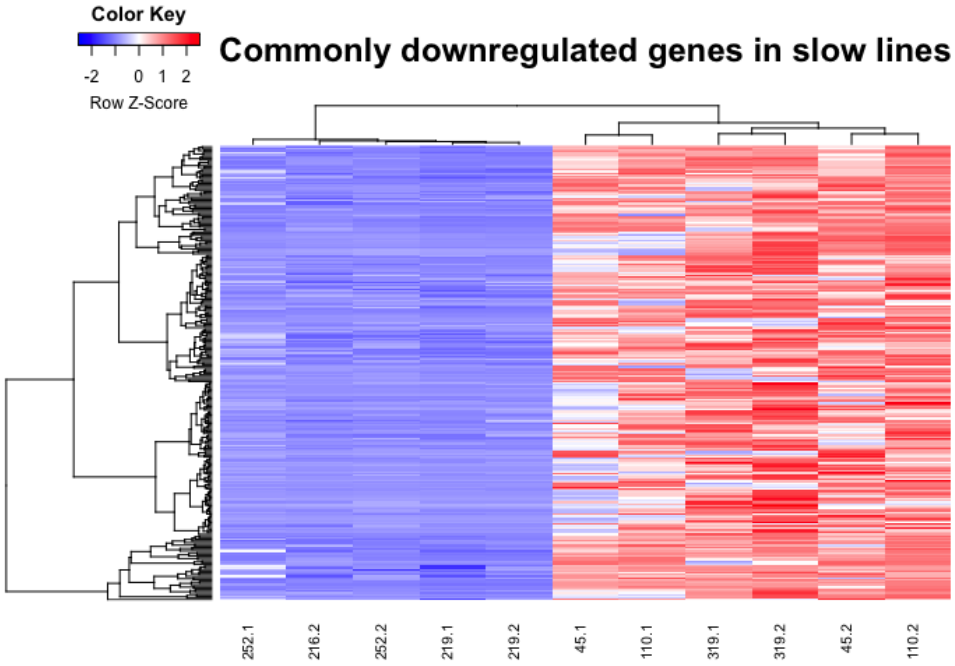


Table 3-3: Pairwise comparisons of differentially expressed genes. D and U indicates number of down-regulated and up-regulated genes, respectively, in each of the comparisons.

	AE319f		AE110f		AE45f	
	D	U	D	U	D	U
AE252s	1298 (66%)	672 (34%)	1170 D (70%)	495 (30%)	1349 (68%)	634 (32%)
AE219s	2442 (80%)	628 (20%)	2271 (82%),	493 (18%)	2505 (79%)	662 (21%)
AE216s	2333 (73%)	855 (27%)	1965 (82%),	446 (18%)	2262 (74%)	788 (26%)

Figure 3-3: Venn diagrams representing individual pairwise comparisons of differentially expressed genes. First three Venn diagrams are comparisons of down-regulated (A) and up-regulated genes (B) between each slow line with all three fast lines and the fourth diagram draws comparisons between results of first three comparisons.

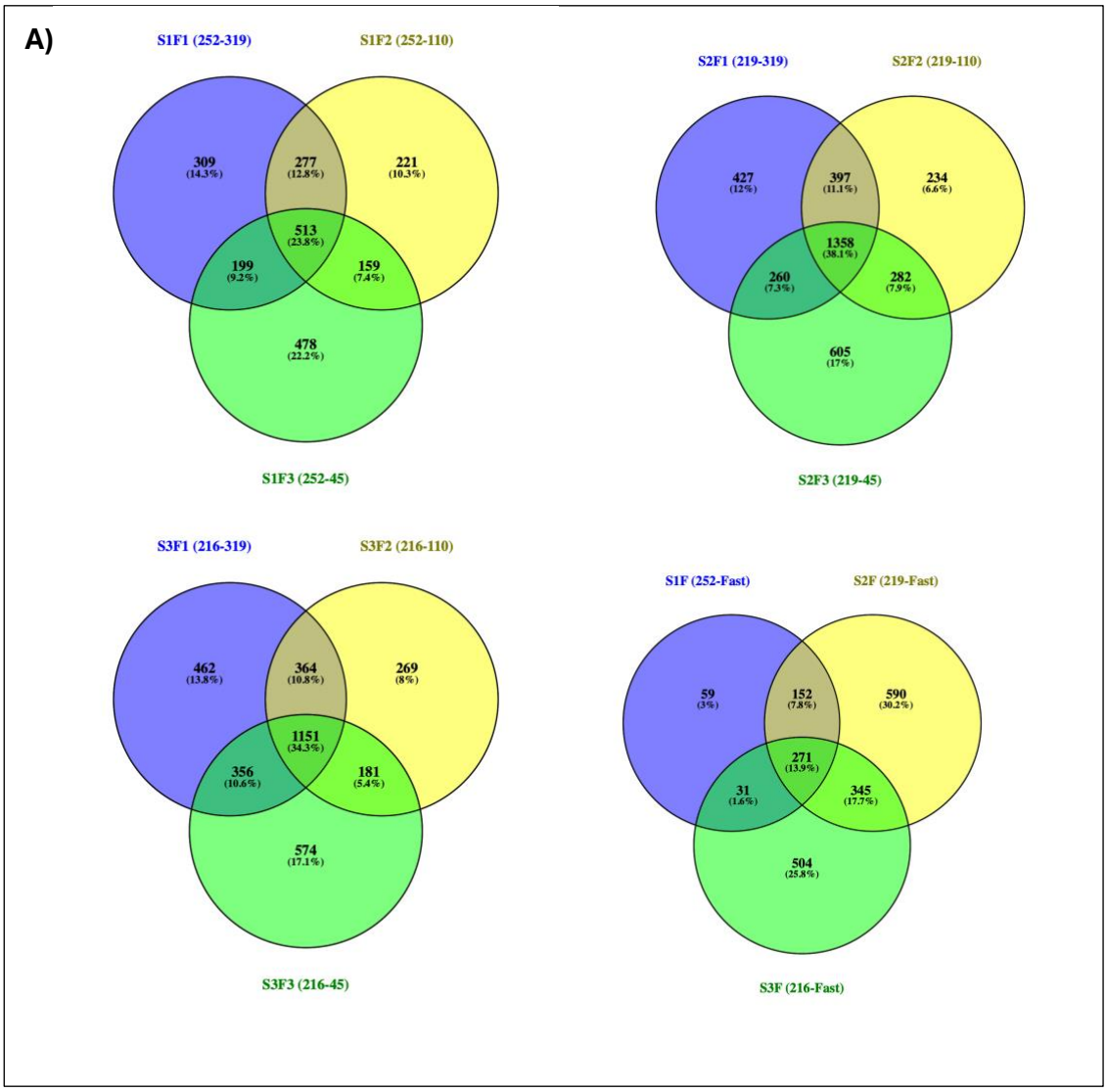


Figure 3-3 (cont'd)

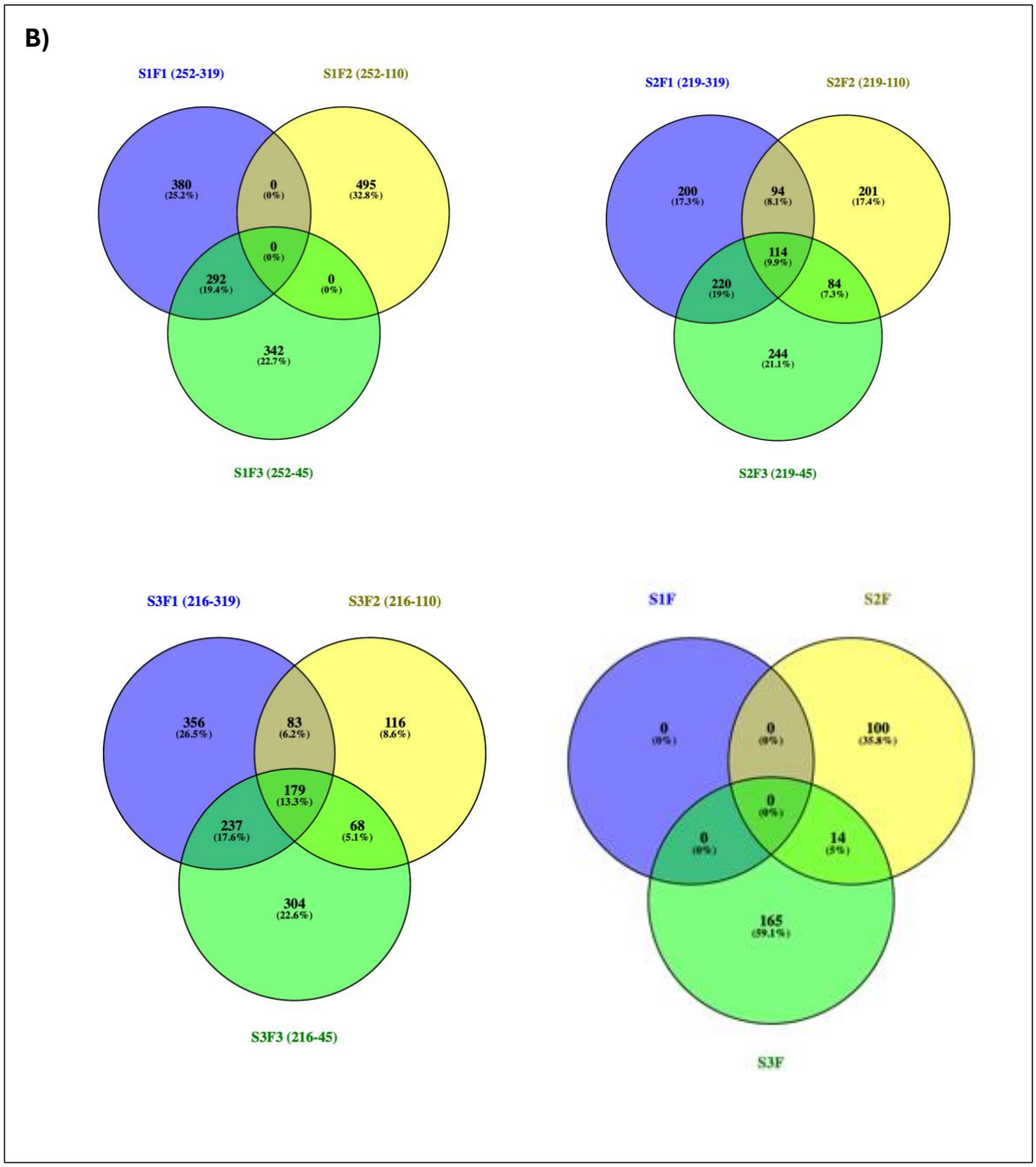


Figure 3-4: Sample clustering dendrogram of all samples (A) and only 11 samples used for the WGCNA analysis (B).

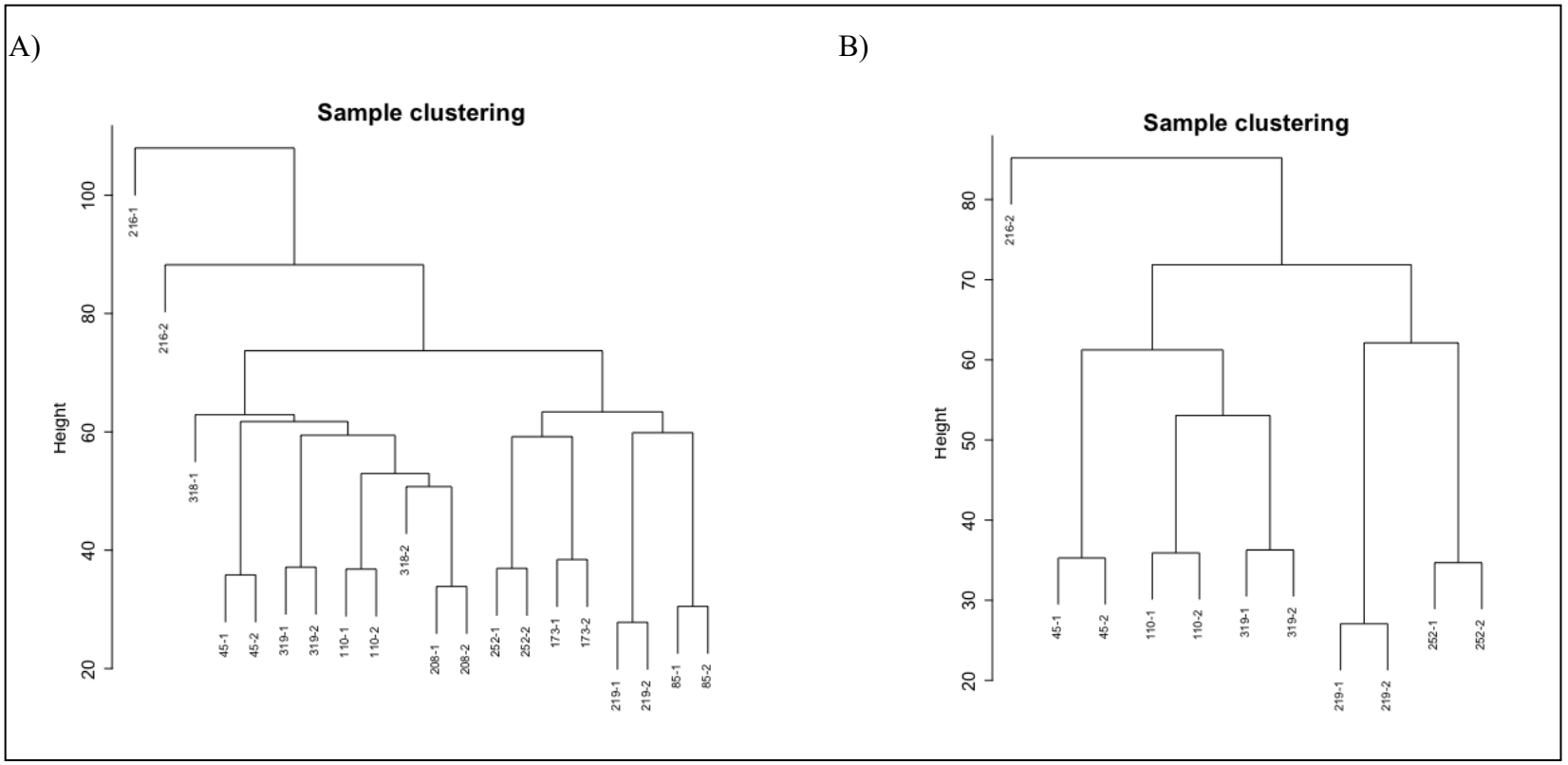


Figure 3-5:Heatmap displaying the significantly correlated modules with the development rate phenotype. Numbers inside each box indicate Pearson correlation coefficient between the module and the phenotype and a p-value in bracket. Red color indicates positive correlation whereas blue color indicates a negative correlation.

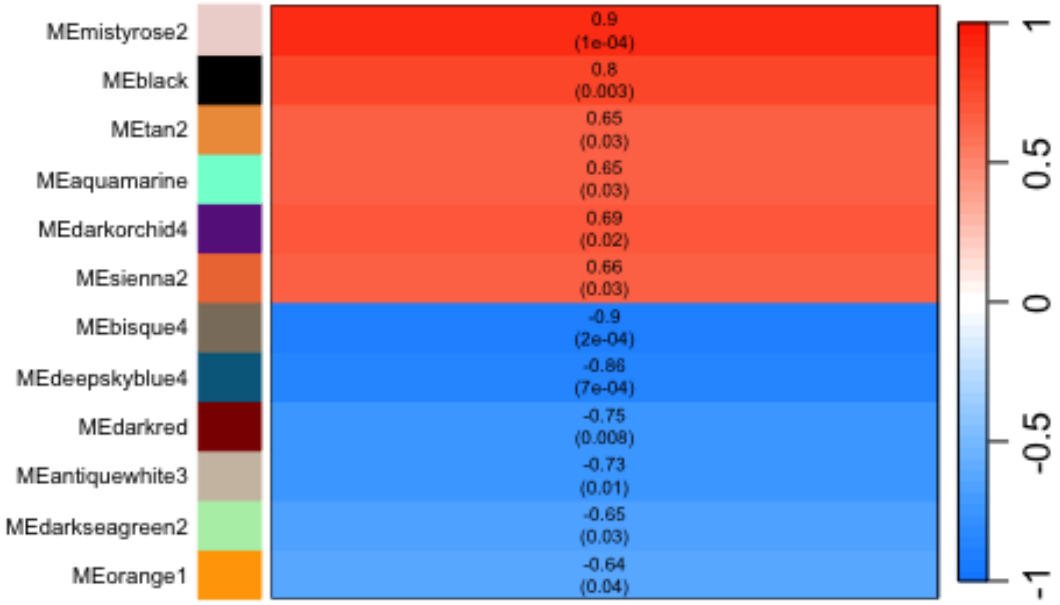


Table 3-4: Summary of significantly correlated modules. Hub genes are defined as genes with module membership & gene significance values greater than equal to 0.8.

Module	Number of total genes	Number of hub genes
Mistyrose2	189	75
Black	847	105
Tan2	54	2
Aquamarine	2429	45
Darkorchid	565	24
Sienna2	967	58
Bisque4	1438	310
Deepskyblue4	326	63
Darkred	689	63
Antiquewhite3	139	10
Darkseagreen2	72	2
Orange1	301	11

Figure 3-6: Bar plots representing eigenvalues of the negatively (A) and positively correlated modules (B) in each of the samples. Phenotypes 1-3 are the fast lines samples; 1- AE319f, 2- AE110f and 3- AE45f and phenotypes 4-6 are the slow line samples; 4- AE252s, 5- AE219s, and 6- AE216s. Eigengene value represents the average expression profile of all genes within the module for each sample, serving as a representative expression value for the entire module.

A)

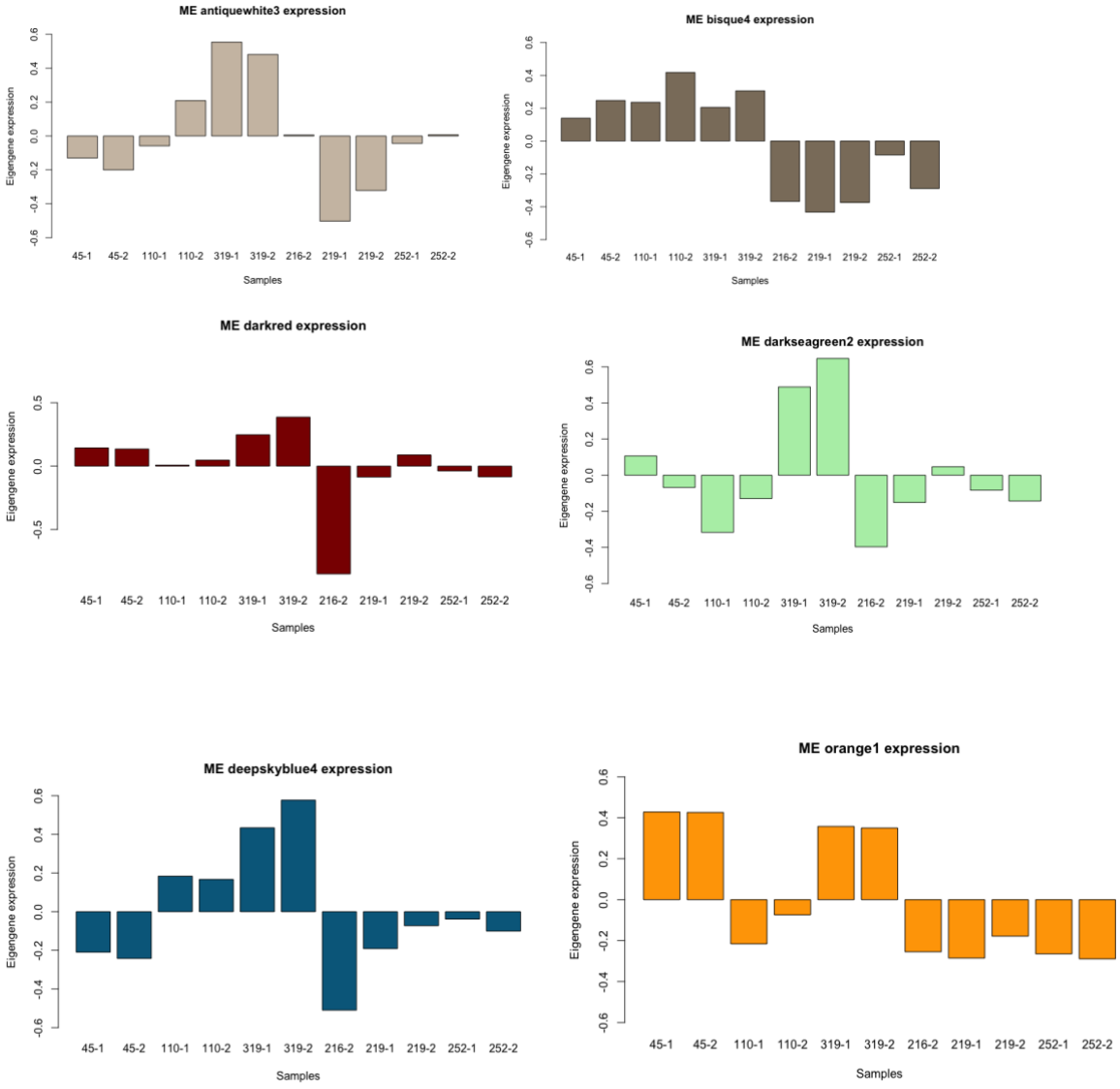


Figure 3-6 (cont'd)

B)

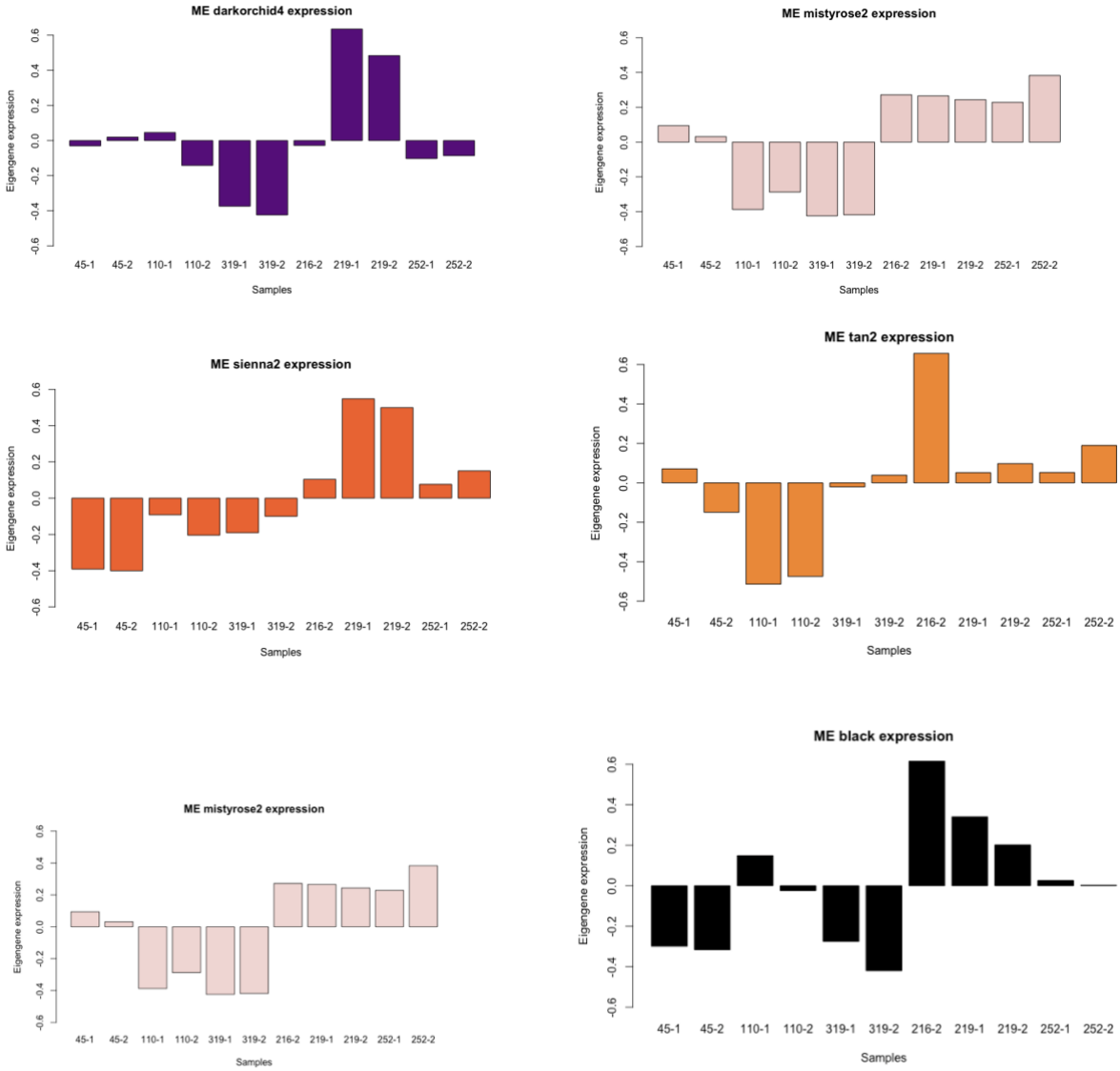


Table 3-5: Functional description of genes common in modules identified in WGCNA and core DEGs. Core DEGs are identified as genes commonly differentially expressed in at least six pairwise comparisons of slow and fast lines.

Module	GeneID	Functional description
bisque4	Peaxi162Scf00001g00481.1	SPX (SYG1/Pho81/XPR1) domain-containing protein
	Peaxi162Scf00002g00332.1	calcium ATPase 2
	Peaxi162Scf00003g02039.1	conserved hypothetical protein [Ricinus communis] gb|EEF43357.1| conserved hypothetical protein [Ricinus communis]
	Peaxi162Scf00003g05227.1	nodulin MtN21 /EamA-like transporter family protein
	Peaxi162Scf00011g00077.1	SRF-type transcription factor family protein [Solanum lycopersicum]
	Peaxi162Scf00016g02234.1	cysteine proteinase1
	Peaxi162Scf00022g00098.1	**MADS-box transcription factor 3**
	Peaxi162Scf00037g01116.1	40S ribosomal protein S10-3
	Peaxi162Scf00045g00142.1	NAD(P)-binding Rossmann-fold superfamily protein
	Peaxi162Scf00069g01326.1	Unknown protein
	Peaxi162Scf00073g00173.1	Unknown protein
	Peaxi162Scf00075g01418.1	**basic helix-loop-helix (bHLH) DNA-binding superfamily protein**
	Peaxi162Scf00079g00129.1	glucose-6-phosphate/phosphate translocator 2
	Peaxi162Scf00083g00516.1	**Protein WUSCHEL**
	Peaxi162Scf00089g01235.1	Quinolinate synthase, chloroplastic
	Peaxi162Scf00089g01857.1	UDP-N-acetylglucosamine--N-acetylmuramyl- pyrophosphoryl-undecaprenol N-acetylglucosamine transferase isoform 1
	Peaxi162Scf00111g00125.1	NADH-ubiquinone oxidoreductase chain 5
	Peaxi162Scf00128g01233.1	HORMA domain-containing protein 1
	Peaxi162Scf00140g00212.1	Mannan endo-1,4-beta-mannosidase 7
	Peaxi162Scf00152g00612.1	squalene monooxygenase 2
	Peaxi162Scf00152g01220.1	actin-11
	Peaxi162Scf00155g00096.1	cationic amino acid transporter 5
	Peaxi162Scf00160g00117.1	RING/U-box superfamily protein
	Peaxi162Scf00171g00044.1	Ornithine decarboxylase
	Peaxi162Scf00174g00101.1	phospholipase D alpha 1
	Peaxi162Scf00198g00148.1	MLP-like protein 28
	Peaxi162Scf00199g01019.1	Pectin lyase-like superfamily protein
	Peaxi162Scf00253g01317.1	Unknown protein
	Peaxi162Scf00253g01317.1	Unknown protein

Table 3-5 (cont'd)

Module	GeneID	Functional description
	Peaxi162Scf00270g00089.1	Unknown protein
	Peaxi162Scf00284g00022.1	Peroxidase superfamily protein
	Peaxi162Scf00288g00815.1	Galactosyltransferase family protein
	Peaxi162Scf00349g00711.1	calmodulin-binding family protein
	Peaxi162Scf00406g00134.1	Glutelin type-A 1 [Morus notabilis]
	Peaxi162Scf00409g00516.1	RmlC-like cupins superfamily protein
	Peaxi162Scf00434g00334.1	GDSL esterase/lipase
	Peaxi162Scf00451g00723.1	Transmembrane amino acid transporter family protein
	Peaxi162Scf00481g00063.1	Zinc transporter 2
	Peaxi162Scf00486g00310.1	alpha/beta-Hydrolases superfamily protein
	Peaxi162Scf00516g00671.1	NAD(P)-binding Rossmann-fold superfamily protein
	Peaxi162Scf00560g00223.1	Aldehyde dehydrogenase family 2 member C4
	Peaxi162Scf00570g00049.1	PIG93, partial [Petunia x hybrida]
	Peaxi162Scf00619g00113.1	Protein kinase superfamily protein
	Peaxi162Scf00620g00814.1	Glucose-methanol-choline (GMC) oxidoreductase family protein
	Peaxi162Scf00623g00041.1	*O*-fucosyltransferase family protein
	Peaxi162Scf00717g00215.1	Calcium-dependent phosphotriesterase superfamily protein
	Peaxi162Scf00739g00426.1	response regulator 17
	Peaxi162Scf00907g00120.1	Major facilitator superfamily protein
	Peaxi162Scf00931g00117.1	DNA binding protein, putative [Ricinus communis] gb|EEF52579.1| DNA binding protein, putative [Ricinus communis]
	Peaxi162Scf00943g00003.1	conserved hypothetical protein [Ricinus communis] gb|EEF30394.1| conserved hypothetical protein [Ricinus communis]
	Peaxi162Scf00945g00013.1	**Auxin-responsive GH3 family protein**
	Peaxi162Scf01002g00114.1	Pollen Ole e 1 allergen and extensin family protein [Theobroma cacao] gb|EOY03810.1| Pollen Ole e 1 allergen and extensin family protein [Theobroma cacao]
	Peaxi162Scf01694g00018.1	K(+)-insensitive pyrophosphate-energized proton pump
	Peaxi162Scf02085g00005.1	DNAse I-like superfamily protein
Sienna2	Peaxi162Scf00129g01043.1	**Homeobox-leucine zipper protein HAT5**
	Peaxi162Scf00140g01136.1	beta glucosidase 11
	Peaxi162Scf00351g00526.1	Unknown protein
	Peaxi162Scf00355g00111.1	Lupus la ribonucleoprotein, putative isoform 2 [Theobroma cacao] gb|EOY34314.1|

Table 3-5 (cont'd)

Module	GeneID	Functional description
	Peaxi162Scf00818g00014.1	Disease resistance protein (CC-NBS-LRR class) family
	Peaxi162Scf01039g00236.1	Protein kinase superfamily protein
Orange1	Peaxi162Scf00129g00543.1	thioredoxin 2
	Peaxi162Scf00169g00182.1	arogenate dehydrogenase
	Peaxi162Scf00287g01112.1	Plasma membrane ATPase 3

Table 3-6: Summary of genes commonly differentially expressed between AE and IA RILs.

Gene	Functional description
Peaxi162Scf00049g01720.1	alpha/beta-Hydrolases superfamily protein
Peaxi162Scf00306g00022.1	Protein of unknown function, DUF584
Peaxi162Scf00118g00042.1	Pectinacetylesterase family protein
Peaxi162Scf00204g01613.1	L-ascorbate oxidase
Peaxi162Scf00204g00116.1	LAG1 longevity assurance homolog 3
Peaxi162Scf00549g00015.1	HXXXD-type acyl-transferase family protein
Peaxi162Scf00241g00053.1	Plant invertase/pectin methylesterase inhibitor superfamily protein
Peaxi162Scf00002g00191.1	heat shock protein 21
Peaxi162Scf00734g00066.1	"Protein kinase family protein"
Peaxi162Scf00666g00042.1	“Protein LIGHT-DEPENDENT SHORT HYPOCOTYLS 10”

CHAPTER 4

IDENTIFICATION OF QUANTITATIVE TRAIT LOCI (QTL) RELATED TO STEVIA DEVELOPMENT RATE AND OTHER LEAF YIELD RELATED TRAITS

INTRODUCTION

Stevia rebaudiana, commonly known as stevia (2n=22), is an important medicinal perennial plant belonging to the Asteraceae family (Goyal et al., 2010). Native to northeast Paraguay (Shock, 1982; Ramesh et al., 2006), stevia leaves produce a group of zero-glycemic, low-calorie sweet-tasting compounds called steviol glycosides (Brandle and Telmer, 2007; Ceunen and Geuns, 2013). These steviol glycosides (SGs) are extracted from the leaves, which can contain up to 30% of these compounds on a dry mass basis (Goyal et al., 2010; Yadav and Guleria, 2012; Ceunen and Geuns, 2013), and are 200-300 times sweeter than sucrose.

In Japan, steviol glycosides have been used as a sweetener in seafoods, soft drinks, and candies since the 1970s (Mizutani and Tanaka, 2001). Beyond their use as sweeteners, stevia has been employed as a weight control agent in obese individuals (Gupta et al., 2013) and as a natural treatment for diabetes in various parts of the world (Shivanna et al., 2013). Stevia products appeal to consumers seeking natural ingredients in their diet. Due to their plant-based origin, steviol glycosides hold great potential as alternatives to sugar and synthetic sweeteners. To meet the growing demand for stevia, it is crucial for plant breeders to develop high SG-yielding cultivars. This will ensure a consistent supply of these beneficial compounds, supporting both consumer health and the food industry's need for natural sweetening agents.

Improvement of stevia through traditional breeding approaches is hampered by its self-incompatibility and low seed germination rate (Yadav et al., 2014; Ucar et al., 2016; Attaya, 2017; Simlat et al., 2018). As an alternative, stevia is clonally propagated by stem cuttings and in-vitro methods to produce genetically uniform plant populations (Goettemoeller and Ching, 1999; Ramesh et al., 2006; Smitha and Umesha, 2012). The above-ground tissue of stevia is harvested, and the leaves are stripped off for the extraction of steviol glycosides. A key strategy

to increase the yield of steviol glycosides is to enhance the rate of leaf production over time, enabling multiple harvests per season. Therefore, understanding the genetic architecture underlying stevia leaf production rate is crucial for breeding high-yielding cultivars. This knowledge will facilitate the development of stevia varieties that can produce more leaves and, consequently, higher amounts of steviol glycosides, meeting the increasing demand for this natural sweetener.

Since stevia is a relatively novel crop, genetic research has predominantly focused on studying the biosynthesis of steviol glycosides (SGs), leaving a significant knowledge gap regarding the genetic mechanisms underlying a broader range of traits related to biomass production. Stevia leaf yield, which corresponds to overall biomass, depends on various morphological traits including leaf size (length and width), rate of leaf production (number of new nodes/leaves produced over time), branch production (primary and secondary branches), maximum and minimum plant canopy width, flowering time, and overall plant vigor. Increased biomass production in stevia would help meet the growing consumer demand for natural sugar alternatives. Thus, investigating the genetic regulation of traits related to biomass production is essential for advancing stevia breeding programs (Hastoy et al., 2019). However, stevia research at the genetic level is constrained by several factors, including the limited availability of germplasm, molecular markers, and a high-resolution linkage map. Overcoming these limitations is crucial for the development of high-yielding stevia cultivars.

The initial genetic linkage map for *Stevia rebaudiana* was constructed using random amplified polymorphic DNA (RAPD) markers (Yao et al., 1999). Due to the limited efficiency of RAPD markers, we developed an improved linkage map using simple sequence repeat (SSR) markers, leveraging RNA sequencing data from young fully expanded leaves, shoot apices,

flowers, and callus tissues (Vallejo and Warner, 2021). This enhanced map enabled the identification of the first quantitative trait loci (QTL) for steviol glycosides, such as Reb D and Reb A, as well as for plant height and vigor, based on phenotypic data from an F_1 mapping population of 161 individuals (Vallejo and Warner, 2021). However, for more precise QTL mapping, a high-density linkage map comprising 11 linkage groups corresponding to the 11 chromosomes of stevia is essential. Genotyping-by-sequencing (GBS)-based single nucleotide polymorphism (SNP) markers are considered optimal for developing such a linkage map (Kho et al., 2021). These markers can be detected in large quantities through automated processes, are relatively abundant, and offer greater genetic stability compared to SSR markers (Delourme et al., 2013; Tsykun et al., 2017). With the availability of a chromosome-level genome assembly, it is now possible to compare the genetic positions of markers with their physical locations on the stevia chromosomes (Xu et al., 2021).

This study aims to overcome the limitations in stevia genetics research and deepen our comprehension of the genetic determinants governing biomass production in stevia. The primary objectives are to construct a high-density linkage map based on SNP markers, evaluate the performance of an F_1 mapping population concerning biomass-related traits across diverse environmental conditions and to pinpoint genomic regions (QTL) linked with these traits. The findings from this investigation will offer valuable insights into the genetic regulation of traits influencing leaf yield in stevia.

MATERIALS AND METHODS

Plant materials

A stevia F_1 mapping population (designated as MSU18-02) was established through a cross between two distinct lines from the MSU stevia program, namely 10-RJR and 10-19, which

exhibit variability in steviol glycoside production (Bahmani, 2021). Subsequently, individuals from this population were propagated clonally. The population, comprising 200 individuals, was planted in a randomized complete block design with three replications at two field trial locations in June 2020 and repeated in June 2021. These trial sites were the MSU Horticultural Teaching and Research Center (HTRC) in Holt, MI, and the MSU Southwest Michigan Research and Education Center (SWMREC) in Benton Harbor, MI. Planting was carried out using raised, plastic-covered rows with drip irrigation installed and utilized as required.

Data collection

At the time of planting, two consecutive newly fully expanded leaves were marked using white paint. Subsequently, after a ten-week period in the field, various morphological parameters were assessed. These included the maximum and minimum width of the plant (in cm), as well as the length and width at the widest point of a young, fully expanded leaf (mm). Additionally, the stem caliper (mm) was recorded at the point of the young, fully expanded leaf. The number of new nodes formed above the marked leaves and the phyllotaxy pattern were also documented: "Opposite" (O) when two leaves arose from a single node, "Alternate" (A) when leaves arose from individual nodes and “Hybrid” (H) when a plant displayed a mix of both O and A phyllotaxy pattern. Furthermore, the number of primary branches (> 3 cm) extending from the main stem was counted. Subjective indices were established to quantify secondary branching, flowering stage, and plant vigor. Secondary branching was rated on a scale ranging from 1 (low) to 5 (high), with 1 indicating no lateral shoots emanating from the primary branches and 5 representing strong secondary branching. Similarly, the flowering stage of the plant was assessed on a scale from 1 to 5, with 1 indicating the absence of visible flower buds and only vegetative growth, and 5 indicating a plant that has been flowering for some time with numerous open

flowers. Plant vigor was rated from 1 to 5 based on the overall volume of the plant canopy, with plants rated as 1 exhibiting minimal branching and consequently very low biomass, while those rated as 5 displayed strong primary and secondary branching, resulting in high overall biomass.

Data analysis

Descriptive statistics, population distributions, and Pearson correlation coefficients were computed and analyzed using SPSS version 27 (IBM; Chicago, IL). Broad-sense heritability (H2) was calculated as follows: $H^2 = \sigma^2(\text{genotype}) / [(\sigma^2 (\text{genotype}) + (\sigma^2 (\text{genotype: location})/2) + (\sigma^2 (\text{genotype: year})/2 + (\text{RESIDUAL}/3*2*2)]$. This calculation was performed using a linear mixed model (lmerMod) in a two-stage model approach, as described previously (Schmidt et al., 2019), in the R programming environment.

Genotyping and linkage map generation

DNA was extracted from leaf samples collected from 238 individuals of the MSU18-02 (F_1) population, and these samples were subsequently sent to the University of Minnesota Genomics Center (UMGC) for genotyping using the genotyping-by-sequencing (GBS) approach. The variant calling and genotyping processes were conducted following standard procedures outlined in the Genome Analysis ToolKit (GATK) software suite (DePristo et al., 2011). Prior to variant calling, GBS sequence reads underwent alignment to a reference assembly of the stevia genome (Xu et al., 2021)using the BWA-MEM alignment tool (Li, 2013). The reference genome comprised 6978 contigs, with 6358 of these contigs assembled into 11 chromosomes. The finalized chromosome-level genome encompassed 3708 scaffolds, boasting a scaffold N50 value of 106.55 Mb and a cumulative length of 1416 Mb. Following BWA-MEM alignment, reads were subjected to filtering to exclude alignments with a mapping quality (MAPQ) score of less

than 20 or those designated as non-primary alignments, thereby eliminating reads that mapped to multiple locations within the genome.

The resulting alignments were further processed using GATK, and variants were called utilizing its HaplotypeCaller utility. Notably, variant calling was restricted to the 11 chromosomes, despite the reads being mapped to the entire reference assembly, inclusive of unassembled contigs. Variants spanning all F_1 individuals and the parental lines were combined, and genotypes were assigned to each individual. Subsequently, variants were subjected to filtering, retaining those with a mean sequence depth ≥ 5, variant quality ≥ 50, genotype quality ≥ 10, and missing rate ≤ 0.25. Variants with a minor allele frequency (MAF) $< 1\%$ or those harboring more than two alleles were excluded from the dataset, yielding final genotype data for all 238 individuals. The genotypes, called in Variant Call Format (VCF), were converted to a format suitable for JoinMap, wherein alleles in the F_1 individuals were labeled as missing if they were absent in the parental lines or exhibited Mendelian inheritance errors. Variant sites with an allele missing rate > 0.01 underwent filtration, resulting in the generation of a final set of loci for constructing the linkage map in JoinMap5 (Ooijen, 2018). Markers displaying segregation distortion ($p \leq 0.05$) and redundant markers were eliminated from subsequent analysis. Linkage groups were delineated using the 'Independence LOD' function, and marker order was determined utilizing the maximum likelihood method. Map distances were calculated employing the Kosambi mapping function (Kosambi, 2016).

QTL identification

MapQTL6.0 (Ooijen, 2009) was used as the tool for identifying QTL positions in the genome linked to all phenotypic traits. QTL were discerned via the interval mapping method employing the regression algorithm with default settings. A genome wide significant LOD score

was established at the relative cumulative value of 0.95 (corresponding to a 95% probability level), determined through 1000 permutation tests. The proportion of total phenotypic variation explained by each QTL (VE%) was one of the outcomes of interval mapping. Subsequently, significant QTL, defined as those exhibiting overlapping regions across at least two environments on the same linkage group, were visualized using MapChart v2.32 (Voorrips, 2002).

RESULTS

SNP marker development

The initial number of raw reads (150 bp paired-end) spanned from 6 million to 36 million pairs across the samples, with a median of 10 million pairs. Post-mapping to the stevia reference genome and subsequent retention of uniquely mapped reads, the count of remaining reads (properly mapped reads) ranged from 6 million to 33 million, with a median of 10 million (individual reads were counted instead of read pairs, considering some reads lacked properly mapped mates) (Table 4-1). Our finalized genotype data for 238 individuals encompassed 181,614 variant sites, exhibiting an average genotyping rate of 0.74. Subsequently, 181,614 variants were filtered to eliminate markers with an allele missing rate of 0.01 or higher, resulting in a total of 11,575 SNPs for linkage map development.

Linkage map generation

A total of 1452 non-redundant markers were assigned to eleven linkage groups, corresponding to the eleven chromosomes. Upon comparing the genetic positions of these markers with their physical chromosome positions, 130 markers with conflicting positions were identified and subsequently removed (Figure 4-1). This curation yielded 1322 markers distributed across 11 linkage groups (Table 4-2), covering a cumulative distance of 2001.8 cM.

Although the average marker density across the map was 6.62 cM, it was notably inflated by linkage group 8, which comprised only 18 markers but spanned 1044.1 cM. This group exhibited the lowest marker density at 58.0 cM (Table 4-2). Notably, approximately 91% of the 841 markers mapped to chromosome 8 displayed segregation distortion and were consequently excluded from linkage group 8 analysis, resulting in only 18 markers on this linkage group in the final map. Excluding linkage group 8, the remaining ten linkage groups collectively spanned 1947.7 cM, with individual linkage groups ranging from 77.1 cM for linkage group 4 to 325.5 cM for linkage group 2 (Table 4-2). The number of SNPs varied from 67 on linkage group 4 to 209 on linkage group 2. The average marker density across the map was 1.48 cM, with individual linkage group marker densities, excluding linkage group 8, ranging from 1.15 cM on linkage group 4 to 1.84 cM on linkage group 7.

Phenotyping

The MSU18-02 population exhibited transgressive segregation across the majority of traits at both locations and in both years (HTRC and SWMREC 2020, and HTRC and SWMREC 2021), with the exception of flowering stage across all environments and secondary branching in 2021, which displayed skewed distributions (Figures 4-2, 4-3, 4-4, and 4-5). Generally, mean values for most traits were higher in 2020 compared to 2021 (Table 4-3). Additionally, the maximum values for minimum and maximum width were higher in 2020 than in 2021. Mean trait values of at least one parent (10-RJR) exceeded the population means for leaf width, primary branching, minimum width, and maximum width. Moreover, the maximum values of primary branching ranged from 1 to 6 in 2020. Across most traits, the mean values were higher for 10-RJR compared to 10-19 and the population means, while the mean of leaves was higher for 10-19 compared to 10-RJR.

Plants exhibited varying phyllotaxy patterns both within and between genotypes across all four locations. Phyllotaxy, the arrangement of leaves on the main stem, is typically categorized as alternate (one leaf per node) or opposite (two leaves per node) (Fleming, 2005; Lee et al., 2009). The percentage of plants displaying opposite phyllotaxy was the highest in all environments, followed by those with an alternate pattern. The occurrence of hybrid (mix of both opposite and alternate on a same plant) phyllotaxy was the least frequent (Table 4-4). In 2021, the percentage of plants with alternate and opposite phyllotaxy were comparable. However, in 2020, the proportion of plants with opposite phyllotaxy was 16% higher at HTRC and 23% higher at SWMREC compared to those with alternate phyllotaxy (Table 4-4).

Plant height displayed significant positive correlations with all traits, except for flowering stage, which did not exhibit a significant correlation with plant height (Table 4-5). Leaf length showed positive correlations with all other traits, except for flowering stage across both environments, and with primary and secondary branching in one of the locations in each year (SWMREC 2020 and HTRC 2021). Similarly, leaf width demonstrated positive correlations with all traits except primary branching and flowering stage in HTRC 2020 and SWMREC 2021. In SWMREC 2020 and HTRC 2021, leaf width exhibited positive correlations with leaves, plant vigor, and stem caliper (stem diameter), while it displayed a negative correlation with flowering stage. Primary branching exhibited positive correlations with secondary branching, minimum and maximum width, and vigor at both locations in 2020. Secondary branching was positively correlated with minimum and maximum canopy width and vigor across both environments and negatively correlated with leaves at both locations in 2020, but positively correlated at SWMREC 2021. Leaves showed a negative correlation with minimum width, maximum width, and plant vigor in HTRC 2020, while the correlation with these traits was positive in SWMREC

2021. Moreover, minimum canopy width demonstrated positive correlations with maximum width, vigor, stem caliper, and flowering stage in at least three environments. Additionally, maximum width displayed positive correlations with vigor and stem caliper across both environments, and vigor was positively correlated with stem caliper in both environments.

Broad-sense heritability estimates were predominantly high (> 0.5) for most traits, except for primary branching, stem caliper, and leaf count, which displayed heritability estimates of 0.33, 0.23, and 0.39, respectively (Table 4-6).

QTL identification

QTL analysis across four environments revealed the presence of at least one QTL for key traits such as maximum width, secondary branching, leaf length, vigor, and flowering stage (Table 4-7). Minimum canopy width and leaf width exhibited QTL in three environments, while leaves and stem caliper displayed QTL in two environments. Single environments showed QTL for plant height and primary branching. Significant QTL were identified for secondary branching on linkage groups 2, 7, and 11, accounting for 8.4% to 15.3% of phenotypic variation (VE%) (Figure 4-6). Another significant QTL was observed for minimum canopy width on linkage group 9, explaining 6.8% to 9.8% of VE%. Leaf width analysis revealed three distinct QTL positions (7.1, 10.6, and 10.4 cM) on linkage group 2 across various environments. Plant vigor analysis identified significant QTL on linkage groups 7 (explaining 10.7% to 11.6% VE%) and 11 (explaining 7.8% to 9.2% VE%). Furthermore, QTL for flowering stage were detected on linkage groups 3, 5, and 8. Notably, a QTL hotspot spanning the genetic interval of 160.90–161.90 cM on linkage group 7 was associated with multiple traits, including secondary branching, minimum width, maximum width, and plant vigor. For leaf count, no significant QTL

were detected except three single QTL at linkage groups 1, 2 and 4 which were identified at both sites in 2021.

DISCUSSION

The breeding of stevia for higher leaf yield is imperative to enhance the production of desirable steviol glycosides, which vary in composition (Ahmad et al., 2020). To achieve this goal, a high-density genetic linkage map is essential for precisely mapping these traits and facilitating marker-assisted selection breeding strategies. Previous efforts utilized linkage maps based on random amplified polymorphic DNA (RAPD) and inter simple sequence repeat (ISSR) markers to assess genetic diversity in stevia germplasm panels (Yao et al., 1999; Heikal et al., 2008; Chester et al., 2013). However, due to their limited reproducibility, these markers were found to be suboptimal. A more recent linkage map based on co-dominant simple sequence repeat (SSR) markers covered a distance of 582 cM across 13 linkage groups (Vallejo and Warner, 2021). Nonetheless, the efficiency and small marker numbers of SSR markers pose constraints on constructing a high-resolution genetic map (Kho et al., 2021). The advent of next-generation sequencing technologies has revolutionized marker development, particularly SNPs, offering a promising avenue for advancing stevia breeding efforts (Tam et al., 2019).

Here, we present the novel stevia linkage map constructed from 1322 SNP markers condensed into 11 linkage groups, corresponding to the 11 chromosomes of the stevia genome. Notably, this map exhibits a markedly higher average marker density compared to previously published maps, with an average spacing of 1.48 cM, in contrast to 6.0 cM and 7.6 cM in earlier studies (Yao et al., 1999; Vallejo and Warner, 2021). Despite this improvement, linkage group 8 stands out for its low marker density, considerable gaps, and a markedly non-linear marker order relative to physical chromosomal positions (refer to Figure 4-1). The presence of these large gaps

could not be rectified by reintegrating segregating distorted markers onto this group. The underlying causes for these gaps remain elusive; they may stem from genotyping errors or insufficient marker segregation between the parental lines within this linkage group. Further investigations are warranted to elucidate the precise factors contributing to this phenomenon. One potential avenue for enhancing the quality of this linkage group could involve the development of an integrated linkage map through the incorporation of SNP and SSR markers previously established.

Stevia ideotype breeding primarily aims to enhance the yield of steviol glycosides (SGs) and dry leaf yield, while also optimizing plant architecture for mechanized harvest, improving regrowth after winter, and enhancing tolerance to pathogens like Septoria leaf spot and weeds (Angelini et al., 2018; Tavarini et al., 2018; Hastoy et al., 2019; Huber and Wehner, 2023). Leaf yield in stevia is influenced by various factors including leaf size, leaf number, branching pattern, canopy dimensions, and leaf-to-stem ratios (Benhmimou et al., 2017; Abdulameer et al., 2018). Beyond leaf yield, the branching pattern and canopy size also play pivotal roles in determining plant architecture and overall vigor, which in turn affect the plant's competitive ability against weeds.

The MSU18-02 F_1 population exhibited a normal distribution across most of the leaf yield traits, implying polygenic control with varying genetic effects. Given its biparental nature, uncovering QTL associated with these traits would significantly augment our understanding of their genetic regulation. Through the identification of SNP markers linked to these QTL regions, marker-assisted selection breeding strategies can be employed to enhance desirable traits in stevia cultivars (Al-Taweel et al., 2021). Moreover, these QTL regions serve as valuable genomic

regions for identifying candidate genes underlying these traits, further advancing our knowledge of stevia genetics and facilitating targeted breeding efforts (Yang et al., 2021).

Overall, a positive correlation was noted among the leaf yield-related traits examined in this study, although some variations in correlation were observed across different environments. Notably, certain traits such as showed positive correlations in one environment but lacked correlation in others. Additionally, there were instances where the correlation between leaf count and other traits reversed across different environments. Nevertheless, a more extensive F_1 population and conducting trials across multiple environments could provide a more comprehensive understanding of the interplay between these traits.

The heritability estimates for most traits in our study ranged from moderate to high, indicating that genetic factors play a significant role in determining the variation observed in these traits. However, it's noteworthy that the heritability estimates for stem caliper were lower in our study compared to a previous investigation where it was reported as 0.75 (Vallejo and Warner, 2021). Similarly, the heritability estimates for traits like plant height, secondary branching, plant vigor, and leaf area were also relatively lower compared to the previous study (Vallejo and Warner, 2021). Conducting further investigations to refine the heritability estimates for these morphological traits would provide valuable insights into their true genetic potential and aid in optimizing breeding strategies (Huber and Wehner, 2023).

Leaf shape and size, encompassing parameters such as length, width, and angle, play pivotal roles in determining key physiological processes such as photosynthetic rate and canopy architecture, ultimately influencing overall plant biomass (Khuluq et al., 2022; Zhou et al., 2024). Consequently, the selection for optimal leaf size is paramount in breeding endeavors aimed at developing high-yielding varieties. However, it is equally crucial to investigate the

correlation between leaf size and steviol glycosides (SGs) content to inform strategic breeding decisions.

Stevioside (ST) and rebuadioside (Reb) A are among the most prevalent types of SGs utilized as sugar substitutes, albeit accompanied by a bitter aftertaste (Gupta et al., 2013). In contrast, Reb D and Reb M, although present in lower concentrations compared to ST and Reb A, offer a similar sweetness profile to Reb A while mitigating bitterness and enhancing taste (Prakash et al., 2014; Vallejo and Warner, 2021). The MSU18-02 population was also phenotyped for various SGs, including Reb A, D, and M, as well as ST and total SGs content across four distinct environments (Bahmani, 2021; Warner et al., unpublished). Analyzing trait correlations revealed intriguing insights, particularly regarding the relationship between leaf size and SGs content. Notably, leaf dimensions, including length and width, exhibited a negative correlation ($p < 0.01$) with Reb M in two of the four environments (HTRC 2020 and SWMREC 2020) (Table 4-8). Similarly, Reb D displayed a negative correlation with leaf size parameters in these environments, although statistical significance was not observed. The absence of correlation between leaf size and both Reb A and ST suggests that other factors may influence SGs concentration, independent of leaf size. It is plausible that SGs content and leaf size are not directly correlated but instead influenced by shared underlying traits related to biomass production. Therefore, careful consideration is warranted when selecting for both traits simultaneously in breeding programs to ensure optimal outcomes.

Leaves exhibit remarkable plasticity in shape and size, a characteristic influenced by diverse environmental factors (Tsukaya, 2005). Among the various determinants of leaf size, phytohormones such as auxins, cytokinins, gibberellins, and brassinosteroids are known to play pivotal roles by modulating cellular processes like cell proliferation and expansion (Wang et al.,

2021). Additionally, the *TCP* (*TEOSINTE BRANCHED*, *CYCLOIDEA*, and *PCF1/2*) transcription factor family, microRNAs including miR319 and miR396, and regulators of transcription factors orchestrate leaf size regulation through intricate and coordinated pathways (Kessler and Sinha, 2004; Wang et al., 2021). Despite the well-established roles of these factors in leaf development across various plant species, the genetic control mechanisms governing these traits remain largely unexplored in stevia. Consequently, there is a critical need for comprehensive investigations to elucidate the genetic underpinnings of leaf size regulation in this economically important crop.

Previous studies have primarily focused on identifying QTL associated with steviol glycoside compounds, with limited attention to leaf yield-related traits, except for overall plant vigor (Vallejo and Warner, 2021). In this study, we present the first set of QTL associated with several key agronomic traits related to leaf yield. These QTL exhibit minor to moderate effects (7-15% VE%) and are characterized by large intervals. Further refinement of these QTL regions through fine mapping approaches could potentially narrow down the genomic regions of interest (Su et al., 2010; Zhang et al., 2019). Interestingly, we observed QTL at different positions on the same linkage group across different environments, suggesting the possibility of genomic regions shifting positions across environments. To enhance the resolution of such QTL regions, future studies may benefit from employing a larger F_1 population and/or adopting a more refined composite interval mapping approach. Additionally, QTL associated with several traits were identified around 137 cM on linkage group 8. However, the utilization of these QTL warrants further validation due to the inconsistencies observed within this linkage group, as discussed above.

The absence of significant QTL for leaf count (development rate), a crucial trait related to stevia biomass, is noteworthy, especially considering its importance in biomass estimation (Benhmimou et al., 2017). This trait exhibited lower broad-sense heritability and the highest level of inconsistency in correlation studies, indicating potential complexity in its genetic regulation. A plausible explanation for this complexity could lie in the inconsistent phyllotaxy patterns observed in stevia. Phyllotaxy, the arrangement of leaves on the main stem, is typically categorized as alternate (one leaf per node) or opposite (two leaves per node) (Fleming, 2005; Lee et al., 2009). While stevia typically exhibits an opposite phyllotactic pattern (Rossi et al., 2018), our study observed instances of both opposite and alternate phyllotaxy patterns, as well as irregular phyllotaxy patterns where the arrangement shifted during plant development, particularly during the transition to the reproductive stage.

The change in phyllotaxy patterns, from opposite to alternate, could be attributed to changes in plastochron ratio and meristem characteristics, phenomena observed in other plant species (Jackson and Hake, 1999; Rutishauser and Peisl, 2001). Plastochron ratio, which measures the radial distances between successive leaf primordia emergence, influences leaf arrangement (Jean and Barab, 1998). Soybean serves as a notable example of a crop undergoing significant shoot architecture changes associated with phyllotaxy alteration during the transition from opposite to alternate patterns in the vegetative phase (Yoshikawa et al., 2013). The expression levels of microRNAs (miR156 and miR172) and their target genes, known to regulate phase changes, play a role in determining phyllotaxy in soybean (Wang et al., 2008; Preston et al., 2016). Additionally, mutants of cytokinin-related genes in maize and rice also exhibit altered phyllotaxy patterns (Giulini et al., 2004), further highlighting the polygenic nature of phyllotaxy regulation (Reinhardt and Kuhlemeier, 2002).

Given these complexities, understanding the genetic control of phyllotaxy in stevia warrants further investigation. Transcriptomic analysis of meristem-related tissues in plants with varying phyllotaxy patterns could provide valuable insights into the underlying genetic mechanisms.

CONCLUSION

In summary, this study marks a significant advancement in stevia breeding efforts by providing essential resources for pre-breeding initiatives. It represents the inaugural establishment of a high-density SNP-based linkage map in stevia, a pivotal tool for pinpointing QTL associated with traits essential for the stevia ideotype. The availability of molecular markers offers a promising avenue for the identification of closely linked candidate genes responsible for regulating stevia leaf yield-related traits. Additionally, the identification of QTL associated with these traits represents a pioneering accomplishment. Moving forward, this research sets the stage for further exploration of genomic regions housing narrower QTL, thereby facilitating the identification of potential candidate genes governing these morphological traits.

Tables & Figures

Table 4-1: Summary statistics of mapping of reads after applying the filtering criteria (MAPQ > 20 and retaining only primary alignment).

Sample	Total reads	Properly mapped reads	Proportion of properly mapped reads
10-19	17894754	8714343	0.486978
10-19TC	38120846	18881283	0.495301
10-RJR	17965456	8458058	0.470796
10-RJRTC	18407756	8968565	0.487217
18-02-001	51688736	24749681	0.478822
18-02-002	33176164	16298230	0.491263
18-02-003	19124020	9950417	0.52031
18-02-004	32836254	16049878	0.488785
18-02-005	21429010	11199114	0.522615
18-02-006	17236578	8237025	0.477881
18-02-007	35122182	17678803	0.503352
18-02-008	36034264	17754030	0.492699
18-02-009	38723628	19034414	0.491545
18-02-010	28429672	13625845	0.479283
18-02-011	25643738	13338078	0.52013
18-02-012	33743096	17020049	0.504401
18-02-013	32054420	15859122	0.494756
18-02-014	22032996	11044525	0.501272
18-02-015	34709156	17463025	0.503124
18-02-016	21123628	10790673	0.510834
18-02-017	17579632	9037168	0.51407
18-02-018	20493308	10468966	0.510848
18-02-019	20883766	10100630	0.483659
18-02-020	25693526	12630201	0.491571
18-02-021	23500376	11888442	0.505883
18-02-022	34172458	17276411	0.505565
18-02-023	20578526	10620452	0.516094
18-02-024	23454370	11354929	0.484129
18-02-025	23834752	11527986	0.483663
18-02-026	22830316	10931919	0.478833
18-02-027	14314336	7131934	0.498237
18-02-028	15354616	7487088	0.487612
18-02-029	14130078	6952402	0.492029

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-030	15901236	8164853	0.513473
18-02-031	17331626	8376943	0.483333
18-02-032	18313612	8868941	0.484281
18-02-033	23895294	12468451	0.521795
18-02-034	26777886	13279503	0.495913
18-02-035	17955358	9163890	0.510371
18-02-036	18281698	8426050	0.460901
18-02-037	46412048	21628661	0.466014
18-02-038	22610060	10634568	0.470347
18-02-039	14771850	7654207	0.518162
18-02-040	21996370	10513528	0.477967
18-02-041	17032786	8268206	0.485429
18-02-042	20681132	9775769	0.47269
18-02-043	72608284	33754267	0.464882
18-02-044	26985428	13164310	0.48783
18-02-045	29287560	14001553	0.478072
18-02-046	28447400	12993874	0.456768
18-02-047	24149284	11236333	0.465286
18-02-048	29573470	13520396	0.45718
18-02-049	16199286	7878658	0.486358
18-02-050	13752170	6559691	0.476993
18-02-051	15007832	7445304	0.496095
18-02-052	13625120	7028449	0.515845
18-02-053	16311636	8032800	0.492458
18-02-054	19042886	10533658	0.553154
18-02-055	18283850	8436376	0.461411
18-02-056	20644226	9861265	0.477677
18-02-057	21339260	10677828	0.500384
18-02-058	29206686	14416165	0.493591
18-02-059	17898950	9426232	0.526636
18-02-060	17011100	8348598	0.490774
18-02-061	42896590	21203176	0.494286
18-02-062	37240392	18703279	0.502231
18-02-063	21371388	10456629	0.489282
18-02-064	35876226	17355958	0.483773

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-065	24911928	11720683	0.470485
18-02-066	27340836	12670605	0.463432
18-02-067	33010450	16293668	0.493591
18-02-068	31037118	14818089	0.477431
18-02-069	41081398	20944203	0.509822
18-02-071	34479996	17228490	0.499666
18-02-072	36400154	16692446	0.458582
18-02-073	21738430	10178822	0.468241
18-02-074	15603084	7429163	0.476134
18-02-075	12553990	6032512	0.480525
18-02-076	15647526	6963657	0.445032
18-02-077	12800006	6319222	0.493689
18-02-078	13736464	6465501	0.470682
18-02-079	19364898	8710329	0.4498
18-02-080	18031568	8416653	0.466773
18-02-081	19779374	10302999	0.520896
18-02-082	19839408	9484067	0.478042
18-02-083	20243128	10236395	0.505673
18-02-084	18006824	7940583	0.440976
18-02-085	19757408	9401491	0.475846
18-02-086	15994378	7810672	0.488339
18-02-087	13034978	6216129	0.476881
18-02-088	15826766	7821408	0.494189
18-02-089	13241156	6853609	0.517599
18-02-090	15873488	7253398	0.45695
18-02-091	16972652	8028839	0.473046
18-02-092	17349306	8464687	0.487898
18-02-093	26989048	13499028	0.500167
18-02-094	20463422	9882142	0.482917
18-02-095	34433376	16113654	0.467966
18-02-096	25762974	12552389	0.487226
18-02-097	30716412	16131167	0.525164
18-02-098	28087438	13615557	0.484756
18-02-099	21513786	11283479	0.524477
18-02-100	28428720	14050205	0.494226
18-02-101	28062480	13376498	0.476668

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-102	32419994	16086930	0.496204
18-02-103	27825660	14703494	0.528415
18-02-104	16040760	7851702	0.489484
18-02-105	37241666	20002982	0.537113
18-02-106	54248968	26954344	0.496864
18-02-107	22087854	11167821	0.505609
18-02-108	17941924	9253429	0.515743
18-02-109	21254196	10941981	0.514815
18-02-110	20765546	11197792	0.539249
18-02-111	18791592	9254349	0.492473
18-02-112	18203110	9227745	0.506932
18-02-113	20499406	10813774	0.527516
18-02-114	20727494	10952674	0.528413
18-02-115	20245166	10378017	0.512617
18-02-116	21872490	11340750	0.518494
18-02-117	29961626	15152325	0.505724
18-02-118	20869092	10373005	0.497051
18-02-119	15534608	7720385	0.49698
18-02-120	14478364	7114785	0.491408
18-02-121	16599590	8441025	0.508508
18-02-122	15869356	8478670	0.534279
18-02-123	14725508	7157103	0.486034
18-02-124	13615646	6732593	0.494475
18-02-125	17736472	8512380	0.479936
18-02-126	15445700	8047021	0.520988
18-02-127	15996438	8081489	0.505206
18-02-128	17732046	9413602	0.530881
18-02-129	26615682	12950734	0.486583
18-02-130	27142988	12885760	0.474736
18-02-131	30780444	15104906	0.490731
18-02-132	27228004	13117474	0.481764
18-02-133	15278890	7369399	0.482326
18-02-134	31039374	16706741	0.538243
18-02-135	27344180	13012263	0.47587
18-02-136	25106052	12095922	0.481793

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-137	30790504	14958053	0.485801
18-02-138	28980968	15062367	0.519733
18-02-139	19000716	8942189	0.470624
18-02-140	29342312	14871952	0.506843
18-02-141	56823002	26951301	0.474303
18-02-142	21072820	10253037	0.486553
18-02-143	17424240	8646881	0.496256
18-02-144	15681938	8455751	0.539203
18-02-145	14199954	7038002	0.495636
18-02-146	15971086	8781363	0.549829
18-02-147	16327300	7930104	0.485696
18-02-148	14438942	7019450	0.486147
18-02-149	17844458	8677207	0.486269
18-02-150	16664160	8774952	0.526576
18-02-151	20326154	10051714	0.494521
18-02-152	17915848	9312839	0.51981
18-02-153	33924524	16012574	0.472006
18-02-155	33556346	16792429	0.500425
18-02-156	27717682	13730650	0.495375
18-02-157	28508732	14844781	0.52071
18-02-158	39874304	19658908	0.493022
18-02-159	28330004	13967619	0.493033
18-02-160	40373934	19677266	0.487375
18-02-161	29089152	15318263	0.526597
18-02-162	22426042	10943211	0.487969
18-02-163	35912798	18509443	0.5154
18-02-164	59892434	29264833	0.488623
18-02-165	19668322	9550478	0.485577
18-02-166	16045916	7629311	0.475467
18-02-167	14335940	6935264	0.483768
18-02-168	15374218	7521375	0.48922
18-02-169	15464558	7904625	0.511145
18-02-170	14817680	7198364	0.485796
18-02-171	16173442	7572186	0.468186
18-02-172	23243100	10661327	0.458688
18-02-173	15072942	7715600	0.511884

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-174	12709162	6073096	0.477852
18-02-175	16917618	8545150	0.505104
18-02-176	29207724	13740488	0.47044
18-02-177	18654478	9254241	0.496087
18-02-178	18626800	9263619	0.497327
18-02-179	14509252	7432111	0.512233
18-02-180	17471908	8716281	0.498874
18-02-181	15595416	8239997	0.52836
18-02-182	13935442	6902556	0.495324
18-02-183	15085828	7282258	0.482722
18-02-185	16591612	8452528	0.509446
18-02-186	12062060	5821126	0.482598
18-02-187	16961964	9131832	0.538371
18-02-188	24602578	11672623	0.474447
18-02-189	26694398	13139722	0.492228
18-02-190	33126406	17085242	0.515759
18-02-191	20572670	10361734	0.503665
18-02-192	35718184	17652547	0.494217
18-02-193	23731464	11645364	0.490714
18-02-194	37497754	17905247	0.477502
18-02-195	23781060	12199166	0.512978
18-02-196	26766274	13451279	0.502546
18-02-197	19985882	10685577	0.534656
18-02-198	21390124	10907710	0.509941
18-02-199	20164132	10688256	0.530063
18-02-200	18756874	9516616	0.507367
18-02-201	19731400	9757015	0.494492
18-02-202	20355152	10191345	0.500676
18-02-203	21712920	11175023	0.514672
18-02-204	14150408	7280458	0.514505
18-02-205	16032844	8049795	0.502082
18-02-206	14176806	7486561	0.528085
18-02-208	18373890	9107310	0.495666
18-02-209	24446600	11822804	0.483618
18-02-210	33335772	16522077	0.495626
18-02-211	19329716	9758251	0.504832
18-02-212	25829732	12666874	0.490399

Table 4-1 (cont'd)

Sample	Total reads	Properly mapped reads	Properly mapped reads (%)
18-02-213	20250972	10511927	0.519083
18-02-214	23376876	11654153	0.498533
18-02-215	56464212	27140322	0.480664
18-02-216	22602792	10922735	0.483247
18-02-217	23710262	11990451	0.505707
18-02-218	14911332	7607645	0.510192
18-02-219	16898080	8406836	0.497502
18-02-220	14693290	7442571	0.506529
18-02-221	16331716	8117625	0.497047
18-02-222	17002306	8403388	0.49425
18-02-223	34840306	17050064	0.489378
18-02-224	27331988	14099171	0.515849
18-02-225	25864884	13277392	0.513337
18-02-226	31844550	16018020	0.503007
18-02-227	26103800	13900143	0.532495
18-02-228	24175818	12166127	0.503235
18-02-229	13654834	7144398	0.523214
18-02-230	19916222	9349335	0.469433
18-02-231	23785374	11750317	0.494014
18-02-232	13910288	6770834	0.48675
18-02-233	17892948	8522394	0.476299
18-02-234	13548406	7302682	0.539007
18-02-235	15376700	7370677	0.479341
18-02-236	21074016	10289962	0.488277
18-02-237	20751294	10522899	0.507096
18-02-238	13809302	6916828	0.500882

Table 4-2. Summary of linkage map generated by genotyping 234 individuals from stevia MSU 18-02 F_1 population.

Linkage group	Length (cM)	Number of markers	Average marker density (cM)
1	201.6	152	1.33
2	325.5	209	1.56
3	168.5	122	1.38
4	77.1	67	1.15
5	151.3	89	1.7
6	240.7	171	1.41
7	266.7	145	1.84
8	1044.1	18	58.0
9	172.2	127	1.35
10	159.1	108	1.47
11	185.0	114	1.62
Total	2991.8	1322	6.62

Figure 4-1: Comparison of 11 linkage groups with eleven chromosomes (Xu et al., 2021) by using AllMaps. Figure on left represents comparison of genetic positions (cM) of each linkage group with the physical positions (Mb) of corresponding chromosomes by straight lines. Figure on right represents the same comparison by dotted plot.

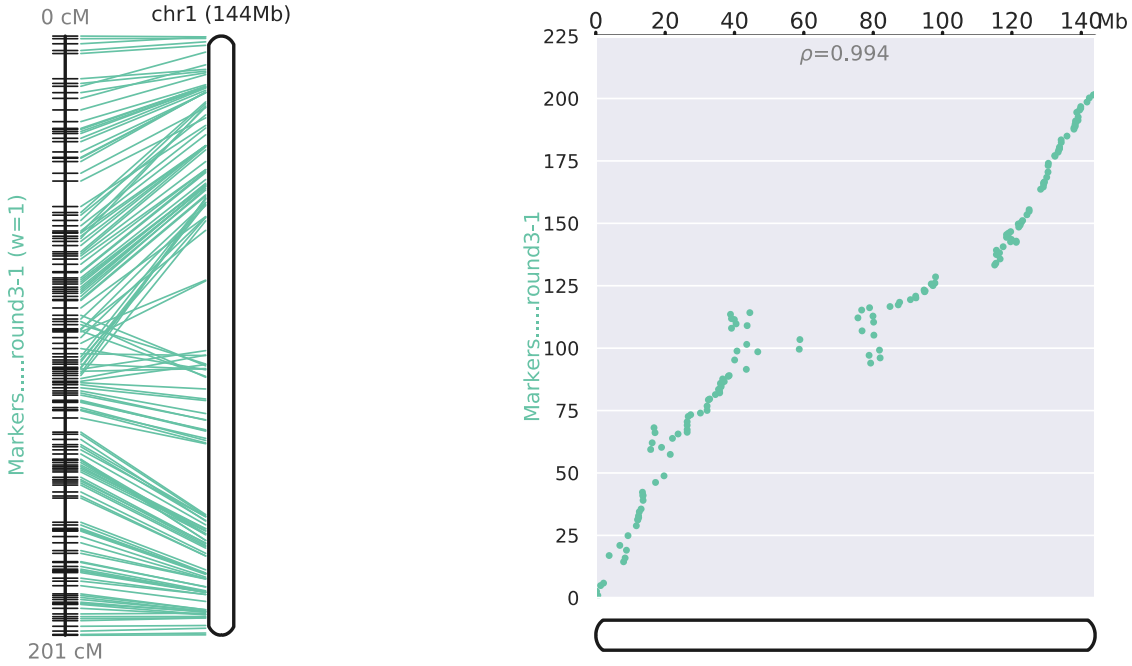


Figure 4-1 (cont'd)

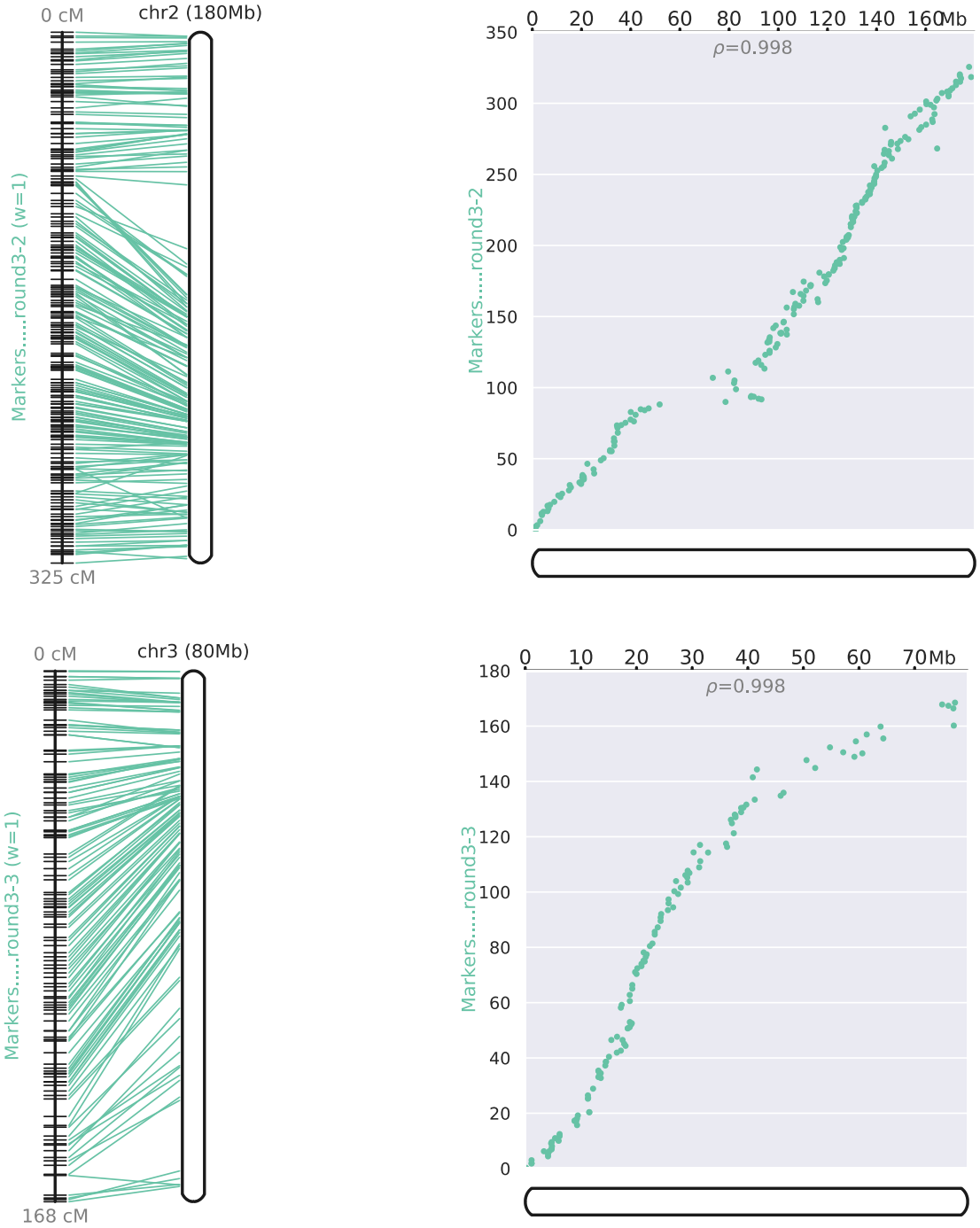


Figure 4-1 (cont'd)

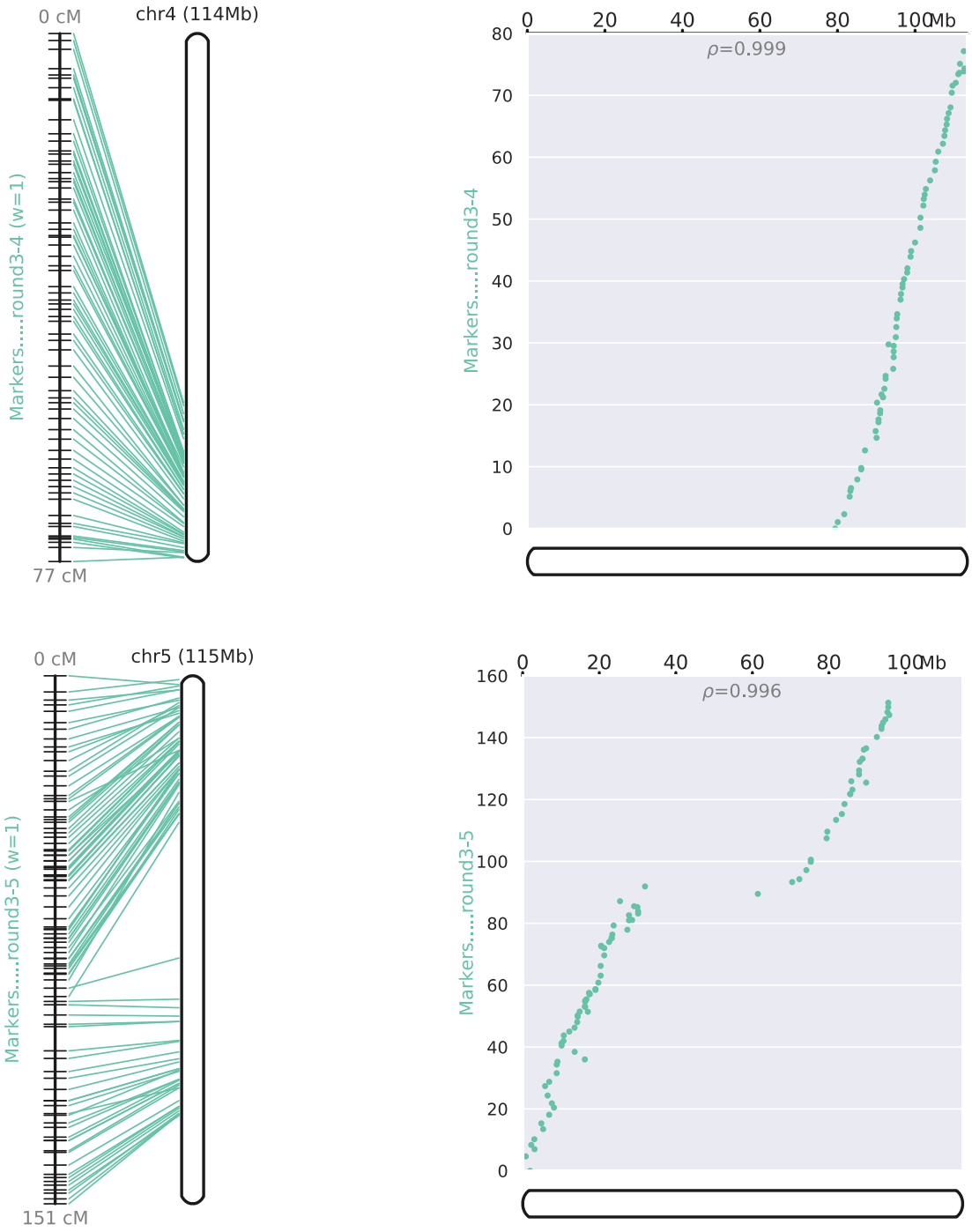


Figure 4-1 (cont'd)

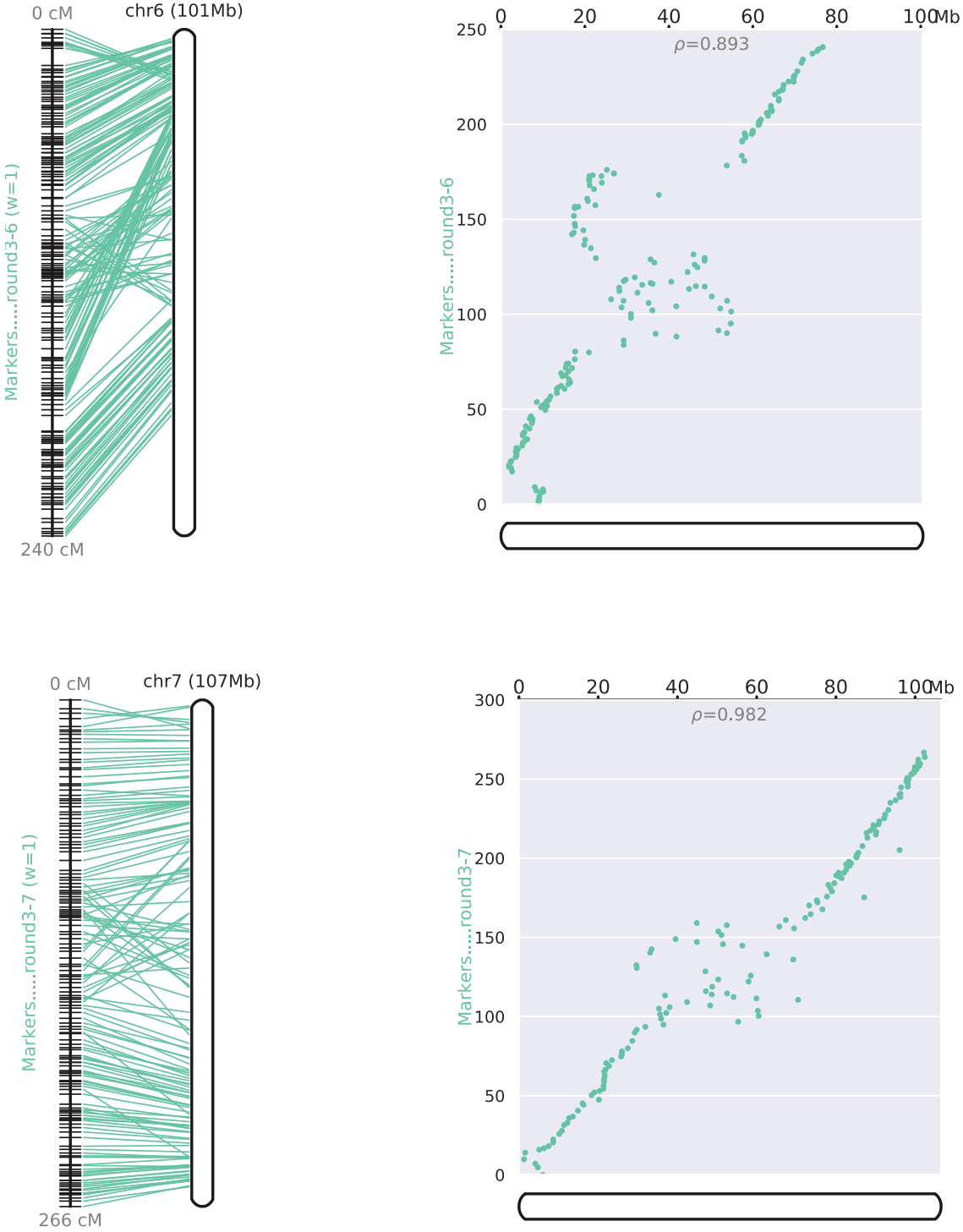


Figure 4-1 (cont'd)

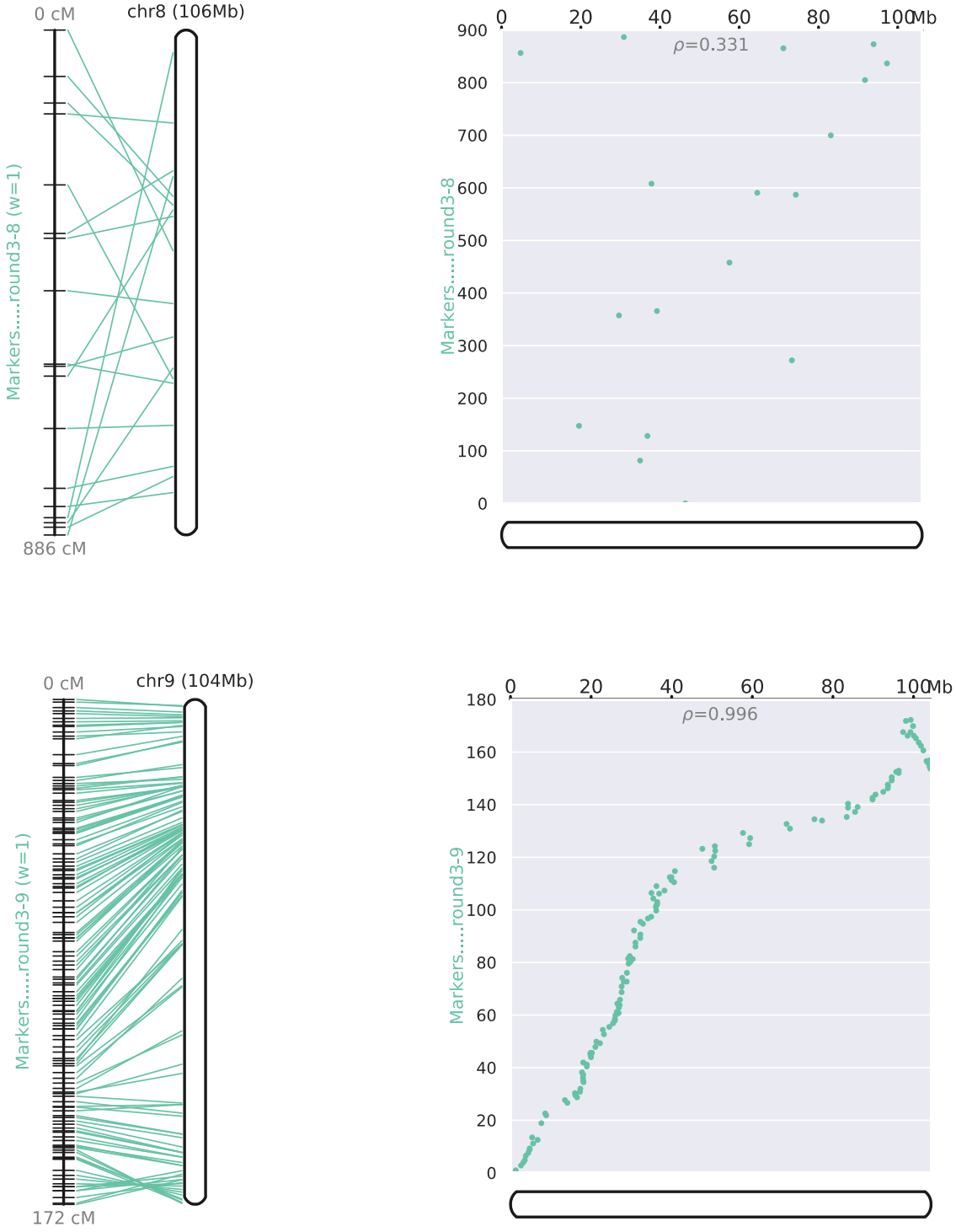


Figure 4-1 (cont'd)

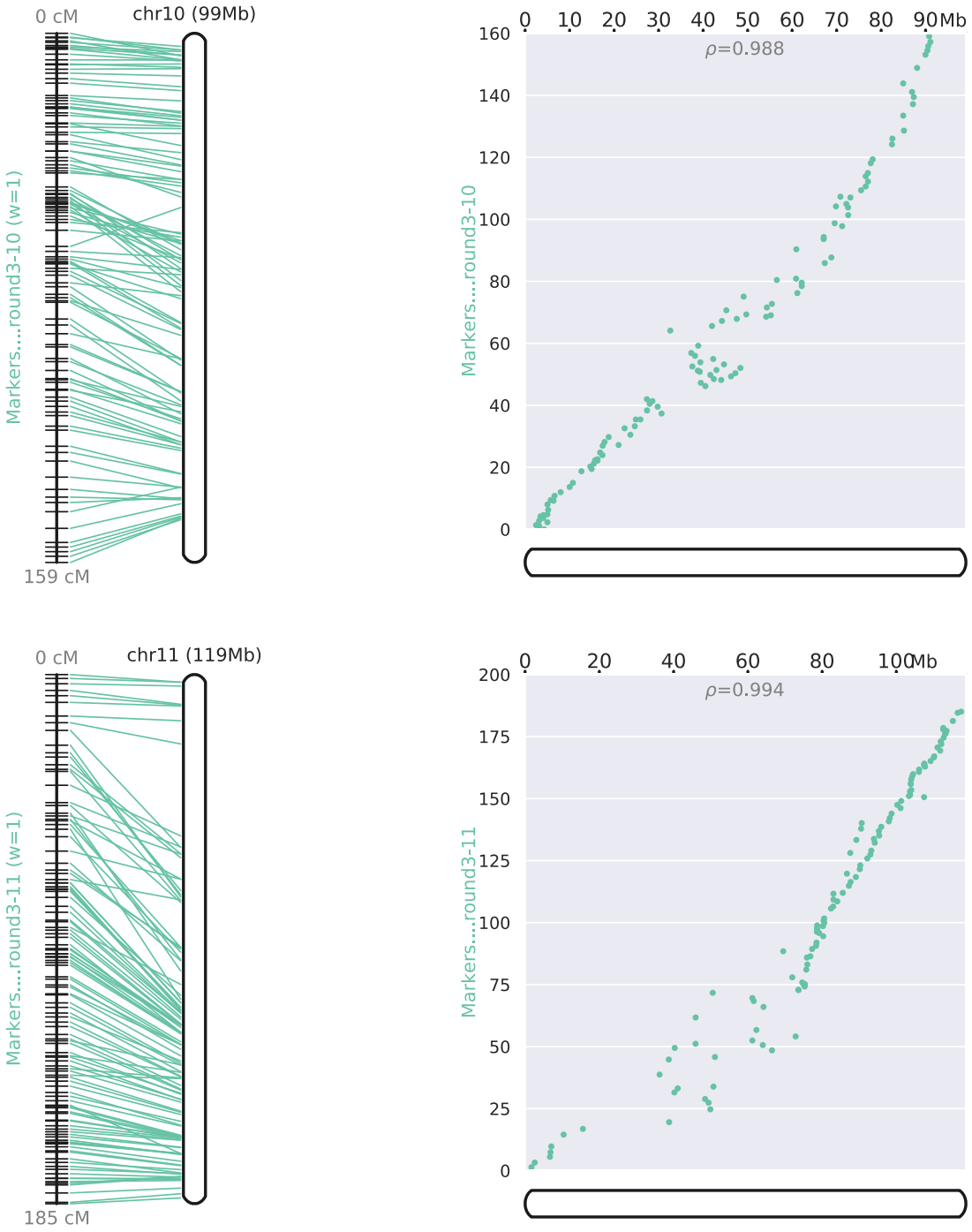


Figure 4-2: Population distribution of MSU18-02 F_1 population for all traits at Horticulture Teaching Research Center (HTRC), Holt, MI in 2020. Each panel represents the population distribution of a single trait. Panels (a) through (k) represent, plant height (a), number of leaves (b), maximum canopy width (c), minimum canopy width (d), leaf length (e), leaf width (f), primary branching (g), secondary branching (h), stem caliper (i), plant vigor (j) and flowering stage (k). Arrows represent parental mean values.

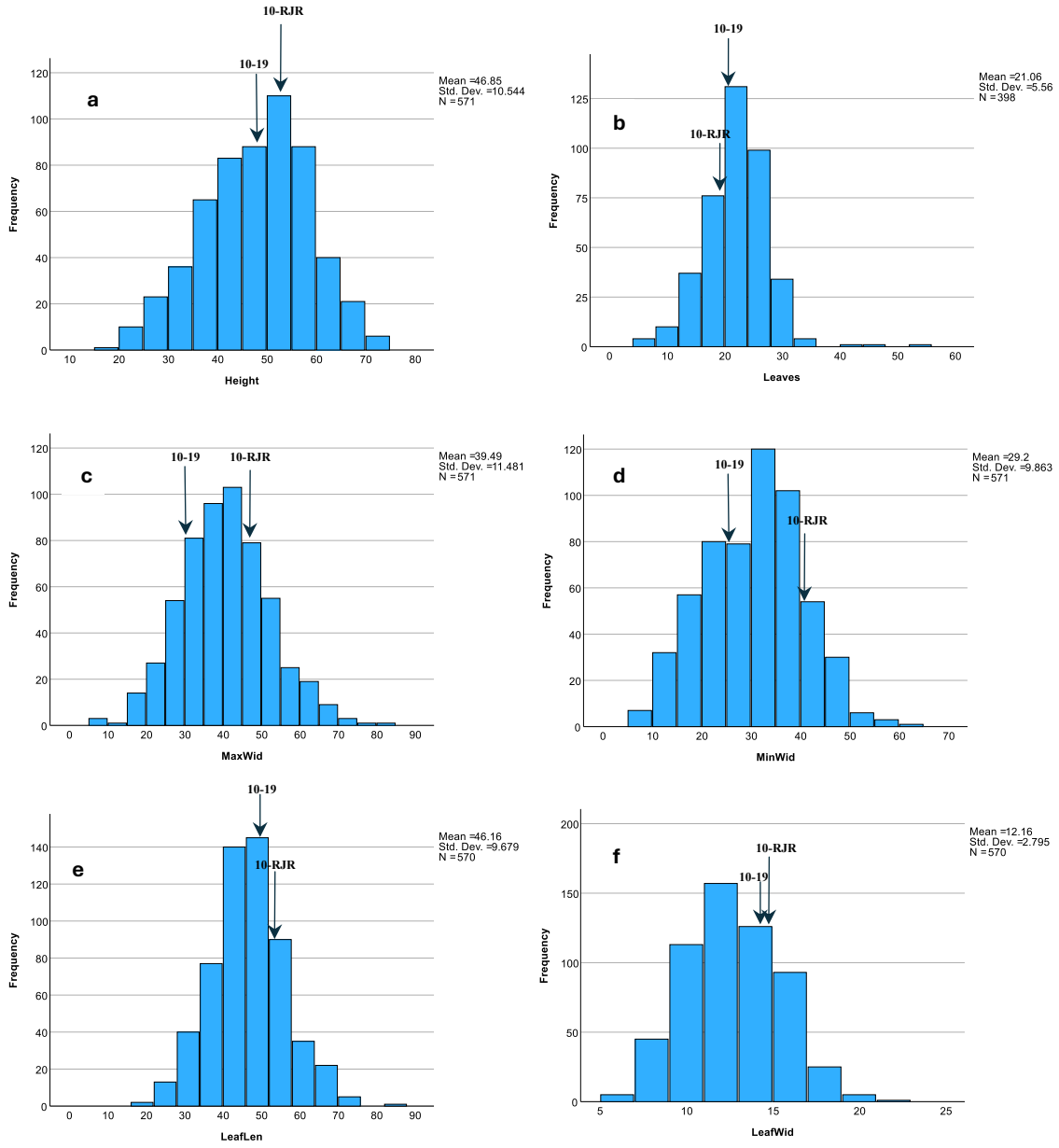


Figure 4-2 (cont'd)

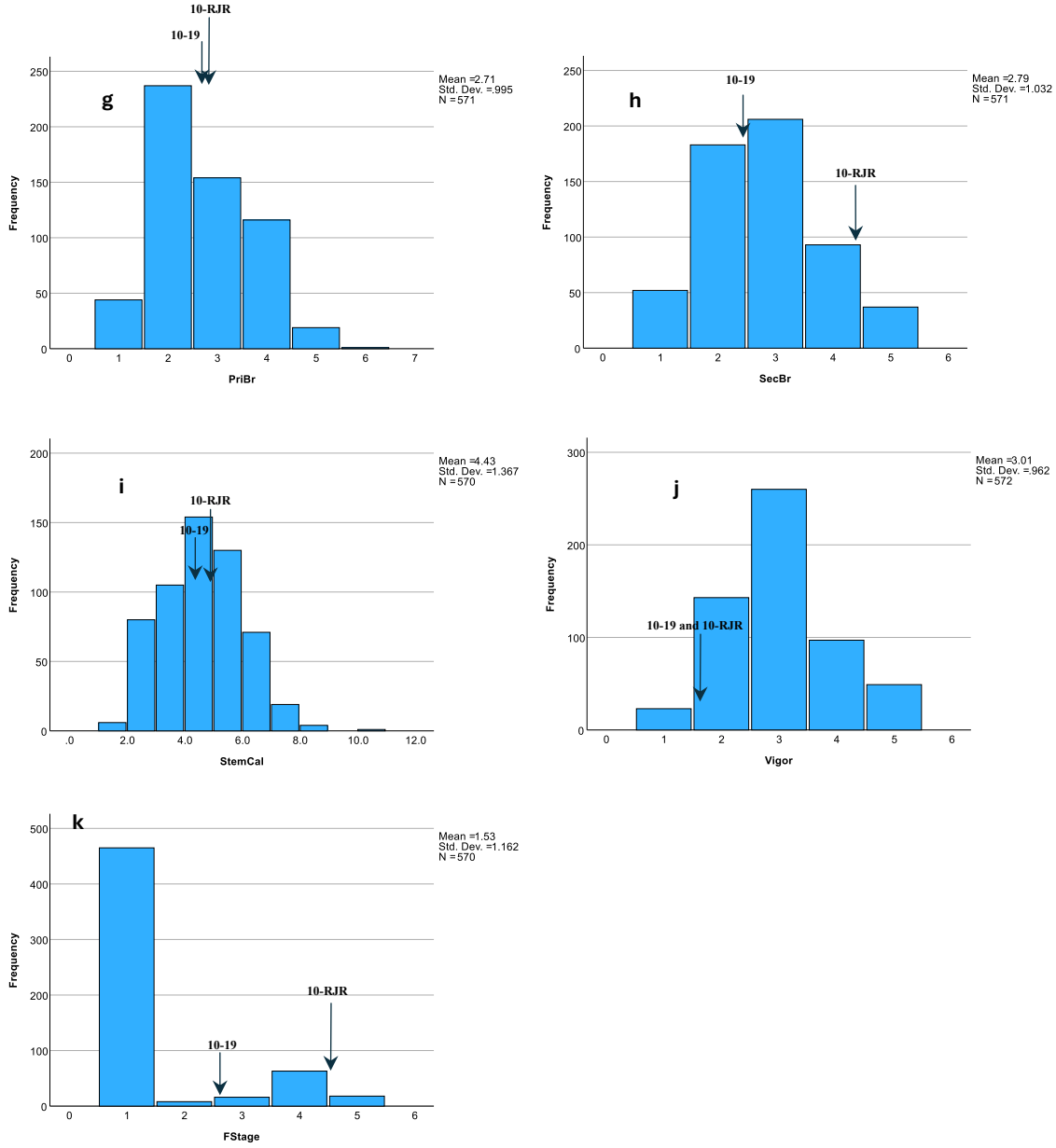


Figure 4-3: Population distribution of MSU18-02 F_1 population for all traits at Southwest Michigan Research and Education Center (SWMREC), Benton Harbor, MI in 2020. Each panel represents the population distribution of a single trait. Panels (a) through (k) represent, plant height (a), number of leaves (b), maximum canopy width (c), minimum canopy width (d), leaf length (e), leaf width (f), primary branching (g), secondary branching (h), stem caliper (i), plant vigor (j) and flowering stage (k). Arrows represent parental mean values.

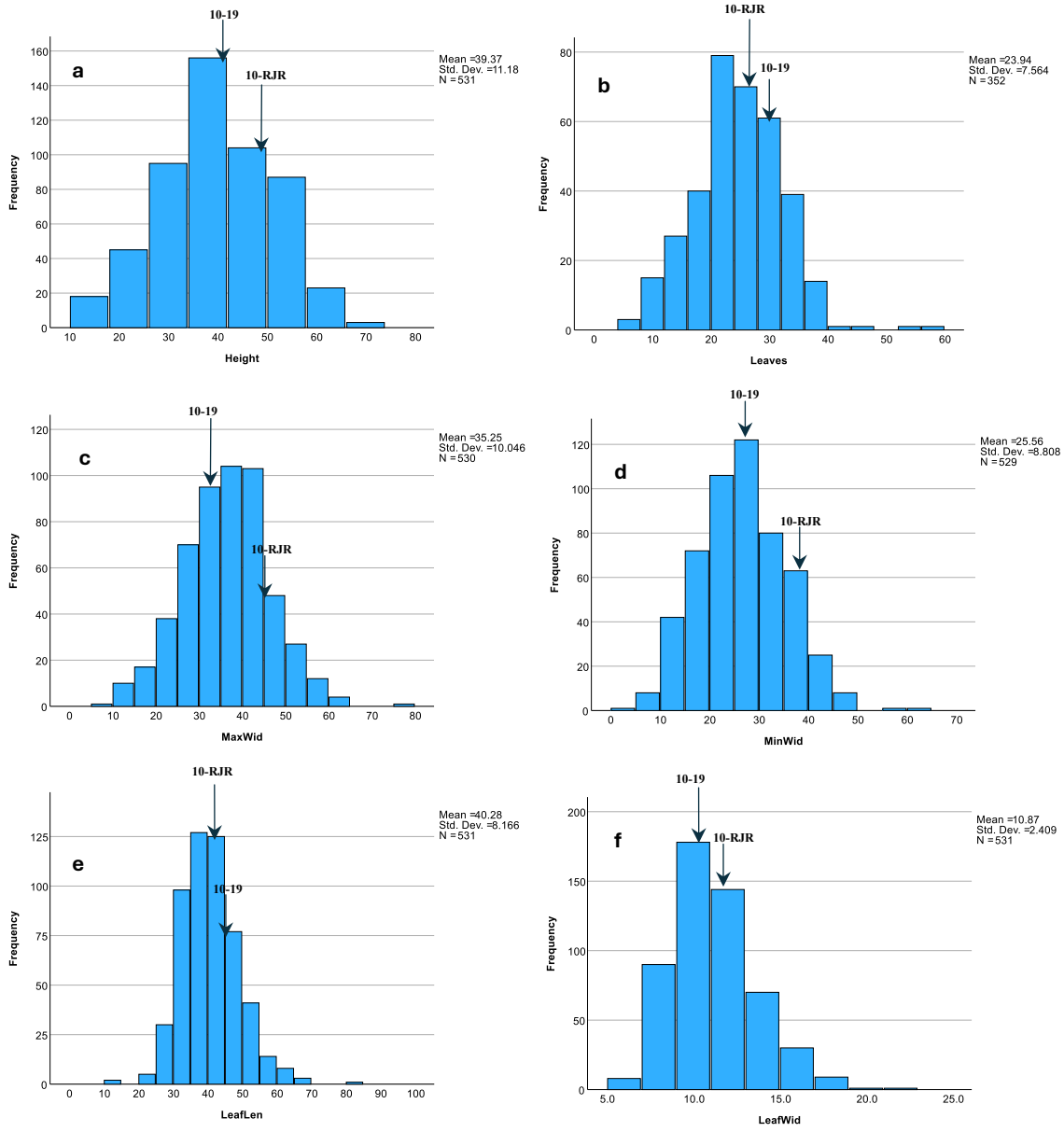


Figure 4-3 (cont'd)

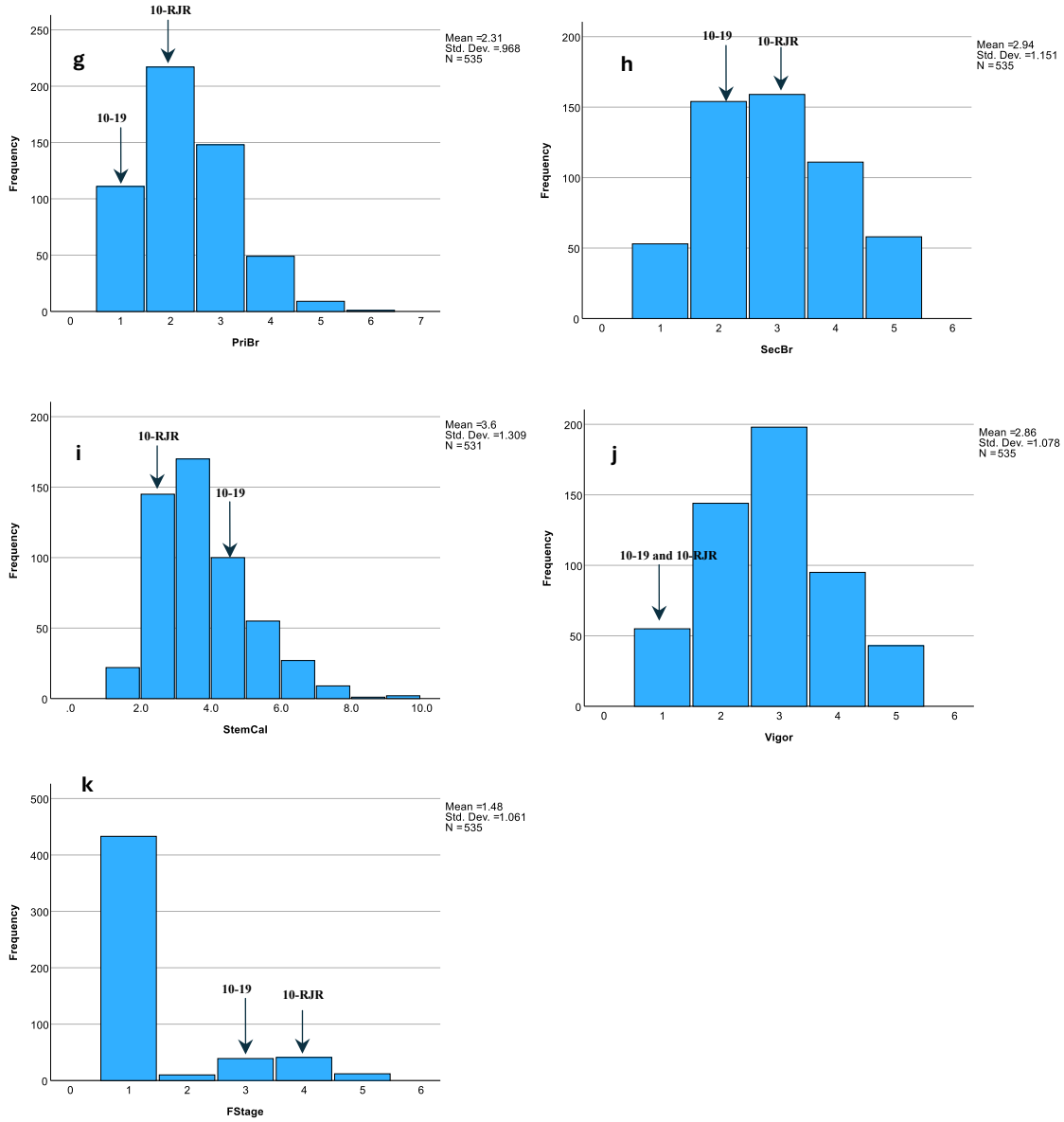


Figure 4-4: Population distribution of MSU18-02 F_1 population for all traits at Horticulture Teaching Research Center (HTRC), Holt, MI in 2021. Each panel represents the population distribution of a single trait. Panels (a) through (i) represent, maximum canopy width (a), minimum canopy width (b), leaf length (c), leaf width (d) number of leaves (e), secondary branching (f), stem caliper (g), plant vigor (h) and flowering stage (i). Arrows represent parental mean values.

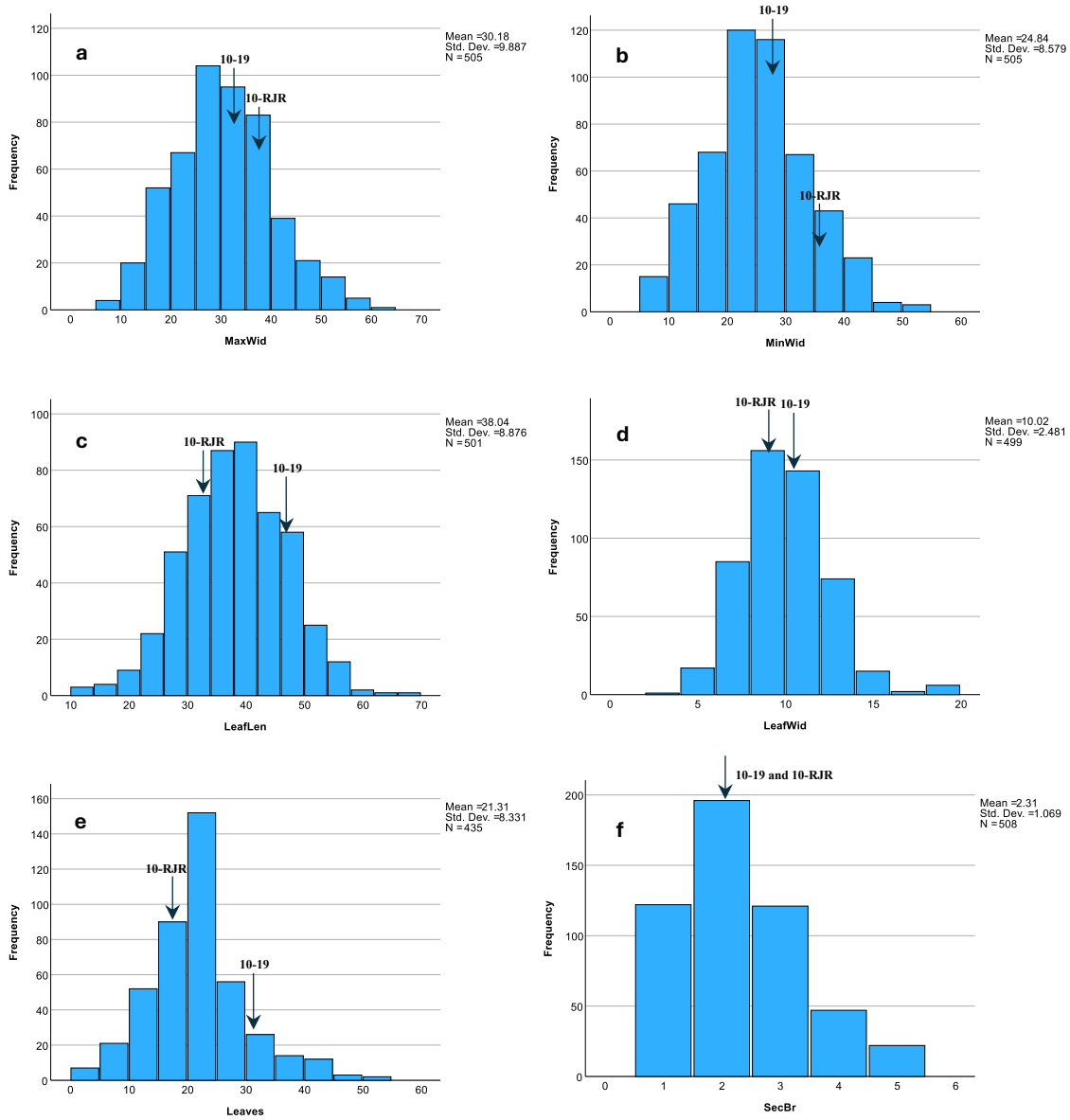
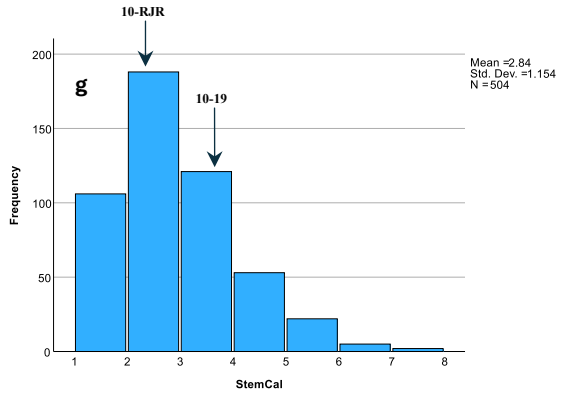
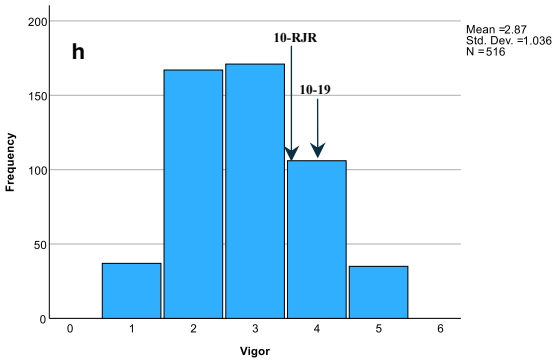


Figure 4-4 (cont'd)





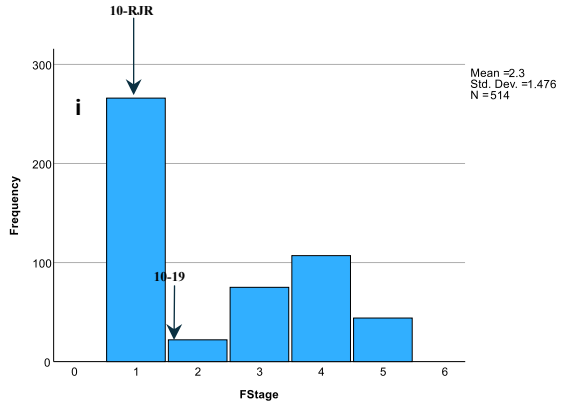


Figure 4-5: Population distribution of MSU18-02 F_1 population for all traits at Southwest Michigan Research and Education Center, Benton Harbor (SWMREC), MI 2021. Each panel represents the population distribution of a single trait. Panels (a) through (i) represent, maximum canopy width (a), minimum canopy width (b), leaf length (c), leaf width (d) number of leaves (e), secondary branching (f), stem caliper (g), plant vigor (h) and flowering stage (i). Arrows represent parental mean values.

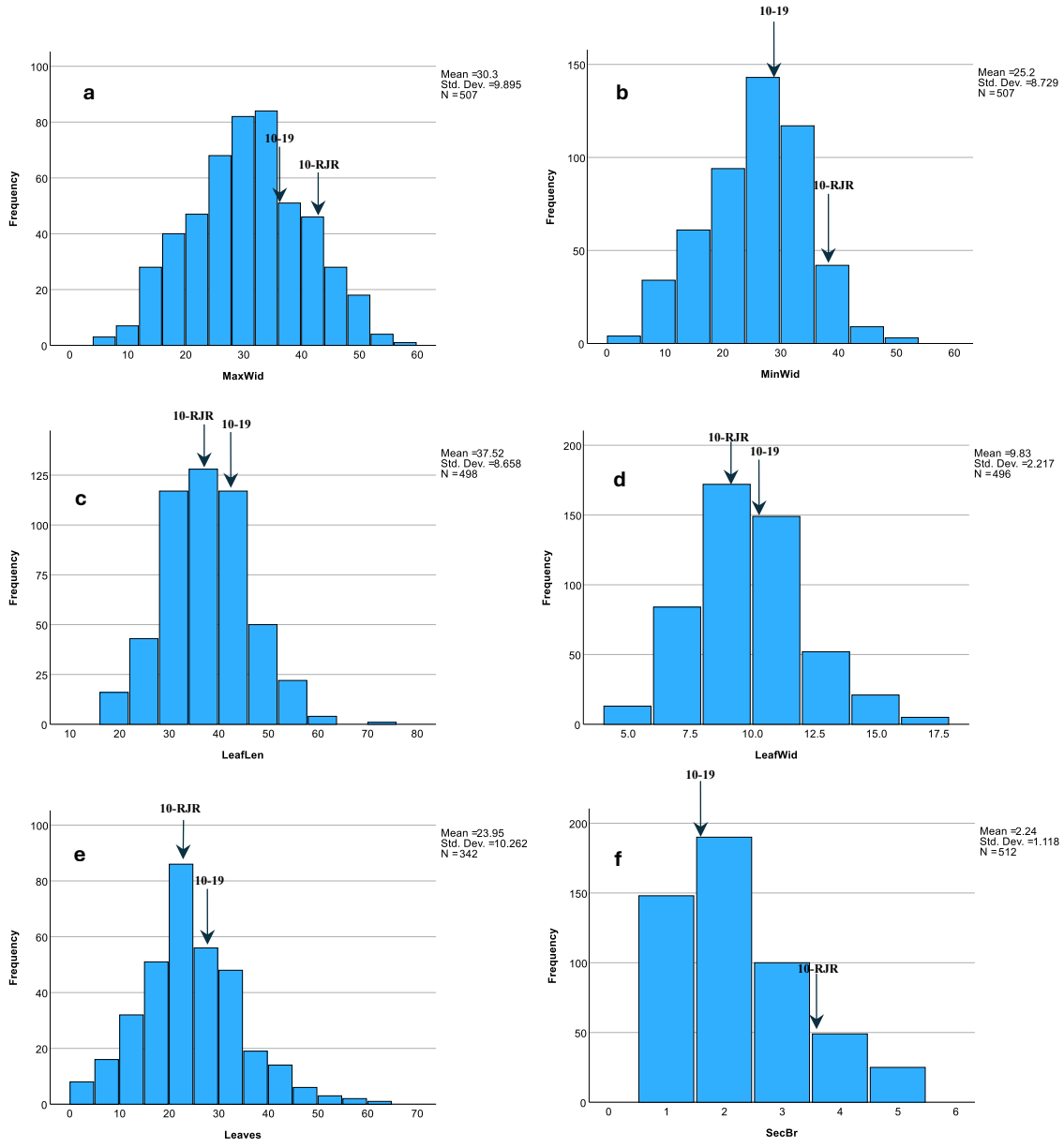
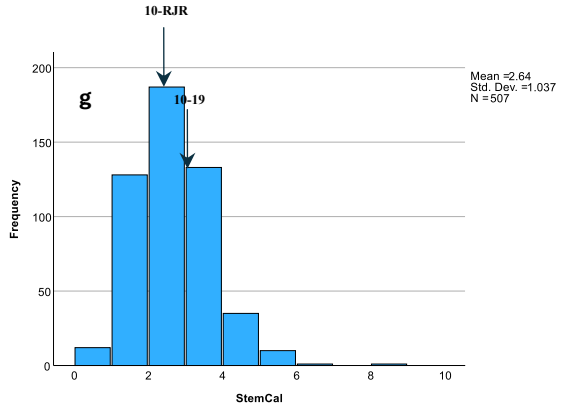
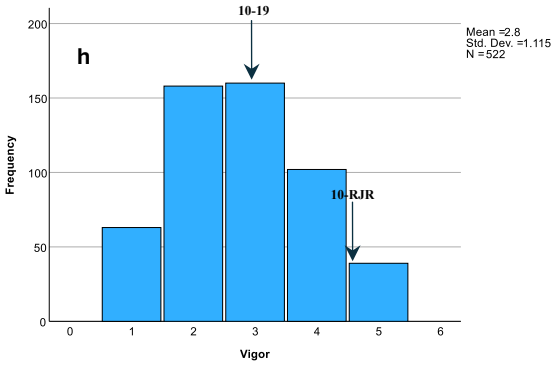


Figure 4-5 (cont'd)





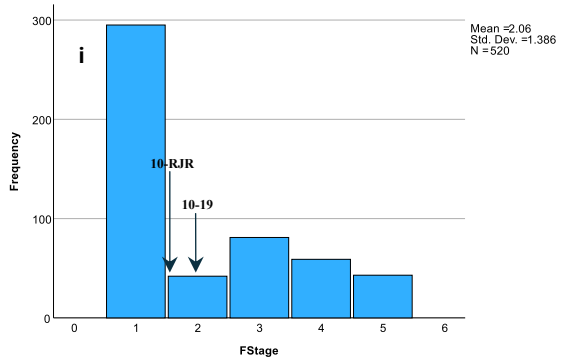


Table 4-3: Descriptive statistics of MSU18-02 F_1 mapping population for 11 leaf yield traits: Leaf length (LeafLen), Leaf width (LeafWid), Primary branching (PriBr), Secondary branching (SecBr), Minimum canopy width (MinWid), Maximum Canopy Width (MaxWid), Number of leaves (Leaves), Stem caliper (StemCal), Plant Vigor (Vig), Flowering stage (FStage) and Plant height (Hght) at two locations (HTRC, Holt, Michigan and SWMREC, Benton Harbor, MI) and two years (2020 and 2021). N represents the total number of progeny individuals for which data is available (across at least two replications to a maximum of three replications at all environments). For the parental lines, N = 3 for all traits. SD represents standard deviation of mean.

	MSU 18-02 population					**Parental means**	
Trait	**N**	**Minimum**	**Maximum**	**Mean**	**SD**	**10-19**	**10-RJR**
			HRTC 2020				
LeafLen (mm)	564	17	82	46.12	9.70	48.00	52.33
LeafWid (mm)	564	5	21	12.13	2.80	14.33	14.67
PriBr	571	1	6	2.71	0.99	2.33	2.67
SecBr (1-5)	571	5	5	2.79	1.03	2.00	4.00
MinWid (cm)	565	5	60	29.15	9.86	25.67	41.33
MaxWid (cm)	565	8	80	39.44	11.50	35.00	52.00
Leaves	398	4	54	21.06	5.56	20.67	19.67
StemCal (mm)	564	1.0	10.0	4.44	1.37	3.10	3.63
Vig (1-5)	570	1	5	1.53	1.16	1.00	1.00
FStage (1-5)	572	1	5	3.01	0.96	2.00	4.00
Hght (cm)	565	16.00	74.00	46.82	10.58	48.00	53.00
			SWMREC 2020				
LeafLen (mm)	534	14.0	83.0	40.23	8.17	45.17	42.63
LeafWid (mm)	534	5.4	21.0	10.86	2.40	11.33	12.33
PriBr	535	1	6	2.31	0.97	1.67	2.33
SecBr (1-5)	535	1	5	2.94	1.15	2.33	3.67

Table 4-3 (cont'd)

		MSU 18-02 population				Parental means	
Trait	**N**	**Minimum**	**Maximum**	**Mean**	**SD**	**10-19**	**10-RJR**
			SWMREC 2020				
MaxWid (cm)	533	7.5	79.0	35.15	10.11	33.83	45.33
MinWid (cm)	532	4.0	61.5	25.48	8.86	27.17	38.83
Leaves	352	4	56.0	23.94	7.56	30.00	27.33
StemCal (mm)	534	1.0	9.3	3.59	1.31	4.60	2.73
Vig (1-5)	535	1	5	1.48	1.06	1.00	1.00
FStage (1-5)	535	1	5	2.86	1.08	3.00	4.00
Hght (cm)	534	10.5	70.0	39.29	11.21	41.40	48.83
			HTRC 2021				
LeafLen (mm)	501	11	66	38.04	8.88	46.33	32.13
LeafWid (mm)	499	3.3	19.6	10.02	2.48	11.00	9.53
SecBr (1-5)	508	1	5	2.31	1.07	2.33	2.33
MinWid (cm)	505	6	50	24.84	8.58	28.33	35.67
MaxWid (cm)	505	8	60	30.18	9.89	33.00	38.97
Leaves	363	2	52	22.01	8.35	31.33	18.00
StemCal (mm)	504	1	8	2.84	1.15	3.83	2.33
Vig (1-5)	516	1	5	2.87	1.04	4.00	3.67
FStage (1-5)	514	1	5	2.30	1.48	1.67	1.00
			SWMREC 2021				
LeafLen (mm)	498	17	73	37.52	8.66	43.53	38.17
LeafWid (mm)	496	4.5	17.9	9.83	2.22	10.37	9.07

Table 4-3 (cont'd)

MSU 18-02 population						**Parental means**	
Trait	**N**	**Minimum**	**Maximum**	**Mean**	**SD**	**10-19**	**10-RJR**
SecBr (1-5)	512	1	5	2.24	1.12	1.67	3.67
MinWid (cm)	507	3	51	25.20	8.73	29.67	37.00
MaxWid (cm)	507	4	57	30.30	9.89	36.00	44.33
Leaves	342	3	62	23.95	10.26	28.67	23.33
StemCal (mm)	507	1	9	2.64	1.04	3.07	2.47
Vig (1-5)	522	1	5	2.80	1.11	3.00	4.67
FStage (1-5)	520	1	5	2.06	1.39	2.00	1.67

Table 4-4: Summary of phyllotaxy patterns of MSU18-02 population at all four environments (years and locations). Numbers represent percentage of plants showing each phyllotaxy type among the total number of plants for which phyllotaxy was recorded.

Environment	Opposite (%)	Alternate (%)	Hybrid (%)
HTRC 2020	55.39	39.13	5.48
SWMREC 2020	51.02	27.54	21.44
HTRC 2021	49.70	48.23	2.03
SWMREC 2021	42.54	45.67	11.79

Table 4-5. Pearson correlation coefficients for 11 traits in stevia biparental cross population MSU18-02 at two locations (HTRC and SWMREC) over two years (2020 and 2021). ** indicates correlation is significant at a p-value of 0.01 and * indicates significant correlation at a p-value of 0.05.

	Hght	**LeafLen**	**LeafWid**	**PriBr**	**SecBr**	**Leaves**	**MinWid**	**MaxWid**	**Vig**	**StemCal**
					HTRC 2020					
LeafLen	.313**									
LeafWid	.271**	.572**								
PriBr	.300**	0.096	0.072							
SecBr	.230**	.246**	.236**	.184*						
Leaves	.194*	-0.082	0.075	0.041	-.333**					
MinWid	.377**	.282**	.295**	.298**	.764**	-.184*				
MaxWid	.418**	.309**	.343**	.313**	.722**	-.217**	.839**			
Vig	.377**	.291**	.304**	.277**	.859**	-.183*	.764**	.746**		
StemCal	.436**	.527**	.530**	-0.065	0.113	0.107	.168*	.200**	.187**	
FStage	-0.107	-0.068	-0.135	-0.111	0.104	-.208*	0.064	0.041	0.024	-0.137
					SWMREC 2020					
LeafLen	.352**									

Table 4-5 (cont'd)

	Hght	LeafLen	LeafWid	PriBr	SecBr	Leaves	MinWid	MaxWid	Vig	StemCal
LeafWid	.270**	.546**								
PriBr	.149*	-0.007	-0.019							
SecBr	.243**	0.103	0.03	.142*						
Leaves	.324**	.190*	0.132	-.164*	-.183*					
MinWid	.425**	.182*	0.055	.290**	.720**	-0.076				
MaxWid										
Vig	.505**	.235**	.178*	.265**	.823**	-0.006	.814**	.807**		
StemCal	.489**	.483**	.589**	0.009	0.009	.332**	0.138	.246**	.261**	
FStage	-0.049	-0.074	-.225**	-.151*	.225**	-0.053	.181*	0.097	0.096	-.195**
HTRC 2021										
LeafWid		.482**								
SecBr		0.106	0.112	.614**						
Leaves		.228**	.210*	0.14	0.059					
MinWid		.220**	0.114	.683**	.731**	0.152				

Table 4-5 (cont'd)

	Hght	LeafLen	LeafWid	PriBr	SecBr	Leaves	MinWid	MaxWid	Vig	StemCal
MaxWid		.216**	0.129	.651**	.694**	0.119	.893**			
Vig		.304**	.244**	.692**	.792**	0.152	.797**	.781**		
StemCal		.554**	.453**	0.107	0.063	.314**	.187*	.172*	.321**	
FStage		-0.048	-.203**	0.028	.239**	-0.015	.267**	.229**	.211**	-0.007
SWMREC 2021										
LeafWid		.517**								
SecBr		.250**	.167*	.648**						
Leaves		.183*	.239**	.199*	.208*					
MinWid		.306**	.170*	.712**	.651**	.246**				
MaxWid		.355**	.249**	.699**	.647**	.328**	.921**			
Vig		.360**	.252**	.752**	.703**	.277**	.784**	.775**		
StemCal		.472**	.434**	.208**	.156*	.319**	.295**	.327**	.408**	
FStage		0.041	-0.103	-0.098	0.017	-0.021	.150*	0.145	0.01	-0.041

Table 4-6. Broad sense heritability (H^2) estimates of leaf-yield traits by utilizing data across all four environments (HTRC 2020, SWMREC 2020, HTRC 2021 and SWMREC 2021).

Trait (unit/index)	H^2
Hght (cm)	0.50
LeafWid (mm)	0.64
LeafLen (mm)	0.63
MinWid (cm)	0.58
MaxWid (cm)	0.52
Leaves (count)	0.39
StemCal (mm)	0.23
PriBr (count)	0.33
SecBr (1-5)	0.60
Vig (1-5)	0.62
FStage (1-5)	0.55

Table 4-7. QTL summary of 11 leaf yield traits for MSU18-02 F_1 mapping population phenotyped at two locations (HTRC, Holt, Michigan and SWMREC, Benton Harbor, MI) and two years (2020 and 2021): Trait abbreviations used for all traits: Leaf length (LL), Leaf width (LW), Primary branching (PBr), Secondary branching (SBr), Minimum canopy width (MiW), Maximum canopy width (MW), Number of leaves (L), Stem caliper (SC), Flowering stage (FS), Plant Vigor (V) and Plant height (H). QTL names start with a q followed by a trait abbreviation, middle part represents the location and year combination HTRC and SWMREC (h and s) and two years, 2020 and 2021 (20 and 21) and the last two digits represent the linkage group and the QTL number on each linkage group. QTL in bold indicate significant QTL (overlapping peaks or regions) at more than one environment for each trait.

Trait	Environment	QTL	LG	Marker	Position (cM)	LOD	LOD threshold	VE %
PriBr	HTRC 2020	qPBrh20.2.1	2	Chr2_45671067	84.13	4.05	2.95	9.3
	HTRC 2020	qPBr.h20.5.1	5	Chr5_21303512	71.61	2.95	2.95	13.6
	HTRC 2020	qPBrh20.8.1	8	Chr8_76558323	357.29	5.12	2.95	11.6
	HTRC 2020	qPBrh20.9.1	9	Chr9_29289415	78.03	3.91	2.95	8.9
	HTRC 2020	qPBrh20.10.1	10	Chr10_70858836	107.29	3.4	2.95	7.8
SecBr	SWMREC 2021	qSBr.s21.1.1	1	Chr1_19664284	47.19	3.08	2.9	8.2
	HTRC 2020	**qSbr.h20.2.1**	2	Chr2_88859523	93.37	3.67	3.0	10.3
	SWMREC 2020	**qSbr.s20.2.1**	2	Chr2_88859523	93.37	4.62	2.9	10.5
	SWMREC 2020	qSBr.s20.5.1	5	Chr5_70413501	93.32	3.28	2.9	7.6
	HTRC 2021	qSBr.h21.5.1	5	Chr5_8875199	31.53	4.89	3.0	12.1
	SWMREC 2021	qSBR.s21.5.1	5	Chr5_14226909	48.03	3.26	2.9	7.8
	HTRC 2020	**qSbr.h20.7.1**	7	Chr7_67368224	161.90	4.52	3.0	9.4
	SWMREC 2020	**qSBr.s20.7.1**	7	Chr7_67368224	161.90	6.90	2.9	15.3
	SWMREC 2021	**qSBrs.s21.7.1**	7	Chr7_67368224	160.90	3.97	2.9	9.9
	HTRC 2020	**qSbr.h20.8.1**	8	Chr8_69370654	134.20	4.1	3.0	8.4
	SWMREC 2020	**qSBr.s20.8.1**	8	Chr8_69370654	136.20	5.31	2.9	12.0
	SWMREC 2020	**qSBr.s20.11.1**	11	Chr11_101141677	145.92	4.08	2.9	9.3
	HTRC 2021	**qSBr.h21.11.1**	11	Chr11_98216818	142.13	3.57	3.0	9
	SWMREC 2021	**qSBr.s21.11.1**	11	Chr11_101141677	146.14	3.55	2.9	8.9
MinWid	HTRC 2020	qMiW.h20.1.1	1	Chr1_121840368	149.71	3.06	2.94	7.2
	SWMREC 2021	qMiW.s21.1.1	1	Chr1_26652684	72.61	4.64	3.0	11.5

Table 4-7 (cont'd)

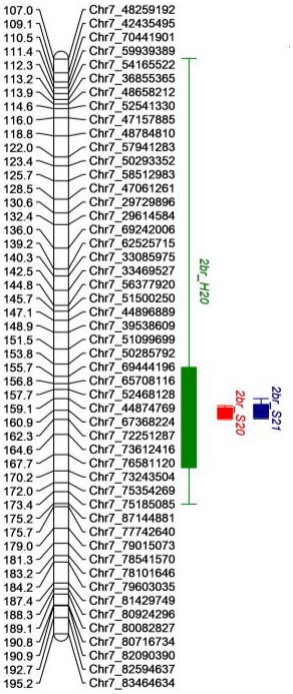
Trait	Environment	QTL	LG	Marker	Position (cM)	LOD	LOD threshold	VE %
	HTRC 2020	qMiW.h20.2.1	2	Chr2_129991039	220.57	3.51	2.94	7.2
	SWMREC 2021	qMiW.s21.5.1	5	Chr5_88787733	133.16	3.58	3.0	9.0
	HTRC 2020	qMiW.h20.7.1	7	Chr7_29729896	130.62	4.71	2.94	10.8
	SWMREC 2020	qMiW.s20.8.1	8	Chr8_69370654	137.20	3.31	2.95	7.7
	SWMREC 2020	qMiW.s20.7.1	7	Chr7_67368224	160.90	3.86	2.95	8.9
	HTRC 2020	**qMiW.h20.9.1**	9	Chr9_103237598	156.57	4.27	2.94	9.8
	SWMREC 2020	**qMiW.s20.9.1**	9	Chr9_103237598	156.57	2.91	2.95	6.8
	SWMREC 2021	qMiW.s20.9.1	9	Chr9_99875743	169.87	3.08	3.0	11.5
MaxWid	HTRC 2020	qMW.h20.1.1	1	Chr1_118883873	144.90	5.73	2.9	13.0
	SWMREC 2020	**qMW.s20.1.1**	1	Chr1_23732853	65.59	3.0	2.9	7.0
	HTRC 2021	qMW.h21.1.1	1	Chr1_12529469	34.40	3.68	3	9.2
	SWMREC 2021	**qMW.s21.1.1**	1	Chr1_26332160	67.29	3.15	2.9	7.9
	HTRC 2020	qMW.h20.7.1	7	Chr7_29729896	130.62	3.85	2.9	8.9
	SWMREC 2020	qMW.s20.7.1	7	Chr7_67368224	161.90	4.16	2.9	9.6
	HTRC 2020	**qMW.h20.8.1**	8	Chr8_69370654	140.20	3.67	2.9	8.5
	SWMREC 2020	**qMW.s20.8.1**	8	Chr8_69370654	137.20	3.7	2.9	8.6
	SWMREC 2020	qMW.s20.9.1	9	Chr9_103971984	154.71	3.23	2.9	7.5
	SWMREC 2021	qMW.s21.1.1	9	Chr9_99875743	169.55	4.26	2.9	10.6
LeafLen	HTRC 2020	qLL.h20.2.1	2	Chr2_137266803	235.39	4.08	3.1	9.4
	HTRC 2021	qLL.h21.8.1	8	Chr8_32893096	223.34	3.3	3.1	8.4
	SWMREC 2020	qLL.s20.9.1	9	Chr9_30645963	91.59	5.8	6.0	13.1
	SWMREC 2021	qLL.s21.9.1	9	Chr9_103971984	154.51	5.13	3.0	12.8
	SWMREC 2021	qLL.s21.9.1	9	Chr9_103971984	154.51	5.13	3.0	12.8
LeafWid	HTRC 2021	qLW.h21.1.1	1	Chr1_192607	0	3.13	3.1	8.0
	HTRC 2020	qLW.h20.2.1	2	Chr2_1373203	2.36	3.05	3.0	7.1
	HTRC 2021	qLW.h21.2.1	2	Chr2_73382612	106.93	4.22	3.1	10.6
	SWMREC 2021	qLW.s21.2.1	2	Chr2_41296939	76.25	4.11	3.1	10.4
Leaves	SWMREC 2021	qL.s21.1.1	1	Chr1_6937018	20.96	4.43	3.0	14.4

Table 4-7 (cont'd)

Trait	Environment	QTL	LG	Marker	Position (cM)	LOD	LOD threshold	VE %
	HTRC 2021	qL.h21.2.1	2	Chr2_40037011	83.83	3.31	3.0	11
	SWMREC 2021	qL.s21.4.1	4	Chr4_81753662	2.31	3.16	3.0	10.5
StemCal	HTRC 2021	qSC.h21.6.1	6	Chr6_69726411	223.94	3.23	2.9	8.2
	SWMREC 2020	qSC.s20.11.1	11	Chr11_40295935	49.46	2.96	3.0	6.9
Vig	SWMREC 2021	qV.s21.1.1	1	Chr1_122391298	149.26	3.12	3 .0	7.8
	HTRC 2020	**qV.h20.2.1**	2	Chr2_90038864	98.11	4.08	3.0	9.3
	HTRC 2021	**qV.h21.2.1**	2	Chr2_88859523	93.20	3.24	3.0	8.0
	HTRC 2021	qV.h21.5.1	5	Chr5_12177923	44.99	3.83	3.0	9.4
	HTRC 2020	**qV.h20.7.1**	7	Chr7_67368224	160.90	4.70	3.0	10.7
	SWMREC 2020	**qV.s20.7.1**	7	Chr7_67368224	160.90	5.15	2.9	11.6
	HTRC 2020	qV.h20.8.1	8	Chr8_69370654	136.20	4.14	3.0	9.5
	SWMREC 2020	qV.s20.8.1	8		111.41	3.61	2.9	8.3
	HTRC 2020	qV.h20.11.1	11	Chr11_93951106	133.74	4.02	3.0	9.2
	SWMREC 2020	qV.s20.11.1	11	Chr11_101141677	122.28	3.25	2.9	7.5
FStage	SWMREC 2020	qFS.s20.1.1	1	Chr1_133317152	175.04	3.86	3.1	8.8
	SWMREC 2020	qFS.s20.3.1	3	Chr3_18803890	60.54	4.37	3.1	9.9
	HTRC 2021	**qFS.h21.3.1**	3	Chr3_45920040	134.38	5.01	3	12.2
	SWMREC 2021	**qFS.s21.3.1**	3	Chr3_45920040	134.81	4.81	2.9	11.8
	HTRC 2021	**qFS.h21.5.1**	5	Chr5_61449999	89.50	3.51	3.0	8.7
	HTRC 2020	qFS.h20.8.1	8	Chr8_32893096	227.34	5.15	3.0	11.6
	SWMREC 2020	qFS.s20.8.1	8	Chr8_32893096	243.34	6.97	3.1	15.4
	HTRC 2021	qFS.h21.8.1	8	Chr8_32893096	304.02	4.07	3.0	10
	SWMREC 2021	qFS.s21.8.1	8	Chr8_76558323	319.02	3.49	2.9	8.7
	HTRC 2021	qFS.h21.11.1	11	Chr11_117450739	185.01	4.51	3.0	11.1
Hght	HTRC 2020	qH.h20.3.1	3	Chr3_22852840	81.37	3.09	2.98	7.2

Figure 4-6: Visualization of QTL on linkage groups 7 and 11 for secondary branching (2br) a) and linkage group 1 for maximum canopy width (MaxWid) (b). QTL names include an acronym for each trait followed by the environment from which data was taken. 'H' and 'S' represents our two locations (HTRC, Holt, Michigan and SWMREC, Benton Harbor, MI, respectively. 20 and 21 represent the years 2020 and 2021.

a)



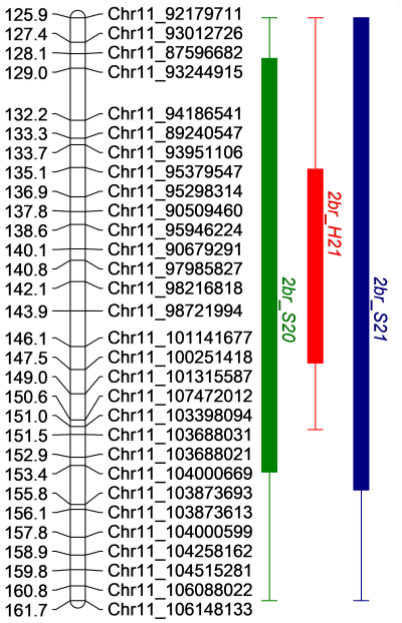


Figure 4-6 (cont'd)

b)

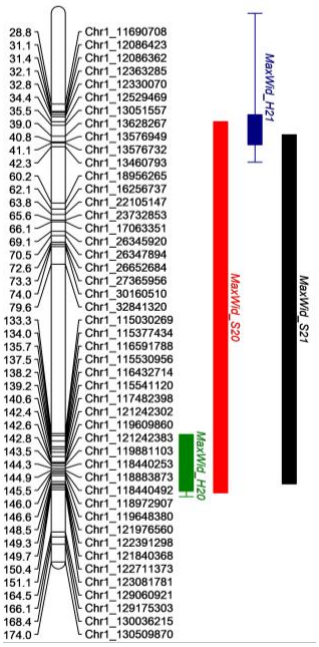


Table 4-8: Pearson correlation coefficients for stevioside (ST), rebaudiosides A, D and M, leaf length and leaf width for stevia biparental MSU18-02 population at two locations (HTRC and SWMREC) over two years (2020 and 2021). This population was also phenotyped for steviol glycosides along with other agronomic traits at two locations (HTRC and SWMREC) and two years (2020 and 2021). ** indicates correlation is significant at a p-value of 0.01 and * indicates significant correlation at a p-value of 0.05.

	ST	**Reb A**	**Reb D**	**Reb M**	**LeafLen**
			HTRC 2021		
Reb A	.404**				
Reb D	0.098	-.351**			
Reb M	-.603**	-.294**	.375**		
LeafLen	0.094	0.143	-0.111	-.215**	
LeafWid	0.121	0.138	-0.013	-.213**	.556**
			SWMREC 2020		
Reb A	0.295**				
Reb D	-0.013	-0.493**			
Reb M	-0.697**	-0.286**	0.342**		
LeafLen	0.103	0.03	-0.035	-.0195**	
LeafWid	0.108	0.014	0.04	-0.214**	0.523**
			HTRC 2021		
Reb A	-0.123				
Reb D	-0.05	-.238**			
Reb M	-.682**	-0.1	.488**		
LeafLen	-0.07	-0.115	0.087	0.088	
LeafWid	-0.114	0.006	.195*	0.159	.458**
			SWMREC 2021		
Reb A	0.042				

Table 4-8 (cont'd)

	ST	Reb A	Reb D	Reb M	LeafLen
Reb D	-0.034	-.430**			
Reb M	-.746**	-.186*	.373**		
LeafLen	0.068	-0.128	0.097	-0.087	
LeafWid	0.053	-0.078	0.08	-0.1	.504**

CHAPTER 5

TRANSCRIPTOMIC ANALYSIS OF STEVIA F_1 LINES WITH CONTRASTING DEVELOPMENT RATES

INTRODUCTION

Stevia rebaudiana (stevia) is a perennial shrub prized for its production of zero-calorie sweeteners called steviol glycosides in the leaves (Sharma et al., 2016). These compounds are incredibly potent, up to 300 times sweeter than sucrose (Gupta et al., 2013) yet they are metabolized safely by the human body without affecting blood sugar levels (Carakostas et al., 2012). Consequently, steviol glycosides are hailed as a healthy alternative to sugar, offering additional benefits such as antioxidant, antihyperglycemic, anti-inflammatory, and anti-cancer properties (Basharat et al., 2021). While native to Paraguay, the majority of commercial stevia production is centered in China (Madan et al., 2010; Ijaz et al., 2015) posing a challenge in meeting global demand for low-calorie foods and beverages. This demand is particularly pressing in countries like the United States, where obesity-related conditions such as diabetes and cardiovascular diseases are prevalent among adults (Moraes et al., 2013; Skinner et al., 2018). As a result, collaborative efforts among researchers aim to establish a sustainable stevia industry within the United States, potentially offering a local solution to a global health challenge.

Stevia thrives in warm climates with well-drained soil and ample sunlight, making regions with such conditions ideal for its cultivation (Ramesh et al., 2006; Libik-Konieczny et al., 2021). Due to challenges with seed germination, stevia is commonly propagated vegetatively through stem cuttings or in vitro methods (Ramesh et al., 2006). During the vegetative stage, above-ground tissue is harvested, and leaves are stripped for glycoside extraction. Since only the above-ground tissue is harvested and the plants are cut while still vegetative, they can continue to grow, allowing multiple harvests during a single growing season. Thus, if the development rate, defined as the rate at which plants produce new nodes or leaves over time before flowering, were

increased, it could allow for more frequent harvests in a growing season, boosting overall yield. Development rate plays a crucial role in determining crop timing and the timing to first yield (Guo et al., 2015; Guo et al., 2017). By breeding for faster-developing varieties, it is possible to achieve an advanced first harvest of stevia. However, to effectively breed for faster development, it is essential to first understand the underlying genetic mechanisms governing development rate.

Plastochron, the time interval between two successive nodes (Lee et al., 2009), is the inverse of development rate. It is controlled by a diverse set of genes encoding various proteins. These include cytochrome P450 (*CYP78A11* and *CYP78A5*) enzymes, RNA-binding protein, MATE cell transporter, glutamate carboxypeptidase, and *N*-acetyltransferase-like protein (Veit et al., 1998; Miyoshi et al., 2004; Kawakatsu et al., 2009; Griffiths et al., 2011; Mimura et al., 2012; Suzuki et al., 2015; Hibara et al., 2021). Additionally, interaction of *miR156/SPL* genes, where *SPL* potentially acts as a leaf-derived signal to suppress the formation of young leaf primordia, as elucidated by Wang et al. (2008). Overexpression of *miR156* in leaf primordia suppresses the function of *SPL* genes, resulting in an accelerated development rate (Wang et al., 2008). Furthermore, phytohormones such as auxins, cytokinin, and gibberellins influence development rate, as discussed previously (Reinhardt et al., 2000; Werner et al., 2003; Mimura et al., 2012). Despite this knowledge, gaps persist in our understanding of development rate regulation due to the pleiotropic effects of genes identified to impact development rate and the diverse nature of encoded products. The evidence suggests that development rate is regulated by a complex mechanism(s), and further research is needed to clarify potential interactions between these pathways and determine whether they are conserved among different species.

The current study was structured to uncover the genetic basis of development rate in stevia by utilizing F_1 lines exhibiting contrasting development rates. The MSU18-02 (10-19 ×10-

RJR) F_1 population was phenotyped for development rate alongside other morphological traits in an open field setting (Chapter 4-4). Subsequently, lines falling within the top and bottom twenty-five percentiles of the development rate spectrum were selected for further assessment in a controlled greenhouse environment. The primary objective of this investigation was to phenotype these selected lines under controlled environmental conditions and analyze the transcriptomes of selected lines to identify genes with differential expression between lines with fast and slow development rates. This approach seeks to deepen our understanding of the genes or gene families involved in regulating development rate in stevia, thus contributing to broader insights into stevia cultivation and breeding efforts.

MATERIALS AND METHODS

Plant materials

For this study, twenty-four lines from the MSU18-02 (10-19 X 10-RJR) F_1 population exhibiting varying development rates (Chapter 4) were chosen. Each line was represented by at least 20 clonally propagated plants, which were transferred to 15.24 cm, 1420.76 cm^3 pots. The plants were arranged in a randomized complete block design with two replications and subjected to long-day conditions (16 hours of light per day) at a temperature of 22°C in two greenhouse compartments. Upon transplantation, the topmost fully expanded leaves on the main stem and two side shoots of each plant were marked with white paint. The number of new nodes produced by each plant was then recorded over a six-week period starting from the date of marking. Data was collected from only eighteen lines, as the remaining six lines had started flowering.

Data analysis

One-way ANOVA and Tukey's Honestly Significant Difference (HSD) post-hoc tests were utilized to examine significant differences in node numbers among the F_1 lines, with a

significance level of α=0.05. These statistical analyses were performed using SPSS v. 27 (IBM; Chicago, IL).

RNA extraction and sequencing

Shoot apex tissue samples of ca. 2 mm length from the tip were harvested and leaf primordia were removed as much as possible with forceps. Samples for each line (within a replication) were pooled, flash frozen with liquid nitrogen, and stored at -80°C. Total RNA extraction was performed using the MagMAX™ Plant RNA Isolation Kit (Catalog #A33784). Subsequently, the RNA samples underwent Quality Control (QC) assessment by running them on TapeStation Analysis Software 3.2 at MSU's Research Technology Support Facility (RTSF) Genomics core facility. Samples with an RNA integrity (RIN) score of $\geq$ 6.0 were selected for sequencing.

Library preparation was carried out using the Illumina Stranded mRNA Prep Ligation kit, incorporating IDT for Illumina TruSeq RNA Unique Dual Indexes, following the manufacturer's protocols. Completed libraries underwent QC and quantification using a combination of Qubit dsDNA HS and Agilent 4200 TapeStation HS DNA1000 assays. The libraries were then pooled in equimolar amounts, and the pool was quantified using the Invitrogen Collibri Quantification qPCR kit. This combined pool was loaded onto one lane of an Illumina S4 flow cell, and sequencing was performed in a 2x150 bp paired-end format using a NovaSeq v1.5, 300 cycle reagent kit. Base calling was conducted using Illumina Real Time Analysis (RTA) v3.4.4, and the output of RTA was demultiplexed and converted to FastQ format using Illumina Bcl2fastq v2.20.0.

Quantification of transcripts

The raw reads were trimmed to remove adaptor sequences and low-quality bases using Trimmomatic version 0.39 (Bolger et al., 2014). Subsequently, the quality of these processed reads was assessed using FastQC (Andrews, 2017). As sequencing for each sample was conducted on two separate lanes, reads from both lanes were merged into a single read. Stevia transcriptome index was constructed using STAR (Dobin et al., 2013), and the reads were aligned to the stevia transcriptome (Vallejo and Warner, 2021) using this index. Samtools feature idxstats was then employed to enumerate the number of reads mapped to each transcript (Danecek et al., 2021).

Identification of differentially expressed genes (transcripts) (DEGs)

Quality control of the samples was conducted through the generation of a heatmap and principal component analysis (PCA) plot. These analyses utilized regularized logarithm (rlog) transformed counts to visualize the relationships between samples based on their gene expression profiles. Any outlier samples detected through these analyses were removed from further analysis to prevent confounding effects. Subsequently, the identification of Differentially Expressed Genes (DEGs) was performed using the DESeq2 package (Love et al., 2014) with default parameters.

Weighted Gene Co-expression Network Analysis (WGCNA)

In the analysis, the WGCNA R package was employed to construct a co-expression network (Pei et al., 2017). Initially, samples were clustered using the 'hclust' function to identify and eliminate any outliers from subsequent analysis, ensuring the robustness of the results. The R function 'pickSoftThreshold' was then utilized to determine the soft threshold power. This was achieved by employing "signed" networks and the "bicor" correlation function to construct the

adjacency matrix. A soft power of 22 was chosen based on a threshold of R^2 fit greater than or equal to 0.85. Next, the Topological Overlap Matrix (TOM) was calculated using the adjacency matrix, and gene dendrograms were generated based on their dissimilarity. Hierarchical clustering and the dynamic tree cut function were applied to detect modules within the coexpression network, with a tree cut height threshold of 0.25 used to cluster the module eigengenes. To relate modules to the development rate trait, gene significance (GS) and module membership (MM) were calculated. Hub genes within each module were identified by applying thresholds of MM >= 0.8 and GS >= 0.8. The corresponding module gene information was then extracted for further analysis, providing valuable insights into the genetic mechanisms underlying the development rate trait in stevia. Venn diagrams were created to visualize the intersection of enriched GO terms between different gene sets using Venny 2.1.0 (Oliveros, 2016).

RESULTS

Selection of slow and fast lines

Of the eighteen evaluated F_1 lines, fourteen, consisting of seven slow and seven fast development rate lines (highlighted in bold), exhibited variation in development rate over a six-week period (Table 5-1). Selecting plants at the extremes of slow and fast development rates proved challenging due to the significant phenotypic variability in development rate. The number of leaves produced by lines 32, 60, and 80 varied between replicates. Notably, lines 30, 71, 75, 165, and 238 consistently produced fewer than 29 leaves, while lines 61, 64, 81, 139, and 171 consistently produced more than 31 leaves at both replications, categorizing them into the slow and fast development rate groups, respectively (Table 5-1). Additionally, lines 103 and 107 were chosen as backup lines in the slow development group, and lines 50 and 58 served as backups in

the fast development group. These backup lines were selected to account for the possibility of losing RNA samples or encountering insufficient RNA quality.

Subsequently, RNA extraction was performed on these fourteen lines, each with two biological replicates. Specifically, five lines were selected from both the fast (denoted with f, e.g. 61f, 64f, 81f, 139f and 171f) and slow (denoted with s, e.g. 30s, 71s, 75s, 165s and 238s) development groups, ensuring that the chosen RNA samples had an RNA integrity (RIN) score of ≥ 6.0 (Table 5-2).

Processing of raw reads

Each sample yielded a minimum of 35 million read pairs, up to a maximum of 65 million read pairs (Table 5-2). Following quality control, at least 99.5% of reads were retained after trimming for adaptor sequences and low-quality reads. The percentage of reads that uniquely mapped to the stevia transcriptome ranged from 70.3% to 76.8%, while mapping to the stevia genome ranged from 39.6% to 71.34% (refer to Table 5-2). Due to better mapping performance, we proceeded with the transcriptome-aligned reads. The number of reads corresponding to each gene in every sample was tallied to identify differentially expressed transcripts.

Differential gene expression analysis

Pooled comparison

After conducting clustering analysis, we identified specific samples (30s, 75s, 238s, 61f, 64f, 81f, and 139f where biological replicates clustered together. Subsequently, principal component analysis (PCA) was performed on these samples (Fig. 5-1A). However, the PCA results indicated that the samples did not distinctly segregate into two clusters based on slow and fast development rates; instead, they exhibited overlapping patterns (Fig. 5-1B). Samples were chosen based on the clustering of biological replicates, leading to the exclusion of sample 61f

from further analysis due to the biological replicates being non-clustered. For the subsequent differential expression analysis, three samples from both the fast (64f, 81f, and 139f) and slow (30s, 75s, and 238s) development rate categories, each with two biological replicates, were selected. In the pooled comparison, 57 transcripts were found to be downregulated, while 114 transcripts were upregulated, in the slow lines compared to fast lines (Fig. 5-2A and 2B). Notably, among the upregulated transcripts in the slow development rate lines were the Cytochrome P450 family gene *CYP78A10* (Locus_2519), *EXPANSIN-LIKE 1* (*EXPL1*) (Locus_18240), *YUC2* (Locus_52753), *CELLULOSE SYNTHASE LIKE D* (Locus_50384), *PPR CONTAINING PROTEIN* (Locus_28383) and F-box transcripts (Locus_32597, Locus_11780 and Locus_33084).

Among the downregulated transcripts were those encoding *PIN-LIKE PROTEIN 3* (Locus_41089) and transcripts related to *FAR1-RELATED_SEQUENCE_5*-like protein (Locus_33270, Locus_33271 and Locus_42243).

Pairwise comparison

Individual comparisons were conducted between each slow and fast line to pinpoint robustly differentially expressed transcripts (DEGs) (Table 5-3). In each comparison, the number of downregulated transcripts was similar to the number of upregulated transcripts. Venn diagrams were utilized to identify core DEGs present in all pairwise comparisons (Fig. 5-3). Core DEGs were defined as transcripts that were consistently differentially expressed in at least six pairwise comparisons, meaning in at least two slow line comparisons to all three fast lines.

Among the core DEGs, 82 downregulated transcripts were identified in the comparisons between slow lines 30s and 75s, while 31 core DEGs were found in the 75s and 238s lines. Interestingly, only one core gene was shared between the 238s and 30s lines compared to all fast

development rate lines (Fig. 5-3A). Additionally, a single core DEG was observed to be upregulated in all slow lines (Fig. 5-3B). This transcript was functionally annotated as one of the *UDP-GLYCOSYLTRANSFERASE* like genes. Among the other upregulated core DEGs, Locus_21363, which encodes a HAD superfamily subfamily IIIB acid phosphatase, was found. A gene from this gene family was previously reported as differentially expressed in petunia genotypes with differing development rates. (Guo et al., 2017). Other core upregulated DEGs (Locus_33084 and Locus_11780) encoded F-box domains with *FBD/LRR-REPEAT PROTEIN* and *F-BOX LIKE PROTEIN*.

Among core downregulated DEGs, transcripts encoding putative *PHYTOHORMONE BINDING PROTEIN-LIKE* (Locus_29612), *FRS5-like* (Locus_42243), *GENERAL TRANSCRIPTION FACTOR 2-RELATED ZINC FINGER PROTEIN* (Locus_40595), and *PUTATIVE LEUCINE-RICH REPEAT DOMAIN L- LIKE PROTEIN* (Locus_28113) were found.

WGCNA

The sample clustering analysis focused on selecting samples with contrasting development rates for further study, while removing outliers. In line with the sample clustering in Figure 5-1A, samples with different development rates did not form two distinct clusters (Fig. 5-4A). For further analysis, samples that grouped with their biological replicates were chosen. Sample 61f was excluded from this analysis because it clustered too far from the other fast development samples. Final samples included three fast development samples (64f, 81f, and 139f) and three slow development samples (30s, 75s, and 238s) for weighted gene co-expression network analysis (Figure 5-B).

After filtering out transcripts with low expression levels or excessive missing data, 31,219 transcripts from a total of 12 samples were retained for further analysis. At a soft power

threshold of 9, 127 co-expression modules were identified. Among these modules, only one, designated "royalblue1", was significantly correlated with the development rate phenotype, showing a negative correlation of -0.85 with this phenotype ($p < 0.01$, Figure 5-5). This indicates that slow development rate has an inverse relationship with the eigengene expression of this module (Figure 5-6), suggesting that as the eigengene expression of this module increases, the likelihood of slow development decreases. The "royalblue1" module contained 47 transcripts, including nine hub transcripts (Table 5-4). One interesting transcript co-expressing in this module was identified as Locus_2494, which corresponds to *MEI2-LIKE PROTEIN 1*, an RNA-binding protein family known to influence development rates in rice and other species (Veit et al., 1998; Miyoshi et al., 2004). The fact that this transcript co-expresses with other transcripts in this module suggests that these genes might share common regulatory mechanisms, participate in similar functions, or respond to the same signaling pathway.

DISCUSSION

The molecular mechanisms underlying leaf initiation have been investigated, particularly concerning the rate of leaf initiation (development rate) and leaf arrangement (phyllotaxy). Existing information highlights the involvement of both biophysical and genetic factors in orchestrating these processes (Reinhardt and Kuhlemeier, 2002; Fleming, 2005; Mimura et al., 2012). However, due to the pleiotropic effects of genes previously implicated in regulation of these traits, significant gaps remain in our understanding of their biochemical activities, interrelationships and conservation of gene functions among species. To address this, our study focused exclusively on unraveling the genetic regulation of development rate in stevia. We aimed to elucidate the specific genetic factors governing this trait and determine if any previously identified genes in other species also regulate development rate in stevia.

Principal component analysis (PCA) analysis revealed that the genotypes selected for fast and slow development rates overlap on the PCA plot, indicating additional natural variation beyond the trait under study. Analysis of gene expression patterns reveals that slow-developing lines exhibit a substantial upregulation of transcripts, with nearly double the number compared to fast-developing lines. This implies that slow development in stevia plants may be attributed to the upregulation of a diverse set of genes, potentially influencing pathways that slow down the rate of leaf emergence or redirect resources to alternative pathways.

The observed differential expression of a *PIN*-like putative auxin efflux carrier family protein (Locus_41089) aligns with previous findings (Reinhardt et al., 2003). *PIN* (pin-formed) proteins are recognized as carriers facilitating the polar transport of auxin, thereby establishing auxin gradients essential for organ initiation (Forestan and Varotto, 2012). Examination of *PIN1* expression suggests that its subcellular polarization leads to localized auxin accumulation at sites of incipient primordia (Adamowski and Friml, 2015), thereby becoming a site for new leaf primordium formation at each plastochron (Reinhardt et al., 2003). The observed downregulation of this transcript in slowly developing plants implies a reduction in polar auxin transport, leading to the creation of auxin minima, consequently delaying leaf initiation. This suggests a prolonged plastochron and a slower development rate. These findings underscore the intertwined genetics governing leaf development rate and arrangement. Notably, genes such as *TE1* and *AMP1* regulate both plastochron and phyllotaxy in maize and *Arabidopsis*, respectively (Veit et al., 1998; Helliwell et al., 2001).

The observed upregulation of a transcript encoding cytochrome P450 protein *CYP78A10*, identified as a homolog of Arabidopsis *KLUH* encoding *CYP78A5/7* (Wang et al., 2008), aligns with previous research findings. Notably, this gene shares orthology with rice *PLA1*, known for

its negative regulation of development rate (Miyoshi et al., 2004). *PLA1* encodes a member of the plant-specific cytochrome P450 monooxygenases subfamily, *CYP78A11*, with expression predominantly observed in young leaf primordia rather than the shoot apical meristem (Miyoshi et al., 2004). Importantly, the study suggests that signals mediated by *PLA1* operate non-cell-autonomously, transmitting from leaf primordia to modulate leaf initiation in the shoot apical meristem (Miyoshi et al., 2004). These findings underscore the conservation of the role of the cytochrome P450 family 78 in regulating development rate, thereby implying a potential similar regulatory mechanism in stevia.

The upregulation of *YUC2*, a key enzyme in the auxin biosynthesis pathway, in the slow development rate lines presents an intriguing observation. *YUC2*, belonging to the flavin-binding monooxygenase family protein, catalyzes the conversion of indole-3-pyruvate (IPA) to indole-3-acetic acid (IAA), a pivotal step in auxin biosynthesis (Dai et al., 2013). This upregulation seemingly contradicts previous evidence suggesting that increased auxin biosynthesis or increased free auxin levels correlate with faster development rates. For instance, mutants of the *Arabidopsis* F-box protein *SLOMO* exhibit reduced auxin levels in the shoot apical meristem (SAM), delaying the formation of an auxin maxima critical for the initiation of subsequent leaf events (Lohmann et al., 2010). Logically, increased auxin biosynthesis, as seen in the upregulation of a biosynthetic gene like *YUC2*, would be expected to accelerate leaf initiation. However, the observation of upregulated auxin biosynthesis genes in slowly developing lines suggests a more nuanced regulatory mechanism. One plausible explanation is that the upregulation of *YUC2* could serve to compensate for the reduced auxin levels in the SAM, thereby maintaining the integrity of leaf initiation events despite the overall slower development

rate. These findings, in conjunction with previous research, underscore the pivotal role of auxin in modulating development rate.

Previous research underscores the role of localized growth modulation, achieved through the modulation of cell wall extensibility, as a critical event in leaf initiation (Pien et al., 2001). Expansins are a family of extracellular proteins that participate in cell wall loosening, consequently altering the physical stress patterns in the meristem. This alteration leads to the acquisition of a new leaf primordium identity as a result of tissue bulging (Fleming et al., 1997). The upregulation of expansin-related genes at the site of a new leaf primordium in tomato further supports this notion (Reinhardt et al., 1998). Additionally, it has been proposed that polar auxin transport activity relies on the cell wall extensibility of expansin proteins (Cosgrove, 2000) or auxin may act as the driver of expansin activity or may regulate expansin-related cells, thereby contributing to leaf initiation events (Kessler and Sinha, 2004). These findings are substantiated by the concurrent upregulation of both auxin biosynthesis and expansin-like transcripts observed in the current study. Additionally, cell wall synthase enzymes such as *CELLULOSE SYNTHASE LIKE D* (*CSLD5*) play a crucial role in maintaining cell proliferation and wall integrity within the SAM (Yang et al., 2016). The increased expression of *CSLD5*, coupled with the upregulation of a gene involved in cell wall loosening, may indicate a compensatory mechanism aimed at regulating SAM cell wall mechanics.

In a prior investigation, a gene from the HAD (haloacid dehalogenase) superfamily subfamily IIIB acid phosphatase family exhibited distinct expression patterns among petunia genotypes characterized by contrasting rates of development (Guo et al., 2017). Although direct evidence linking this gene family to plant development rate is lacking, its observed differential expression in the current study (Locus_21363) provides additional support for its putative role.

This differential expression suggests plausible role in cell signaling pathways, possibly through post-transcriptional dephosphorylation mechanisms (Sanyal et al., 2020). The inferred function of post-transcriptional dephosphorylation by members of the HAD superfamily subfamily IIIB acid phosphatase family may implicate regulatory roles in pivotal signaling molecules or proteins governing plant developmental processes. While further research is needed to elucidate the precise mechanisms by which this gene family impacts development rate, it emerges as a potential candidate gene family for future functional studies.

Several genes belonging to the pentatricopeptide repeat (PPR) containing protein family exhibited differential expression. Notably, a transcript from this family was also identified in a highly correlated module, royalblue1, through WGCNA. This finding is consistent with previous research in petunia, where a gene from the PPR family (Peaxi162Scf01021g00215.1 – PPR superfamily protein) was found to be located near genomic scaffolds harboring SNP markers associated with a development rate QTL (Guo et al., 2017). While QTL regions may contain numerous genes, the proximity of this PPR gene to the identified QTL provides a compelling rationale for further investigation into this gene family. Additionally, our latest study on petunia AE Recombinant Inbred Lines (RILs) with varying development rates corroborated the differential expression of this gene (Chapter 3). The PPR family comprises RNA-binding proteins characterized by repeated RNA motifs, instrumental in RNA binding and metabolism (Barkan and Small, 2014). These proteins are primarily localized to organelles such as mitochondria and chloroplasts (Lurin et al., 2004). They modulate gene expression post-transcriptionally and are implicated in embryogenesis, gametogenesis, and seed development, crucial processes governed by cell division and hormonal signals (Liu et al., 2013). Disruption in the expression of PPR proteins can impair plant metabolism, impacting energy balance and

hormonal signaling pathways (Liu et al., 2010; Barkan and Small, 2014). Mutations in PPR genes have led to lethal or defective embryos and albino seedlings, via disruption of cell proliferation or primary metabolites production in plastids (Tzafrir et al., 2004; Lu et al., 2011; Li et al., 2018). The available evidence suggests that the *PPR* gene family regulates genes involved in reproductive processes, demanding a detailed exploration of the physical and structural changes occurring in the shoot apical meristem preceding flowering initiation. The pleiotropic effects of genes related to the vegetative phase may contribute to the signaling that activates these genes. However, the specific mechanism of action of this gene family in our trait of interest remains elusive and warrants further investigation.

Transcripts related to phytochrome A (phyA) signaling, such as *FAR-RELATED SEQUENCE 5* (Ma and Li, 2018), were found to be differentially expressed. Studies have highlighted their involvement in light signal transduction, photomorphogenesis, circadian clock regulation, flowering time control, shoot meristem maintenance, and floral development (Wang and Deng, 2002; Lin et al., 2007; Li et al., 2011; Li et al., 2016). These findings align with the gene ontology terms identified in our petunia AE RILs exhibiting contrasting development rates, as described in Chapter 3. The enrichment of gene ontology terms associated with flowering initiation suggests the activation of flowering signals, prompting further investigation into the molecular changes occurring in the SAM. Elucidating the intricate interplay of phyA signaling and other regulatory pathways in the SAM is crucial for understanding the mechanisms governing the transition to flowering and the broader regulation of processes related to plant development rate.

Additionally, the observed differential expression of transcripts associated with the putative F-box domain containing leucine-rich repeat regions, alongside the upregulation of

CDPK (calcium-dependent protein kinase) involved in phytohormone signaling, suggests that protein degradation and calcium-dependent signaling cascades (Xu and Huang, 2017; Matsushima et al., 2019) might play a crucial role in regulating development rate.

WGCNA was utilized to identify a group of co-expressed genes potentially associated with the regulation of development rate. Although the correlation between the identified module and the phenotype was robust (0.85), the precise relationship between all genes within the module and development rate control remain challenging to interpret. This difficulty likely stemmed from the presence of noise generated by other pathways, influenced by the inherent variability among genotypes. Significantly, a transcript belonging to the *MEI2-like 1* protein family (Locus_46362), recognized as an ortholog of *Arabidopsis* AML clade gene (AML1), was discovered within the royalblue 1 module. The MEI2-like gene family encodes RNA-binding protein characterized by a highly conserved RNA-binding motif that was initially identified in the MEI2 gene of the fission yeast *S. pombe* (Hirayama et al., 1997). Previous studies have identified associated members of this gene family, such as *PLA2* and *TE1*, with the negative regulation of development rate in rice and maize (Veit et al., 1998; Mimura et al., 2012). AML1, falling into the AML14 clade, one of two sister clades within the *MEI2-like* gene family (Kaur et al., 2006), further supports the significance of our results outlined in Chapter 2, where a potential candidate gene from petunia also belonged to the AML14 clade (refer to Figure 2-9). These findings underscore the importance of investigating the AML14 clade for its potential role in regulating development rate.

CONCLUSION

In summary, our results indicate that leaf initiation is a complex process orchestrated by several factors, starting in the SAM through auxin polar transport, cell wall mechanics, and

communication signals from emerging leaf primordia to the SAM. Auxin, in particular, plays a pivotal role in this process, encompassing its biosynthesis and signaling pathways. Furthermore, we found evidence suggesting that the regulation of development rate in stevia is likely influenced by the cytochrome P450 subfamily 78 (*CYP78A)*, potentially through the metabolism of an unidentified substrate. Our identification of a differentially expressed transcript belonging to an AML clade underscores the need for further investigation into the potential role of this clade in leaf initiation and development rate control. Overall, the regulation of development rate involves a complex interplay of genetic, endogenous (hormonal and signaling peptides), and environmental factors. Understanding these intricate mechanisms is essential for elucidating the molecular basis of leaf development rate in plants like stevia.

Tables & Figures

Table 5-1: Average leaf number and standard deviation of F_1 genotypes collected after six weeks of marking the leaves in both replications. N represents number of plants on which data was collected for each genotype. Blank cells indicate that we could not record the leaf number data of certain genotypes as they were already flowering. Genotypes marked bold were selected for shoot apex tissue collection and RNA extraction under slow and fast development rate categories.

Genotype	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem	Mean ± S.D (N) Side branch	Mean ± S.D (N) Main stem
	Rep 1		Rep 2	
30	25.07 ± 2.1 (15)	28.31 ± 3.4 (13)	25.25 ± 3.7 (16)	28.33 ± 4.9 (12)
32	26.82 ± 6.2 (17)	32.25 ± 4.1 (4)	32.38 ± 3.5 (13)	32.33 ± 3.4 (6)
50	31.46 ± 3.7 (13)		32.24 ± 3.8 (21)	32.33 ± 5.7 (3)
58	30.00 ± 5.1 (11)	33.75 ± 5.9 (11)	33.33 ± 5.3 (18)	34.08 ± 5.4 (12)
60	30.46 ± 3.4 (13)		31.86 ± 6.5 (14)	29.67 ± 3.2 (6)
61	31.09 ± 6.3 (11)	36.40 ± 4.0 (5)	35.33 ± 5.8 (12)	40.43 ± 4.6 (7)
64	34.13 ± 2.3 (8)	36.33 ± 5.9 (15)	33.31 ± 6.1 (13)	33.90 ± 8.0 (10)
71	28.00 ± 5.2 (5)		27.20 ± 6.4 (5)	26.00 ± 3.6 (4)
75	23.50 ± 4.1 (8)		23.83 ± 4.5 (12)	24.63 ± 3.2 (8)
80	30.25 ± 8.4 (4)		33.67 ± 9.2 (6)	31.50 ± 6.4 (4)
81	32.17 ± 2.4 (6)	35.20 ± 3.6 (5)	33.29 ± 5.3 (7)	38.40 ± 3.3 (5)
103	28.20 ± 5.4 (10)	28.71 ± 6.1 (7)	27.31 ± 4.7 (13)	31.67 ± 4.6 (6)
107	28.50 ± 3.9 (10)	32.29 ± 2.4 (7)	27.92 ± 4.6 (12)	32.50 ± 1.5 (6)
139	34.35 ± 6.6 (17)	40.70 ± 5.2 (10)	32.68 ± 6.8 (19)	36.50 ± 7.5 (12)
165	21.20 ± 3.0 (5)		24.63 ± 3.2 (8)	29.33 ± 3.1 (3)
171	31.71 ± 3.9 (7)	34.75 ± 1.7 (4)	34.71 ± 4.1 (7)	
219	29.33 ± 3.1 (3)	40.40 ± 5.2 (5)		
238	25.33 ± 3.2 (9)	24.83 ± 5.0 (6)	28.88 ± 6.2 (8)	31.00 ± 6.3 (5)

Table 5-2: Summary of twenty RNA samples and their biological replicates including the RNA integrity number (RIN), lane information, number of raw reads generated, number of reads after merging the two lanes, final number of reads that survived trimming, percentage of reads uniquely mapped to the *Stevia rebaudiana* genome and Stevia transcriptome.

F_1 genotype	Biological Replicate	RIN score	Lane	Number of raw read pairs	Number of merged read pairs	Final number of read pairs after trimming (%)	Percent of uniquely mapped read pairs to the genome (%)	Percent of uniquely mapped read pairs to the transcriptome (%)
61f	1	7.3	1	10,574,631	41031869	40815717	67.55	71.33
			2	30,457,238		(99.47%)		
	2	8.1	1	16,517,612	65009431	64793036	39.66	75.59
			2	48,491,819		(99.67%)		
64f	1	7.3	1	8,744,782	34860527	34730355	67.33	71.23
			2	26,115,745		(99.63%)		
	2	7.6	1	12,063,850	46564894	46358515	68.62	71.50
			2	34,501,044		(99.56%)		
81f	1	7.2	1	10,729,327	42152834	41952732	67.48	70.98
			2	31,423,507		(99.53%)		
	2	7.7	1	10,329,115	40186180	39984073	67.95	72.79
			2	29,857,065		(99.50%)		

Table 5-2 (cont'd)

F_1 genotype	Biological Replicate	RIN score	Lane	Number of raw read pairs	Number of merged read pairs	Final number of read pairs after trimming (%)	Percent of uniquely mapped read pairs to the genome (%)	Percent of uniquely mapped read pairs to the transcriptome (%)
139f	1	7.1	1	11,683,019	44606161	44403710	67.18	71.51
			2	32,923,142		(99.55%)		
	2	7.9	1	12,618,943	48764149	48596877	69.6	73.56
			2	36,145,206		(99.66%)		
171f	1	7.5	1	11,925,397	46654498	46495489	69.93	72.64
			2	34,729,101		(99.66%)		
	2	9.2	1	11,694,108	44682079	44521028	69.94	73.05
			2	32,987,971		(99.64%)		
30s	1	7.2	1	12,834,101	48413335	48221756	68	71.57
			2	35,579,234		(99.60%)		
	2	8.0	1	11,398,749	43857584	43699250	69.06	72.44
			2	32,458,835		(99.64%)		
71s	1	7.6	1	10,990,446	42488550	42305578	68.07	70.35
						(99.57%)		

Table 5-2 (cont'd)

F_1 genotype	Biological Replicate	RIN score	Lane	Number of raw read pairs	Number of merged read pairs	Final number of read pairs after trimming (%)	Percent of uniquely mapped read pairs to the genome (%)	Percent of uniquely mapped read pairs to the transcriptome (%)
			2	31,498,104				
	2	7.9	1	12,505,036	47565928	47411705	70.40	74.12
			2	35,060,892		(99.68%)		
75s	1	9.2	1	12,887,895	49261849	49070087	68.23	70.90
			2	36,373,954		(99.61%)		
	2	7.3	1	11,509,199	45270477	44995816	68.04	71.16
			2	33,761,278		(99.39%)		
165s	1	8.6	1	13,502,595	53700177	53516286	68.63	71.49
			2	40,197,582		(99.66%)		
	2	7.6	1	13,482,300	50105277	49911293	71.34	76.83
			2	36,622,977		(99.61%)		
238s	1	7.9	1	11,114,876	42820728	42674709	69.06	72.54
						(99.66%)		

Table 5-2 (cont'd)

F_1 genotype	Biological Replicate	RIN score	Lane	Number of raw read pairs	Number of merged read pairs	Final number of read pairs after trimming (%)	Percent of uniquely mapped read pairs to the genome (%)	Percent of uniquely mapped read pairs to the transcriptome (%)
			2	31,705,852				
	2	7.7	1	10,779,307	41390735	41190954	68.29	72.35
			2	30,611,428		(99.55%)		

Figure 5-1: Quality control of samples based on gene expression profiles by using regularized logarithm normalization counts of samples in a heatmap (A) and on a principal component analysis plot (B). The x-axis represents the PC1 and the percentage of variance explained and y-axis represents PC2 and the percentage of variance explained.

A)

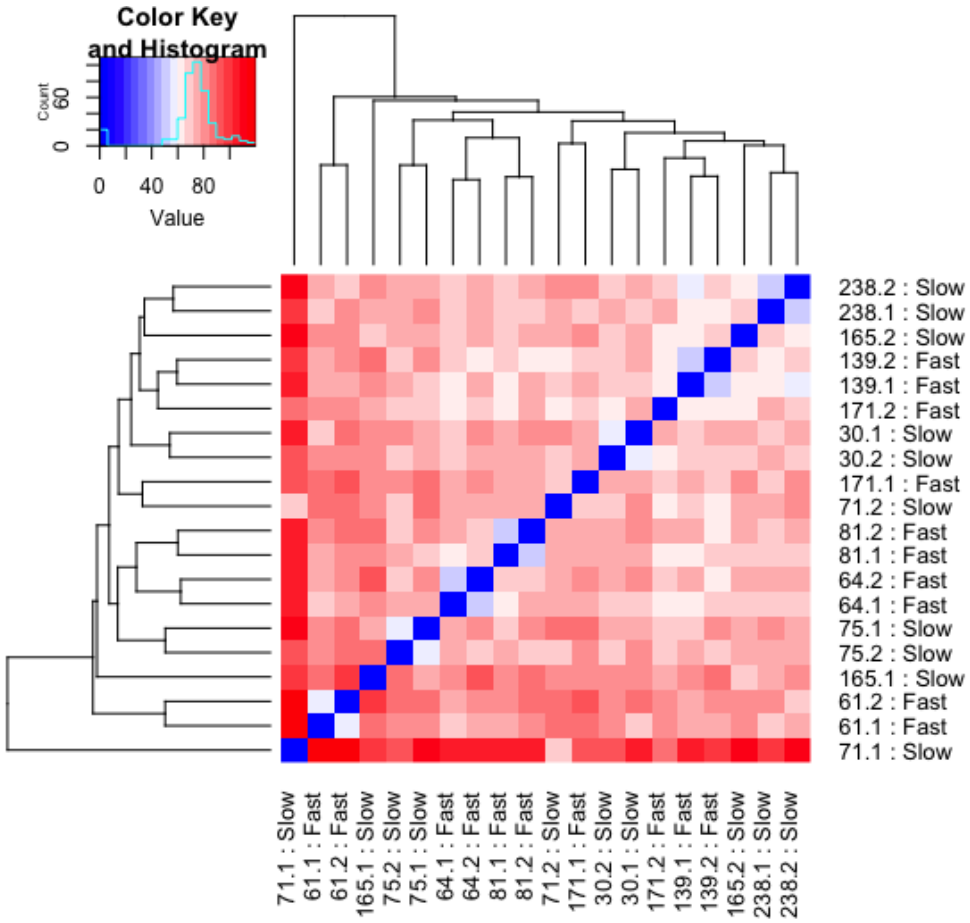


Figure 5-1 (cont'd)

B)

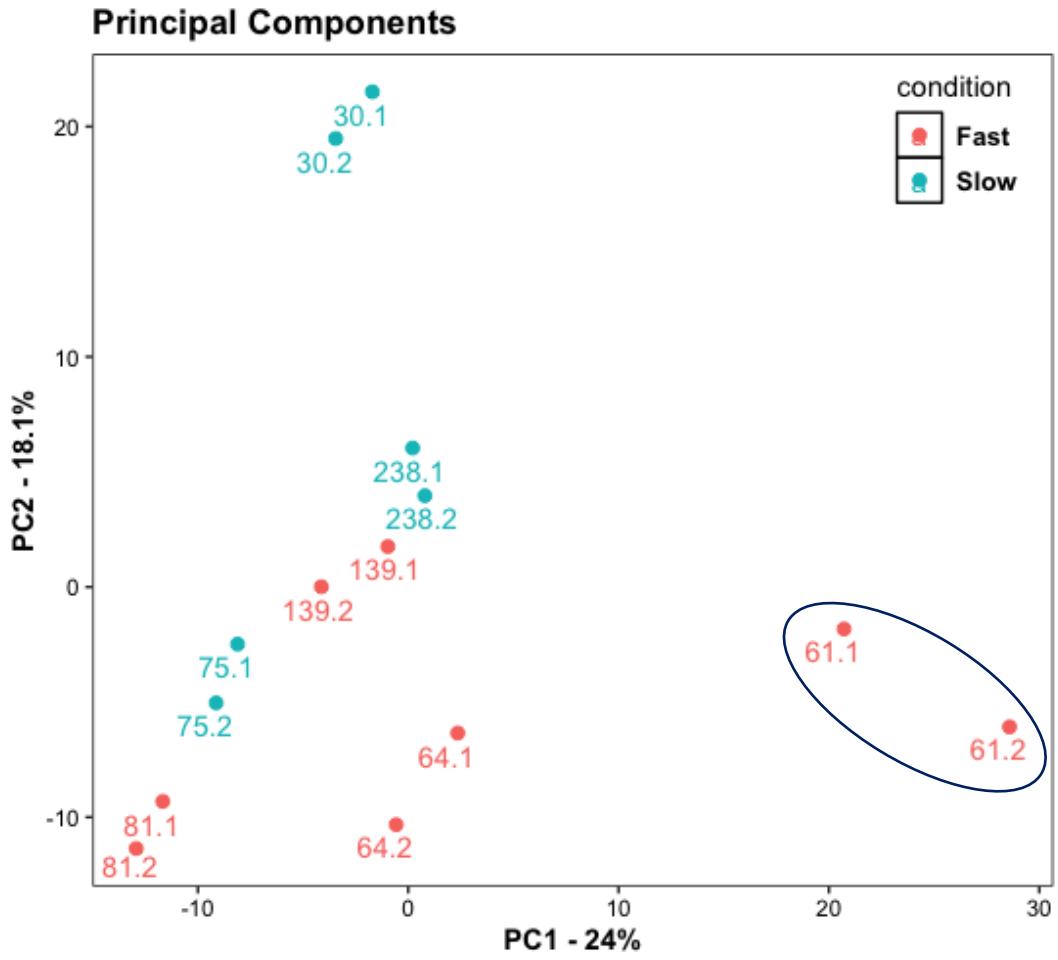
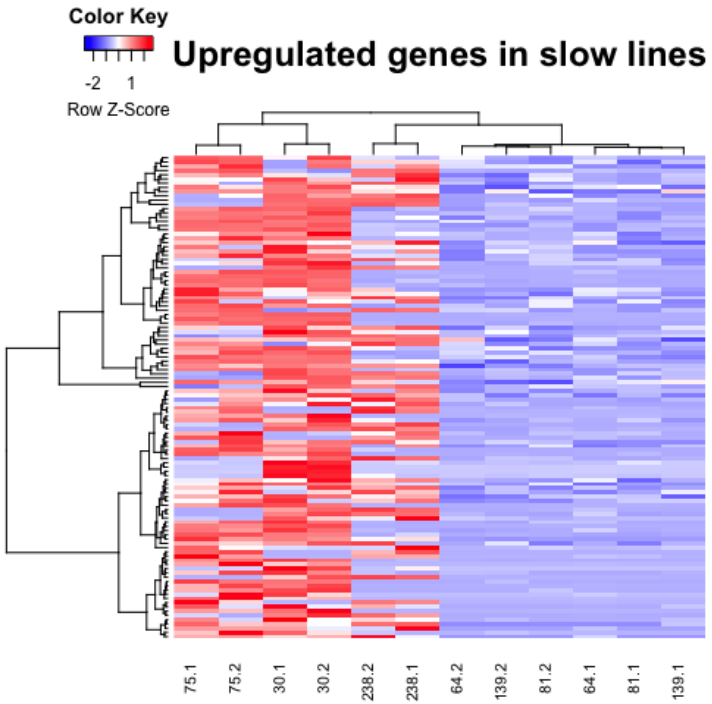


Figure 5-2: Heatmaps of transcripts upregulating (A) and downregulating (B) in the slow lines as compared to the fast lines. Y-axis represents regularized logarithm normalization counts of genes in each of the samples represented on X-axis.

A)



B)

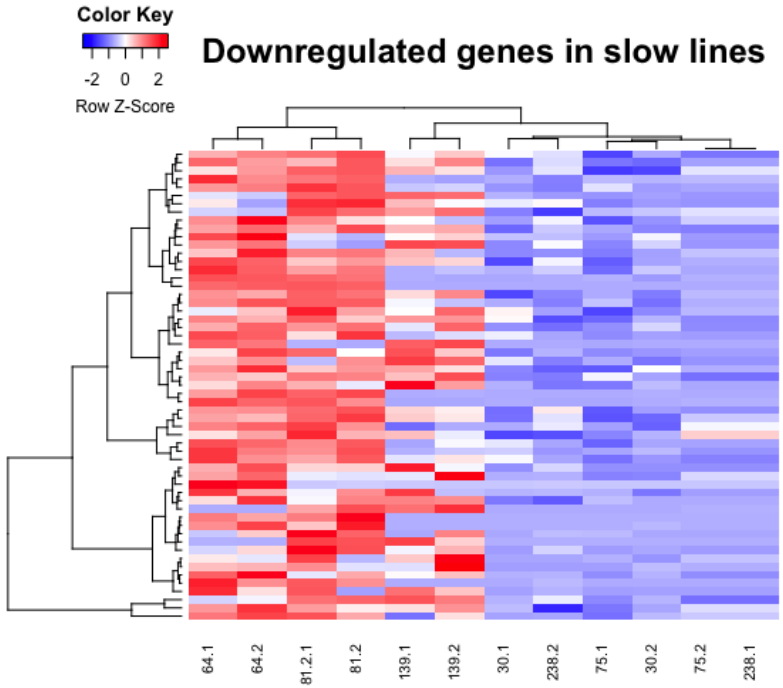


Table 5-3: Pairwise comparisons of differentially expressed transcripts. D and U indicates number of down-regulated and up-regulated transcripts, respectively, in each of the comparisons.

Slow/Fast	64f		81f		139f	
	D	U	D	U	D	U
30s	754 (48%)	822 (52%)	866 (46%)	1001 (54%)	751 (58%)	534 (42%)
75s	886 (54%)	744 (46%)	859 (52%)	778 (48%)	1041 (64%)	597 (36%)
238s	664 (43%)	890 (57%)	503 (40%)	745 (60%)	432 (54%)	364 (46%)

Figure 5-3: Venn diagrams representing individual pairwise comparisons of differentially expressed transcripts. First three Venn diagrams are comparisons of down-regulated (A) and up-regulated transcripts. (B) between each slow line with all three fast lines and the fourth diagram draws comparisons between results of first three comparisons.

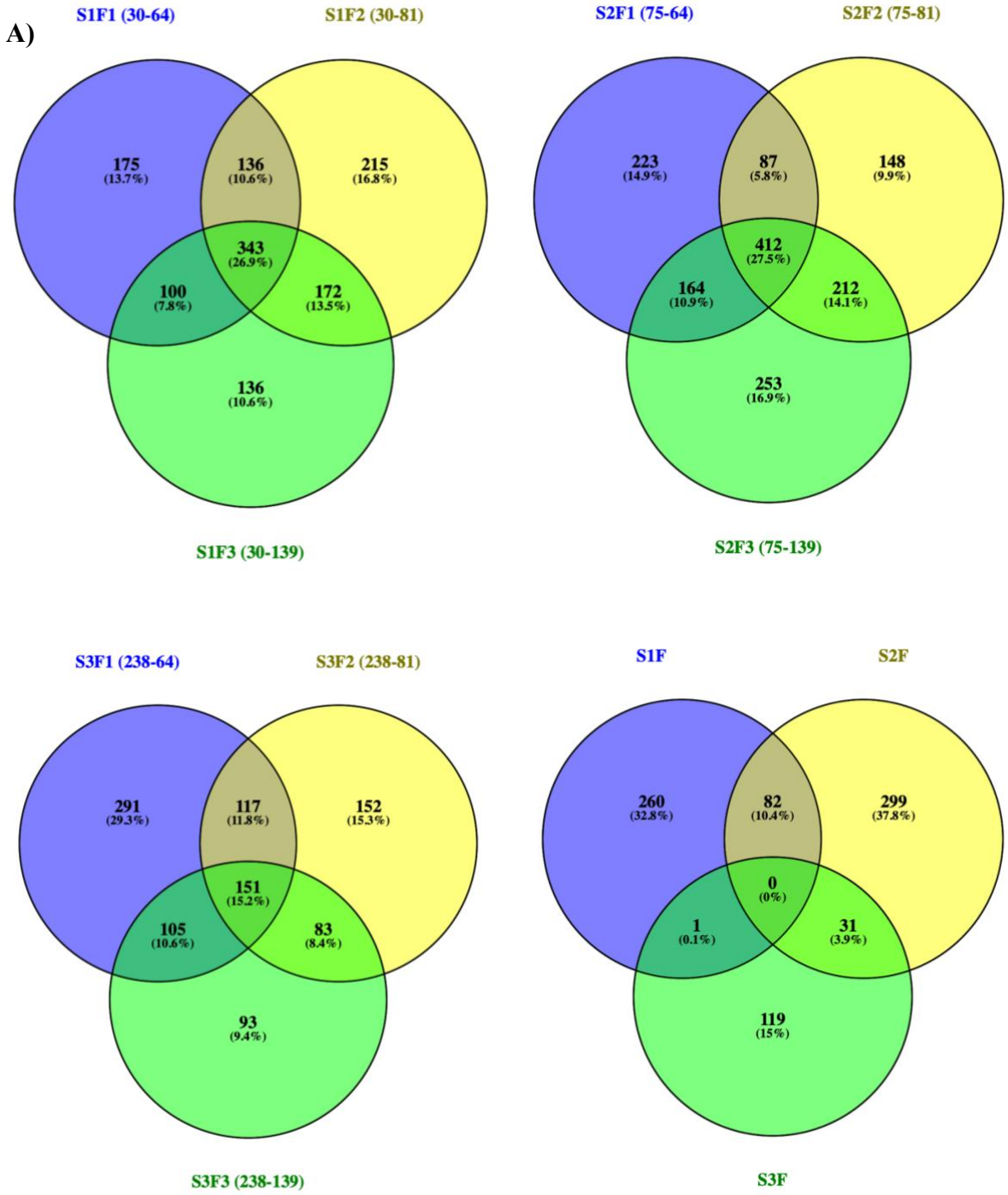


Figure 5-3 (cont'd)

B)

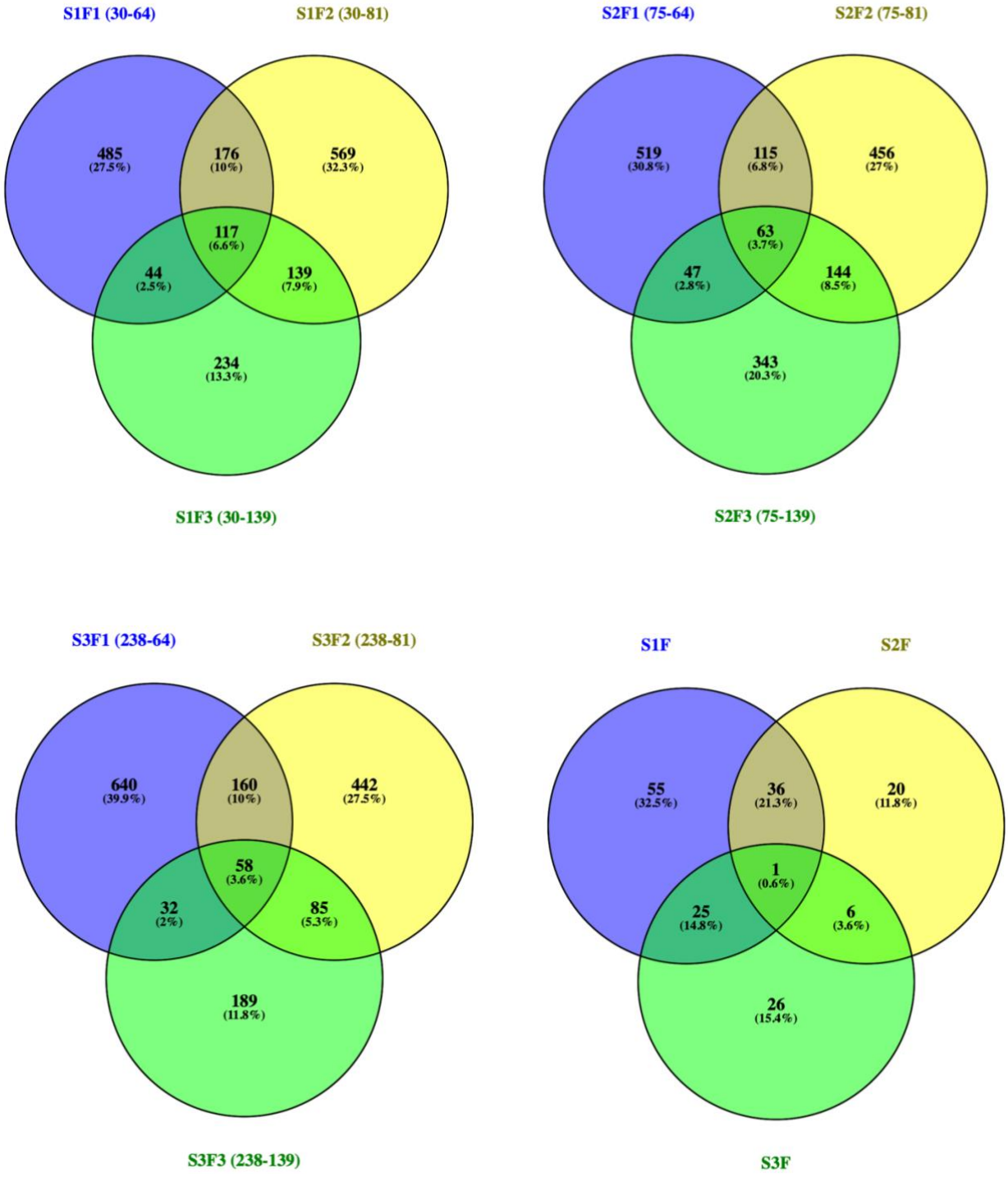


Table 5-4: Summary of differentially expressed transcripts identified in the pooled comparison of slow vs fast development rate lines (LFC ≥ 1.5 and α =0.05) . Positive and negative values of log2Fold change indicate that the transcripts are upregulated and downregulated, respectively, in the slow lines as compared to the fast lines. Only the annotated genes are listed in the table.

Transcript ID	Functional description	Log2foldchange
Locus_21363_Transcript_1/1_Confidence_1.000_Length_542	**|_Symbols:_|_HAD_superfamily,_subfamily_IIIB_acid_phosphatase_|_chr5:17712433-17714046_FORWARD_LENGTH=272_AT5G44020.1**	3.8478
Locus_14053_Transcript_1/1_Confidence_1.000_Length_1332	|_Symbols:_|_NAD(P)-linked_oxidoreductase_superfamily_protein_|_chr1:22071410-22073067_REVERSE_LENGTH=326_AT1G59960.1	2.3480
Locus_24596_Transcript_1/1_Confidence_1.000_Length_1555	|_Symbols:_|_PLC-like_phosphodiesterases_superfamily_protein_|_chr5:2575152-2576770_REVERSE_LENGTH=372_AT5G08030.1	1.6173
Locus_23473_Transcript_2/2_Confidence_0.750_Length_2101	|_Symbols:_|_Transmembrane_amino_acid_transporter_family_protein_|_chr2:17167561-17170145_REVERSE_LENGTH=536_AT2G41190.1	1.6205
Locus_26073_Transcript_1/1_Confidence_1.000_Length_1651	|_Symbols:_|_Transmembrane_amino_acid_transporter_family_protein_|_chr4:17935533-17936843_FORWARD_LENGTH=436_AT4G38250.1	2.7352
Locus_10697_Transcript_1/1_Confidence_1.000_Length_426	|_Symbols:_|_unknown_protein_BEST_Arabidopsis_thaliana_protein_match_is:_unknown_protein_(TAIR:AT5G01970.1)_Has_246_Blast_hits_to_244_proteins_in_61_species:_Archae_-_0_Bacteria_-_8_Metazoa_-_78_Fungi_-_10_Plants_-_117_Viruses_-_0_Other_Eukaryotes_-_33_(source:_NCBI_BLink)._|_chr1:10543177-10544418_FORWARD_LENGTH=389_AT1G30050.1	2.8999

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_38764_Transcript_1/1_Confidence_1.000_Length_1703	|_Symbols:_ATCDPK1,_CPK10,_CDPK1,_AtCPK10_|_calcium-dependent_protein_kinase_1_|_chr1:6523468-6525736_REVERSE_LENGTH=545_AT1G18890.1	1.7637
Locus_17315_Transcript_1/2_Confidence_0.667_Length_1616	|_Symbols:_ATCEL2,_CEL2_|_cellulase_2_|_chr1:613386-616103_REVERSE_LENGTH=501_AT1G02800.1	3.2695
Locus_18240_Transcript_1/1_Confidence_1.000_Length_828	**|_Symbols:_ATEXLA1,_EXPL1,_ATEXPL1,_ATHEXP_BETA_2.1,_EXLA1_|_expansin-like_A1_|_chr3:16896238-16897189_FORWARD_LENGTH=265_AT3G45970.1**	1.5918
Locus_20009_Transcript_1/1_Confidence_1.000_Length_1091	|_Symbols:_AtHSD5,_HSD5_|_hydroxysteroid_dehydrogenase_5_|_chr4:6268363-6270179_FORWARD_LENGTH=389_AT4G10020.1	1.9269
Locus_27883_Transcript_1/1_Confidence_1.000_Length_213	|_Symbols:_ATPAP29,_PAP29_|_purple_acid_phosphatase_29_|_chr5:25328237-25329616_FORWARD_LENGTH=389_AT5G63140.1	2.7988
Locus_43193_Transcript_1/1_Confidence_1.000_Length_1271	|_Symbols:_ATSPO11-1_|_Spo11/DNA_topoisomerase_VI,_subunit_A_protein_|_chr3:4231560-4234192_REVERSE_LENGTH=362_AT3G13170.1	1.5077
Locus_2519_Transcript_5/5_Confidence_0.667_Length_1790	**|_Symbols:_CYP78A10_|_cytochrome_P450,_family_78,_subfamily_A,_polypeptide_10_|_chr1:27866667-27868368_REVERSE_LENGTH=537_AT1G74110.1**	2.1893

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_52753_Transcript_1/1_Confidence_1.000_Length_216	**|_Symbols:_YUC2_|_Flavin-binding_monooxygenase_family_protein_|_chr4:7721840-7723616_REVERSE_LENGTH=415_AT4G13260.1**	4.3116
Locus_52753_Transcript_1/1_Confidence_1.000_Length_216	**|_Symbols:_YUC2_|_Flavin-binding_monooxygenase_family_protein_|_chr4:7721840-7723616_REVERSE_LENGTH=415_AT4G13260.1**	4.3116
Locus_20121_Transcript_2/2_Confidence_0.667_Length_551	Barwin-like_endoglucanase_[Artemisia_annua]_&gt_gb|PWA96312.1|_Barwin-like_endoglucanase_[Artemisia_annua]_PWA47986	6.4189
Locus_19379_Transcript_1/1_Confidence_1.000_Length_1318	Glycosyl_hydrolase_family_100_[Artemisia_annua]_PWA59137	2.5538
Locus_50384_Transcript_1/1_Confidence_1.000_Length_205	**cellulose synthase like_protein_D5_[Lactuca_sativa]_&gt_gb|PLY89076.1|_hypothetical_protein_LSAT_9X27001_[Lactuca_sativa]_XP_023759090**	7.5023
Locus_11780_Transcript_1/1_Confidence_1.000_Length_1268	**F-box_protein_At5g07610-like_[Helianthus_annuus]_&gt_gb|OTG05796.1|_hypothetical_protein_HannXRQ_Chr12g0377621_[Helianthus_annuus]_XP_021996051**	2.3391
Locus_30118_Transcript_1/1_Confidence_1.000_Length_629	heat_shock_70_kDa_protein_18-like_isoform_X2_[Cynara_cardunculus_var._scolymus]_XP_024981473	8.2274

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_47762_Transcript_1/1_Confidence_1.000_Length_222	hypothetical_protein_E3N88_23486_[Mikania_micrantha]_KAD4585885	1.7247
Locus_44072_Transcript_1/1_Confidence_1.000_Length_252	hypothetical_protein_E3N88_44369_[Mikania_micrantha]_KAD0371268	5.4773
Locus_48349_Transcript_1/1_Confidence_1.000_Length_292	hypothetical_protein_E3N88_45324_[Mikania_micrantha]_KAC9735356	6.1228
Locus_11266_Transcript_1/1_Confidence_1.000_Length_465	hypothetical_protein_LSAT_7X88080_[Lactuca_sativa]_PLY63425	2.0246
Locus_49632_Transcript_1/1_Confidence_1.000_Length_580	hypothetical_protein_LSAT_9X34960_[Lactuca_sativa]_PLY73047	6.1875
Locus_21578_Transcript_2/2_Confidence_0.800_Length_741	LEAF_RUST_10_DISEASE-RESISTANCE_LOCUS_RECEPTOR-LIKE_PROTEIN_KINASE-like_2.1_isoform_X3_[Lactuca_sativa]_XP_023756050	4.5235
Locus_53753_Transcript_1/1_Confidence_1.000_Length_270	immediate_early_response_3-interacting_protein_1-like_[Helianthus_annuus]_&gt_ref|XP_021973880.1|_immediate_early_response_3-interacting_protein_1-like_[Helianthus_annuus]_XP_021973878	1.6363
Locus_7732_Transcript_3/8_Confidence_0.583_Length_1372	PAZ_domain-containing_protein_[Artemisia_annua]_PWA66503	1.5456
Locus_28383_Transcript_1/1_Confidence_1.000_Length_282	**Pentatricopeptide_repeat-containing_protein_[Artemisia_annua]_PWA38618**	1.5498

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_77_Transcript_1/1_Confidence_1.000_Length_509	phylloplanin-like_[Helianthus_annuus]_XP_022029707	1.6521
Locus_15669_Transcript_1/2_Confidence_0.667_Length_800	phylloplanin-like_isoform_X2_[Helianthus_annuus]_XP_022029705	1.6565
Locus_40853_Transcript_1/1_Confidence_1.000_Length_372	probable_disease_resistance_protein_At5g66900_[Helianthus_annuus]_XP_022012802	1.6984
Locus_19579_Transcript_3/3_Confidence_0.778_Length_3869	proteasome_activator_subunit_4-like_[Helianthus_annuus]_&gt_gb|OTG31764.1|_putative_proteasome_activating_protein_[Helianthus_annuus]_XP_022028770	1.7375
Locus_3307_Transcript_3/3_Confidence_0.750_Length_1575	protein_kinase-like_domain-containing_protein_[Artemisia_annua]_PWA88905	2.0272
Locus_28314_Transcript_2/3_Confidence_0.714_Length_794	protein_NLP3-like_isoform_X2_[Helianthus_annuus]_&gt_gb|OTF85241.1|_hypothetical_protein_HannXRQ_Chr17g0537831_[Helianthus_annuus]_XP_022025386	2.2288
Locus_20080_Transcript_1/1_Confidence_1.000_Length_1136	protein_STRICTOSIDINE_SYNTHASE-LIKE_6-like_[Helianthus_annuus]_&gt_gb|OTG23966.1|_putative_strictosidine_synthase_[Helianthus_annuus]_XP_022037031	2.2684
Locus_40881_Transcript_1/1_Confidence_1.000_Length_433	putative_bulb-type_lectin_domain,_Thaumatin_[Helianthus_annuus]_OTG31983	2.2909

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_32597_Transcript_1/1_Confidence_1.000_Length_325	**putative_F-box_domain,_Leucine-rich_repeat_domain,_L_domain-like_protein_[Helianthus_annuus]_OTF85159**	2.3680
Locus_33084_Transcript_1/1_Confidence_1.000_Length_1805	**putative_F-box/FBD/LRR-repeat_protein_At4g13965_[Helianthus_annuus]_XP_022023416**	2.4412
Locus_51430_Transcript_1/1_Confidence_1.000_Length_231	putative_germin-like_protein_2-1_[Helianthus_annuus]_&gt_gb|OTG15939.1|_putative_germin,_RmlC-like_cupin_domain_protein_[Helianthus_annuus]_XP_021978932	2.7419
Locus_23135_Transcript_2/2_Confidence_0.667_Length_1075	putative_isoprenoid_synthase_domain-containing_protein_[Helianthus_annuus]_OTG07247	2.8896
Locus_56584_Transcript_1/1_Confidence_1.000_Length_219	putative_PGG_domain-containing_protein_[Helianthus_annuus]_OTG30422	2.9092
Locus_10843_Transcript_1/1_Confidence_1.000_Length_1130	putative_protein_kinase-like_domain,_Concanavalin_A-like_lectin/glucanase_domain_protein_[Helianthus_annuus]_OTG17372	2.9925
Locus_47898_Transcript_1/1_Confidence_1.000_Length_302	putative_spo11/DNA_topoisomerase_VI_subunit_A,_Heavy_metal-associated_domain,_HMA_[Helianthus_annuus]_OTG26727	3.0356
Locus_13762_Transcript_1/1_Confidence_1.000_Length_733	putative_ubiquitin_[Helianthus_annuus]_OTG34721	3.6878

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_28600_Transcript_1/1_Confidence_1.000_Length_1246	**putative_zinc_finger,_CCHC-type_[Helianthus_annuus]_OTG34013**	4.1550
Locus_14198_Transcript_1/1_Confidence_1.000_Length_460	RecName:_Full=2S_seed_storage_protein__AltName:_Full=2S_albumin_storage_protein__Flags:_Precursor_[Helianthus_annuus]_&gt_emb|CAA29699.1|_HaG5_protein_[Helianthus_annuus]_P15461	4.1987
Locus_14348_Transcript_1/1_Confidence_1.000_Length_320	ubiquitin_carboxyl-terminal_hydrolase_22-like_[Cynara_cardunculus_var._scolymus]_XP_024966816	4.2396
Locus_47420_Transcript_1/1_Confidence_1.000_Length_1760	UDP-glycosyltransferase_84B1-like_[Helianthus_annuus]_XP_021985591	4.2600
Locus_23909_Transcript_1/1_Confidence_1.000_Length_616	uncharacterized_protein_LOC110866443_[Helianthus_annuus]_XP_021971282	4.4133
Locus_42111_Transcript_1/1_Confidence_1.000_Length_246	uncharacterized_protein_LOC110871979_isoform_X3_[Helianthus_annuus]_XP_021976431	4.4608
Locus_34412_Transcript_1/1_Confidence_1.000_Length_611	uncharacterized_protein_LOC110882522_[Helianthus_annuus]_&gt_ref|XP_021986896.1|_uncharacterized_protein_LOC110883463_[Helianthus_annuus]_XP_021986212	5.0361
Locus_27532_Transcript_1/1_Confidence_1.000_Length_894	uncharacterized_protein_LOC110886255_[Helianthus_annuus]_&gt_gb|OTG12461.1|_putative_ulp1_protease_family,_C-terminal_catalytic_domain-containing_protein_[Helianthus_annuus]_XP_021989724	5.0380

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_33515_Transcript_1/1_Confidence_1.000_Length_1330	uncharacterized_protein_LOC110907935_isoform_X1_[Helianthus_annuus]_XP_022008540	5.2078
Locus_38817_Transcript_1/1_Confidence_1.000_Length_220	uncharacterized_protein_LOC110902245_[Helianthus_annuus]_XP_022004644	5.0755
Locus_21382_Transcript_2/3_Confidence_0.700_Length_360	uncharacterized_protein_LOC110911583_[Helianthus_annuus]_XP_022011896	5.2413
Locus_40940_Transcript_1/1_Confidence_1.000_Length_231	uncharacterized_protein_LOC111891235_[Lactuca_sativa]_XP_023743072	5.7445
Locus_51810_Transcript_1/1_Confidence_1.000_Length_317	UniRef100_B9SCJ0_Xyloglucan_endotransglucosylase/hydrolase_protein_2,_putative_n=1_Tax=Ricinus_communis_RepID=B9SCJ0_RICCO_4756433	5.8174
Locus_8028_Transcript_1/1_Confidence_1.000_Length_1020	UniRef100_C6TBS7_Putative_uncharacterized_protein_n=1_Tax=Glycine_max_RepID=C6TBS7_SOYBN_4857141	5.9515
Locus_6852_Transcript_1/1_Confidence_1.000_Length_1778	UniRef100_F6H021_Putative_uncharacterized_protein_n=1_Tax=Vitis_vinifera_RepID=F6H021_VITVI_5117528	5.9980
Locus_7881_Transcript_1/1_Confidence_1.000_Length_1555	UniRef100_F6H064_Putative_uncharacterized_protein_n=1_Tax=Vitis_vinifera_RepID=F6H064_VITVI_5117569	6.1610
Locus_21514_Transcript_1/1_Confidence_1.000_Length_975	UniRef100_F6H5H6_Putative_uncharacterized_protein_n=1_Tax=Vitis_vinifera_RepID=F6H5H6_VITVI_5119328	6.2466
Locus_51498_Transcript_1/1_Confidence_1.000_Length_265	**UniRef100_F8S1H8_Cytochrome_P450_n=1_Tax=Helianthus_annuus_RepID=F8S1H8_HELAN_5134371**	6.3407

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_4983_Transcript_1/1_Confidence_1.000_Length_218	UniRef100_G7LEY4_Auxilin-like_protein_n=1_Tax=Medicago_truncatula_RepID=G7LEY4_MEDTR_5210753	6.4058
Locus_30838_Transcript_2/2_Confidence_0.727_Length_541	UniRef100_Q6Y0Z7_RGC2-like_protein_(Fragment)_n=1_Tax=Helianthus_annuus_RepID=Q6Y0Z7_HELAN_5543929	6.5441
Locus_17846_Transcript_1/1_Confidence_1.000_Length_1512	vinorine_synthase-like_[Lactuca_sativa]_&gt_gb|PLY78746.1|_hypothetical_protein_LSAT_9X45161_[Lactuca_sativa]_XP_023772659	6.6833
Locus_21409_Transcript_2/3_Confidence_0.667_Length_909	zinc_finger_BED_domain-containing_protein_RICESLEEPER_2-like_[Helianthus_annuus]_XP_021995786	6.7179
Locus_30431_Transcript_1/1_Confidence_1.000_Length_399	|_Symbols:_|_GDSL-like_Lipase/Acylhydrolase_superfamily_protein_|_chr1:10044603-10046379_REVERSE_LENGTH=390_AT1G28580.1	-1.6098
Locus_40595_Transcript_1/1_Confidence_1.000_Length_304	**|_Symbols:_|_General_transcription_factor_2-related_zinc_finger_protein_|_chr3:11593924-11595441_REVERSE_LENGTH=505_AT3G29763.1**	-2.4695
Locus_50799_Transcript_1/1_Confidence_1.000_Length_277	|_Symbols:_|_Plant_protein_1589_of_unknown_function_|_chr3:20473876-20474705_REVERSE_LENGTH=95_AT3G55240.1	-1.6467

Table 5-4 (cont'd)

Transcript ID	Functional description	Log2foldchange
Locus_27393_Transcript_1/1_Confidence_1.000_Length_279	ankyrin_repeat-containing_domain,_PGG_domain_protein_[Artemisia_annua]_PWA40442	-1.5130
Locus_33463_Transcript_1/1_Confidence_1.000_Length_210	|_Symbols:_CYP72A14_|_cytochrome_P450,_family_72,_subfamily_A,_polypeptide_14_|_chr3:4934478-4936462_FORWARD_LENGTH=512_AT3G14680.1	-1.5720
Locus_55032_Transcript_1/1_Confidence_1.000_Length_219	ankyrin_repeat-containing_protein_ITN1-like_[Helianthus_annuus]_XP_021980816	-3.4371
Locus_9395_Transcript_9/10_Confidence_0.156_Length_2191	CALMODULIN-BINDING_PROTEIN60_[Artemisia_annua]_PWA94216	-2.0728
Locus_40373_Transcript_1/1_Confidence_1.000_Length_239	hypothetical_protein_C1H46_032516_[Malus_baccata]_TQD81913	-4.1233
Locus_46394_Transcript_1/1_Confidence_1.000_Length_269	hypothetical_protein_E3N88_34488_[Mikania_micrantha]_KAD3066608	-2.4523
Locus_44884_Transcript_1/1_Confidence_1.000_Length_655	hypothetical_protein_E3N88_40340_[Mikania_micrantha]_KAD2393363	-1.6421
Locus_2225_Transcript_1/1_Confidence_1.000_Length_1153	nodulin-related_protein_1-like_[Helianthus_annuus]_&gt_gb|OTG12295.1|_putative_protein_involved_in_response_to_salt_stress_[Helianthus_annuus]_XP_021989583	-1.6812
Locus_33271_Transcript_1/1_Confidence_1.000_Length_603	protein_FAR1-RELATED_SEQUENCE_5-like_[Helianthus_annuus]_XP_021973751	-3.1070

Table 5-4 (cont'd)

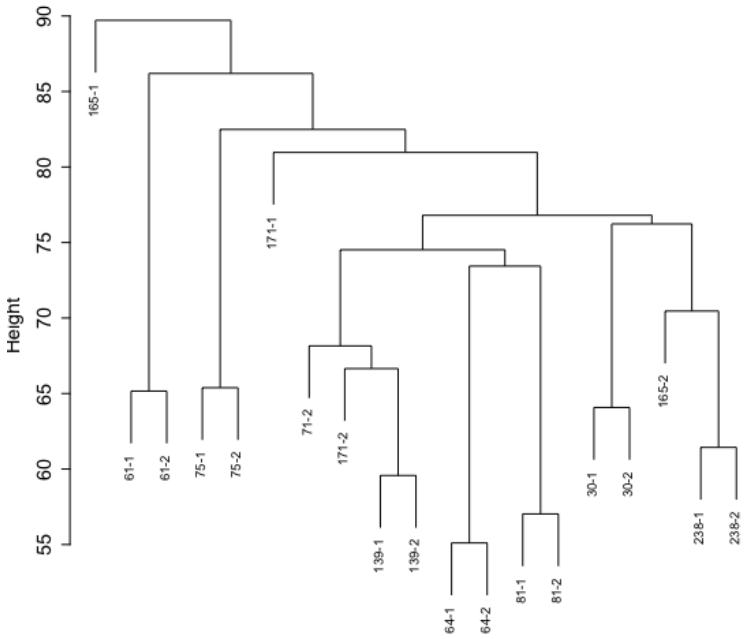
Transcript ID	Functional description	Log2foldchange
Locus_33271_Transcript_1/1_Confidence_1.000_Length_603	protein_FAR1-RELATED_SEQUENCE_5-like_[Helianthus_annuus]_XP_021973751	-3.1070
Locus_6353_Transcript_1/1_Confidence_1.000_Length_599	probable_glutathione_S-transferase_isoform_X1_[Helianthus_annuus]_XP_021988109	-4.2768
Locus_33270_Transcript_1/1_Confidence_1.000_Length_574	protein_FAR1-RELATED_SEQUENCE_5-like_[Helianthus_annuus]_XP_021971663	-3.2480
Locus_42243_Transcript_1/1_Confidence_1.000_Length_214	protein_FAR1-RELATED_SEQUENCE_5-like_[Helianthus_annuus]_XP_022030249	-3.1375
Locus_41089_Transcript_1/1_Confidence_1.000_Length_269	**protein_PIN-LIKES_3-like_[Helianthus_annuus]_&gt_gb|OTF98280.1|_putative_auxin_efflux_carrier_family_protein_[Helianthus_annuus]_XP_022009931**	-2.4274
Locus_43619_Transcript_1/1_Confidence_1.000_Length_204	putative_AMP-dependent_synthetase/ligase_[Helianthus_annuus]_OTG20787	-4.2277
Locus_41902_Transcript_1/1_Confidence_1.000_Length_597	putative_RNA-directed_DNA_polymerase,_eukaryota_[Helianthus_annuus]_OTG18874	-1.6221
Locus_50854_Transcript_1/1_Confidence_1.000_Length_265	putative_RNA-directed_DNA_polymerase,_eukaryota_[Helianthus_annuus]_OTG24384	-5.2285
Locus_4587_Transcript_3/3_Confidence_0.625_Length_764	uncharacterized_protein_LOC110889668_isoform_X2_[Helianthus_annuus]_XP_021992924	-2.5417

Table 5-4 (cont'd)

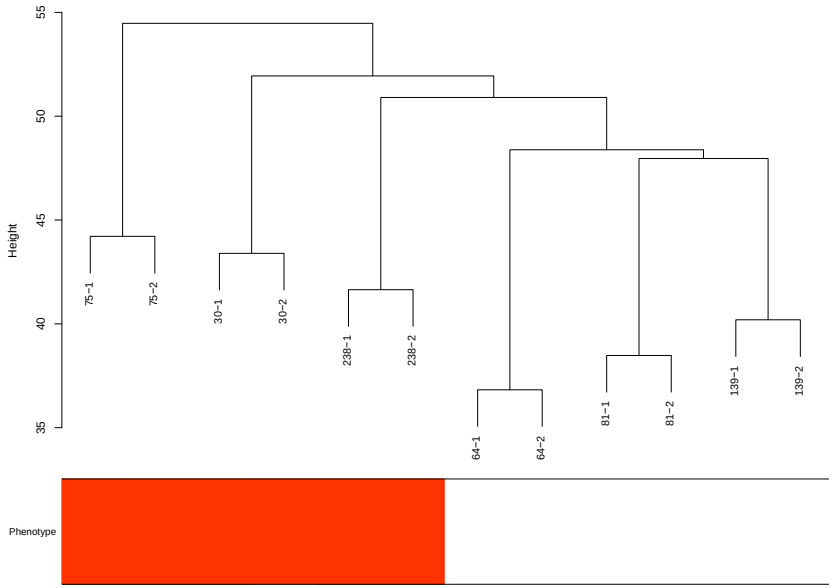
Transcript ID	Functional description	Log2foldchange
Locus_20620_Transcript_1/1_Confidence_1.000_Length_701	uncharacterized_protein_LOC110937409_[Helianthus_annuus]_XP_022035518	-3.5405
Locus_50636_Transcript_1/1_Confidence_1.000_Length_275	uncharacterized_protein_LOC110890546_[Helianthus_annuus]_XP_021993858	-5.5652
Locus_37706_Transcript_1/1_Confidence_1.000_Length_246	UniRef100_C6ZLB7_NBS-LRR_resistance-like_protein_RGC260_n=3_Tax=Helianthus_annuus_RepID=C6ZLB7_HELAN_4861248	-1.8742
Locus_8515_Transcript_1/1_Confidence_1.000_Length_226	UniRef100_F6I0B3_Putative_uncharacterized_protein_n=1_Tax=Vitis_vinifera_RepID=F6I0B3_VITVI_5129176	-1.6925
Locus_27976_Transcript_1/1_Confidence_1.000_Length_1553	**UniRef100_Q309D0_P450_mono-oxygenase_n=1_Tax=Stevia_rebaudiana_RepID=Q309D0_STERE_5487509**	-2.3982
Locus_3755_Transcript_7/7_Confidence_0.500_Length_1714	UniRef100_Q6VAB0_UDP-glycosyltransferase_85C2_n=1_Tax=Stevia_rebaudiana_RepID=Q6VAB0_STERE_5542241	-1.5443

Figure 5-4: Sample clustering dendrogram of all samples (A) and only 12 samples used for the WGCNA analysis (B).

A)



B)



*Samples falling under the red spectrum of phenotype are slow development rate lines.

Figure 5-5: Heatmap displaying the modules correlated with the development rate phenotype. Numbers inside each box indicate Pearson correlation coefficient between the module and the phenotype and a p-value in bracket. Red color indicates positive correlation whereas blue color indicates a negative correlation.

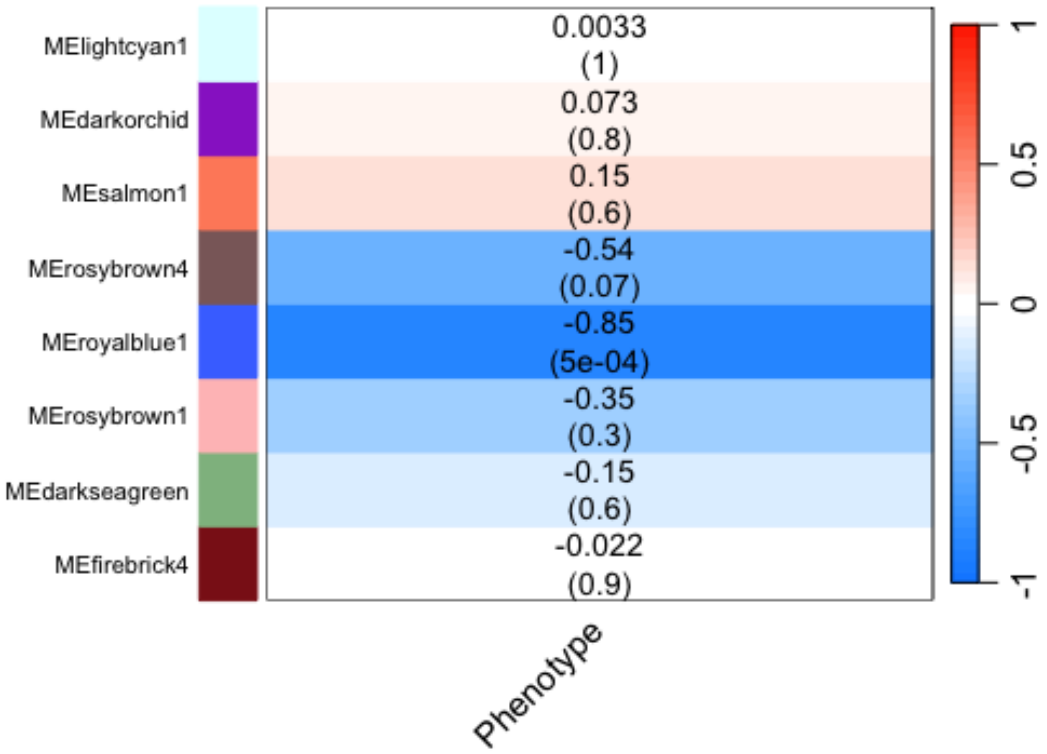


Figure 5-6: Eigenvalues of the royalblue1 module in all samples. Samples 30s, 75s and 238s are the slow development rate lines and 64f, 81f and 139f are the fast development rate lines. Each sample has two biological replicates. Bar plots represent eigenvalues of royalblue1 module individually in all samples.

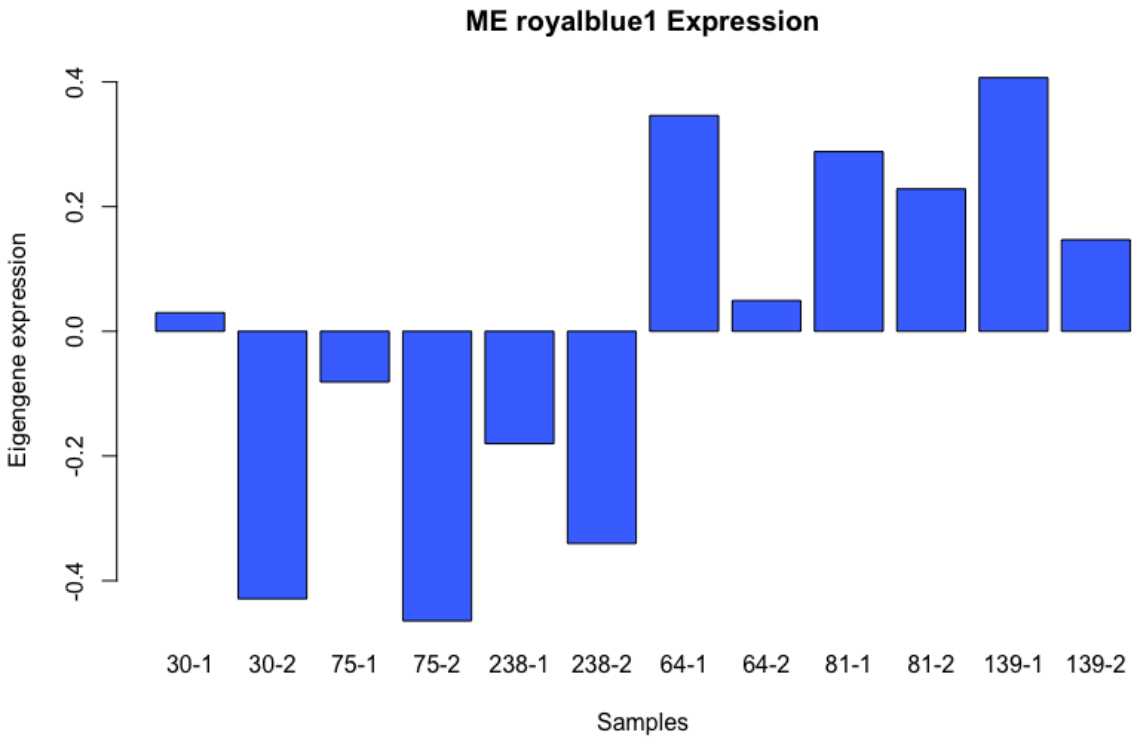


Table 5-5: Functional description of 47 transcripts in module Royalblue1. Hub genes (module membership and gene significance greater than 0.8) for this module are highlighted in bold.

Transcript ID	Functional description
Locus_179_Transcript_1/1_Confidence_1.000_Length_1051	|_Symbols:_|_Protein_of_unknown_function_(DUF579)_|_chr5:26819019-26819972_FORWARD_LENGTH=317_AT5G67210.1
Locus_596_Transcript_2/2_Confidence_0.750_Length_2117	|_Symbols:_emb2004_|_RNI-like_superfamily_protein_|_chr1:3461771-3465590_FORWARD_LENGTH=605_AT1G10510.1
Locus_1168_Transcript_2/2_Confidence_0.667_Length_706	|_Symbols:_|_Bacterial_sec-independent_translocation_protein_mttA/Hcf106_|_chr5:10784142-10785677_REVERSE_LENGTH=147_AT5G28750.1
Locus_1340_Transcript_3/3_Confidence_0.571_Length_1646	|_Symbols:_|_Sulfite_exporter_TauE/SafE_family_protein_|_chr2:10977174-10979677_FORWARD_LENGTH=476_AT2G25737.1
Locus_1603_Transcript_1/1_Confidence_1.000_Length_1694	**|_Symbols:_ENO1_|_enolase_1_|_chr1:27839465-27841901_REVERSE_LENGTH=477_AT1G74030.1**
Locus_1838_Transcript_4/4_Confidence_0.556_Length_1418	|_Symbols:_OASC,_ATCS-C_|_O-acetylserine_(thiol)_lyase_isoform_C_|_chr3:22072668-22075345_REVERSE_LENGTH=430_AT3G59760.3
Locus_1988_Transcript_4/4_Confidence_0.375_Length_1634	UniRef100_Q6VAB0_UDP-glycosyltransferase_85C2_n=1_Tax=Stevia_rebaudiana_RepID=Q6VAB0_STERE_5542241
Locus_2494_Transcript_1/2_Confidence_0.857_Length_2237	protein_MEI2-like_1_[Helianthus_annuus]_&gt_ref|XP_021984220.1|_protein_MEI2-like_1_[Helianthus_annuus]_&gt_gb|OTG16663.1|_putative_RNA_recognition_motif_2,_Nucleotide-binding_alpha-beta_plait_domain_protein_[Helianthus_annuus]_XP_021984219

Table 5-5 (cont'd)

Transcript ID	Functional description
Locus_3314_Transcript_1/1_Confidence_1.000_Length_1121	uncharacterized_protein_LOC110885730_[Helianthus_annuus]_&gt_gb|OTG11807.1|_putative_UBA-like_protein_[Helianthus_annuus]_XP_021989131
Locus_3690_Transcript_1/2_Confidence_0.750_Length_1082	|_Symbols:_|_INVOLVED_IN:_biological_process_unknown_LOCATED_IN:_endomembrane_system_EXPRESSED_IN:_22_plant_structures_EXPRESSED_DURING:_13_growth_stages_CONTAINS_InterPro_DOMAIN/s:_Mannose-6-phosphate_receptor,_binding_(InterPro:IPR009011),_Glucosidase_II_beta_subunit-like_(InterPro:IPR012913)_Has_30201_Blast_hits_to_17322_proteins_in_780_species:_Archae_-_12_Bacteria_-_1396_Metazoa_-_17338_Fungi_-_3422_Plants_-_5037_Viruses_-_0_Other_Eukaryotes_-_2996_(source:_NCBI_BLink)._|_chr5:13354552-13356725_REVERSE_LENGTH=282_AT5G35080.1
Locus_4490_Transcript_1/1_Confidence_1.000_Length_1057	|_Symbols:_|_Ribosomal_protein_L21_|_chr1:13208777-13210246_FORWARD_LENGTH=220_AT1G35680.1
Locus_5097_Transcript_3/3_Confidence_0.714_Length_1148	|_Symbols:_ADK1_|_adenylate_kinase_1_|_chr5:25393274-25394817_REVERSE_LENGTH=246_AT5G63400.1
Locus_5976_Transcript_4/5_Confidence_0.182_Length_1793	|_Symbols:_|_Aldolase-type_TIM_barrel_family_protein_|_chr5:4302080-4304212_REVERSE_LENGTH=438_AT5G13420.1
Locus_6029_Transcript_1/2_Confidence_0.750_Length_827	|_Symbols:_HYD1_|_C-8,7_sterol_isomerase_|_chr1:6949160-6950135_FORWARD_LENGTH=223_AT1G20050.1
Locus_6275_Transcript_2/2_Confidence_0.667_Length_1041	**|_Symbols:_CLPP4,_NCLPP4_|_CLP_protease_P4_|_chr5:18396351-18397586_FORWARD_LENGTH=292_AT5G45390.1**

Table 5-5 (cont'd)

Transcript ID	Functional description
Locus_6391_Transcript_1/1_Confidence_1.000_Length_829	|_Symbols:_|_Ribosomal_protein_L17_family_protein_|_chr3:20067672-20068385_REVERSE_LENGTH=211_AT3G54210.1
Locus_7502_Transcript_2/2_Confidence_0.833_Length_2135	|_Symbols:_ATIMD2,_IMD2_|_isopropylmalate_dehydrogenase_2_|_chr1:30287833-30290126_FORWARD_LENGTH=405_AT1G80560.1
Locus_8431_Transcript_1/1_Confidence_1.000_Length_689	**|_Symbols:_|_CONTAINS_InterPro_DOMAIN/s:_Ribosomal_protein_L53,_mitochondrial_(InterPro:IPR019716)_Has_50_Blast_hits_to_50_proteins_in_19_species:_Archae_-_0_Bacteria_-_0_Metazoa_-_6_Fungi_-_0_Plants_-_42_Viruses_-_0_Other_Eukaryotes_-_2_(source:_NCBI_BLink)._|_chr5:15854188-15854771_REVERSE_LENGTH=127_AT5G39600.1**
Locus_8522_Transcript_2/2_Confidence_0.750_Length_1332	**|_Symbols:_RSZ22a,_At-RSZ22a_|_RNA_recognition_motif_and_CCHC-type_zinc_finger_domains_containing_protein_|_chr2:10449837-10450860_FORWARD_LENGTH=196_AT2G24590.1**
Locus_9055_Transcript_1/2_Confidence_0.800_Length_2067	NAC_domain-containing_protein_53-like_[Helianthus_annuus]_&gt_gb|OTF99367.1|_putative_NAC_domain_containing_protein_53_[Helianthus_annuus]_XP_022006101
Locus_9661_Transcript_4/4_Confidence_0.500_Length_1620	**|_Symbols:_TBL39_|_TRICHOME_BIREFRINGENCE-LIKE_39_|_chr2:17717498-17719921_REVERSE_LENGTH=367_AT2G42570.1**
Locus_10701_Transcript_1/1_Confidence_1.000_Length_677	PREDICTED:_probable_CCR4-associated_factor_1_homolog_6_isoform_X1_[Gossypium_hirsutum]_XP_016738488
Locus_10748_Transcript_1/1_Confidence_1.000_Length_1665	**|_Symbols:_|_Ypt/Rab-GAP_domain_of_gyp1p_superfamily_protein_|_chr2:13086147-13088991_REVERSE_LENGTH=440_AT2G30710.1**

Table 5-5 (cont'd)

Transcript ID	Functional description
Locus_11433_Transcript_3/3_Confidence_0.714_Length_1930	|_Symbols:ATLCB1,_LCB1,_EMB2779,_FBR11_|_long-chain_base1_|_chr4:17218598-17221124_FORWARD_LENGTH=482_AT4G36480.2
Locus_11858_Transcript_2/2_Confidence_0.750_Length_418	transketolase_family_protein_[Artemisia_annua]_PWA81515
Locus_11872_Transcript_1/1_Confidence_1.000_Length_906	**|_Symbols:_PANC,_PTS,_ATPTS_|_homolog_of_bacterial_PANC_|_chr5:19803823-19805041_REVERSE_LENGTH=310_AT5G48840.1**
Locus_12634_Transcript_1/1_Confidence_1.000_Length_933	|_Symbols:_|_unknown_protein_FUNCTIONS_IN:_molecular_function_unknown_INVOLVED_IN:_biological_process_unknown_LOCATED_IN:_chloroplast_thylakoid_membrane,_chloroplast_EXPRESSED_IN:_22_plant_structures_EXPRESSED_DURING:_13_growth_stages_Has_42_Blast_hits_to_42_proteins_in_19_species:_Archae_-_0_Bacteria_-_0_Metazoa_-_0_Fungi_-_0_Plants_-_40_Viruses_-_0_Other_Eukaryotes_-_2_(source:_NCBI_BLink)._|_chr3:19109118-19109842_FORWARD_LENGTH=181_AT3G51510.1
Locus_13598_Transcript_1/1_Confidence_1.000_Length_1164	|_Symbols:_LysoPL2_|_lysophospholipase_2_|_chr1:19651378-19652576_FORWARD_LENGTH=332_AT1G52760.1
Locus_15440_Transcript_3/4_Confidence_0.667_Length_1685	protein_indeterminate-domain_9-like_[Cynara_cardunculus_var._scolymus]_XP_024983804
Locus_16062_Transcript_2/2_Confidence_0.750_Length_1369	UniRef100_D8WUJ0_WRKY_transcription_factor_n=1_Tax=Artemisia_annua_RepID=D8WUJ0_ARTAN_5024140
Locus_16381_Transcript_1/2_Confidence_0.750_Length_339	|_Symbols:_|_Ankyrin_repeat_family_protein_|_chr2:1036192-1037536_REVERSE_LENGTH=240_AT2G03430.1

Table 5-5 (cont'd)

Transcript ID	Functional description
Locus_16799_Transcript_3/3_Confidence_0.667_Length_1258	|_Symbols:_|_phytanoyl-CoA_dioxygenase_(PhyH)_family_protein_|_chr2:221316-223187_FORWARD_LENGTH=283_AT2G01490.1
Locus_18115_Transcript_1/2_Confidence_0.667_Length_954	**|_Symbols:_|_HSP20-like_chaperones_superfamily_protein_|_chr5:23725936-23727528_REVERSE_LENGTH=158_AT5G58740.1**
Locus_21076_Transcript_1/1_Confidence_1.000_Length_234	hypothetical_protein_HannXRQ_Chr17g0560531_[Helianthus_annuus]_OTF87324
Locus_21780_Transcript_1/1_Confidence_1.000_Length_428	**Armadillo_[Artemisia_annua]_PWA78229**
Locus_22670_Transcript_1/1_Confidence_1.000_Length_504	protein_NUCLEAR_FUSION_DEFECTIVE_6,_chloroplastic/mitochondrial-like_[Cynara_cardunculus_var._scolymus]_XP_024995105
Locus_22833_Transcript_1/3_Confidence_0.714_Length_826	Armadillo-like_helical_[Cynara_cardunculus_var._scolymus]_KVH93747
Locus_28230_Transcript_3/3_Confidence_0.571_Length_544	putative_mitogen-activated_protein_(MAP)_kinase_kinase_kinase_Ste11,_Cryptococcus_[Helianthus_annuus]_OTF97831
Locus_30719_Transcript_3/3_Confidence_0.667_Length_852	|_Symbols:_|_Pyridoxal_phosphate_phosphatase-related_protein_|_chr4:14496164-14497310_FORWARD_LENGTH=245_AT4G29530.1
Locus_36329_Transcript_1/1_Confidence_1.000_Length_202	F-box/FBD/LRR-repeat_protein_At1g13570-like_[Helianthus_annuus]_XP_022020378

Table 5-5 (cont'd)

Transcript ID	Functional description
Locus_37894_Transcript_1/1_Confidence_1.000_Length_284	UniRef100_B9N178_Predicted_protein_n=1_Tax=Populus_trichocarpa_RepID=B9N178_POPTR_4733137
Locus_44863_Transcript_1/2_Confidence_0.667_Length_254	transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980057.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980058.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980059.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980060.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980061.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_&gt_ref|XP_024980062.1|_transcription_initiation_factor_TFIID_subunit_6_isoform_X1_[Cynara_cardunculus_var._scolymus]_XP_024980056
Locus_45167_Transcript_1/1_Confidence_1.000_Length_231	putative_quinoprotein_alcohol_dehydrogenase like superfamily_[Helianthus_annuus]_OTG27617
Locus_46362_Transcript_1/1_Confidence_1.000_Length_332	|_Symbols:_TPPF_|_Haloacid_dehalogenase like_hydrolase_(HAD)_superfamily_protein_|_chr4:7365480-7367346_REVERSE_LENGTH=368_AT4G12430.1
Locus_48727_Transcript_1/1_Confidence_1.000_Length_231	putative_pentatricopeptide_repeat_protein_[Helianthus_annuus]_OTG02960

BIBLIOGRAPHY

Abdulameer DA, Osman MB, Sulaiman Z, Yusop MR, Abdullah S, Azizi P, Muttaleb QA (2018) Assessment of *Stevia rebaudiana* Bertoni genotypes via morpho-agronomic traits under two light conditions. American Journal of Plant Sciences 9: 1403

Adamowski M, Friml J (2015) PIN-dependent auxin transport: action, regulation, and evolution. The Plant Cell 27: 20-32

Adams S, Hadley P, Pearson S (1998) The effects of temperature, photoperiod, and photosynthetic photon flux on the time to flowering of petunia 'Express Blush Pink'. Journal of the American Society for Horticultural Science 123: 577-580

Ahmad J, Khan I, Blundell R, Azzopardi J, Mahomoodally MF (2020) *Stevia rebaudiana* Bertoni.: an updated review of its health benefits, industrial applications and safety. Trends in Food Science & Technology 100: 177-189

Al-Taweel S, Azzam C, Khaled K, Abdel-Aziz R (2021) Improvement of stevia (*Stevia rebaudiana* Bertoni) and steviol glycoside through traditional breeding and biotechnological approaches. Sabrao J. Breed. Genet 53: 88-111

Anderson GH, Alvarez ND, Gilman C, Jeffares DC, Trainor VC, Hanson MR, Veit B (2004) Diversification of genes encoding *mei2-like* RNA binding proteins in plants. Plant molecular biology 54: 653-670

Anderson NO (2006) Flower breeding and genetics: issues, challenges and opportunities for the 21st century. Springer Science & Business Media

Anderson SL, Somers DE, Millar AJ, Hanson K, Chory J, Kay SA (1997) Attenuation of phytochrome A and B signaling pathways by the *Arabidopsis* circadian clock. The Plant Cell 9: 1727-1743

Ando T, Nomura M, Tsukahara J, Watanabe H, Kokubun H, Tsukamoto T, Hashimoto G, Marchesi E, Kitching IJ (2001) Reproductive isolation in a native population of *Petunia sensu* Jussieu (*Solanaceae*). Annals of Botany 88: 403-41

Angelini LG, Martini A, Passera B, Tavarini S (2018) Cultivation of *Stevia rebaudiana* Bertoni and associated challenges. Springer Cham, Switzerland: 35-85

Attaya A (2017) 10 An efficient protocol of *Stevia rebaudiana* regeneration for large-scale production. Egyptian Journal of Agronomy, 39(1):117-125

Bahmani K (2021) Genetic analysis of important metabolites in fennel and stevia. Michigan State University

Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S, Werck-Reichhart D (2011) Cytochromes P450. The *Arabidopsis* Book/American Society of Plant Biologists 9

Bao S, Hua C, Shen L, Yu H (2020) New insights into gibberellin signaling in regulating flowering in *Arabidopsis*. Journal of Integrative Plant Biology 62: 118-131

Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. Annual review of plant biology 65: 415-442

Basharat S, Huang Z, Gong M, Lv X, Ahmed A, Hussain I, Li J, Du G, Liu L (2021) A review on current conventional and biotechnical approaches to enhance biosynthesis of steviol glycosides in *Stevia rebaudiana*. Chinese Journal of Chemical Engineering 30: 92-104

Benedito VA, Visser PB, Angenent GC, Krens FA (2004) The potential of virus-induced gene silencing for speeding up functional characterization of plant genes. Genetics and Molecular Research 3: 323-341

Benhmimou A, Ibriz M, Al Faïz C, Douaik A, Khiraoui A, Amchra FZ, Lage M (2017) Productivity of new sweet plant in Morocco (*Stevia rebaudiana* Bertoni) under water stress. Journal of Medicinal Plants Studies 5: 126-131

Bennypaul HS, Mutti JS, Rustgi S, Kumar N, Okubara PA, Gill KS (2012) Virus-induced gene silencing (VIGS) of genes expressed in root, leaf, and meiotic tissues of wheat. Functional & integrative genomics 12: 143-156

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120

Bombarely A, Moser M, Amrad A, Bao M, Bapaume L, Barry CS, Bliek M, Boersma MR, Borghi L, Bruggmann R (2016) Insight into the evolution of the *Solanaceae* from the parental genomes of *Petunia hybrida*. Nature plants 2: 1-9

Brandle J, Telmer P (2007) Steviol glycoside biosynthesis. Phytochemistry 68: 1855-1863

Broderick SR, Jones ML (2014) An optimized protocol to increase virus-induced gene silencing efficiency and minimize viral symptoms in petunia. Plant molecular biology reporter 32: 219-233

Caplan J, Dinesh-Kumar S (2006) Using viral vectors to silence endogenous genes. Current protocols in microbiology 1: 16I. 16.11-16I. 16.13

Carakostas M, Prakash I, Kinghorn AD, Wu CD, Soejarto DD (2012) Steviol glycosides. Alternative sweeteners: 159-180

Ceunen S, Geuns JM (2013) Steviol glycosides: chemical diversity, metabolism, and function. Journal of natural products 76: 1201-1228

Chester K, Tamboli ET, Parveen R, Ahmad S (2013) Genetic and metabolic diversity in *Stevia rebaudiana* using RAPD and HPTLC analysis. Pharmaceutical biology 51: 771-777

Chuck GS, Brown PJ, Meeley R, Hake S (2014) Maize *SBP-box* transcription factors *unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral primordia initiation. Proceedings of the national academy of sciences 111: 18775-18780

Coculo D, Lionetti V (2022) The plant invertase/pectin methylesterase inhibitor superfamily. Frontiers in plant science 13: 863892

Cosgrove DJ (2000) Loosening of plant cell walls by expansins. Nature 407: 321-326

Cosgrove DJ (2016) Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes. Journal of experimental botany 67: 463-476

Cristina MS, Petersen M, Mundy J (2010) Mitogen-activated protein kinase signaling in plants. Annual review of plant biology 61: 621-649

Czauderna F, Fechtner M, Dames S, AyguÈn H, Klippel A, Pronk GJ, Giese K, Kaufmann J (2003) Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. Nucleic acids research 31: 2705-2716

Dai X, Mashiguchi K, Chen Q, Kasahara H, Kamiya Y, Ojha S, DuBois J, Ballou D, Zhao Y (2013) The biochemical mechanism of auxin biosynthesis by an *Arabidopsis YUCCA* flavin-containing monooxygenase. Journal of Biological Chemistry 288: 1448-1457

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM (2021) Twelve years of SAMtools and BCFtools. Gigascience 10: giab008

de Souza A, Hull PA, Gille S, Pauly M (2014) Identification and functional characterization of the distinct plant pectin esterases *PAE8* and *PAE9* and their deletion mutants. Planta 240: 1123-1138

Delourme R, Falentin C, Fomeju BF, Boillot M, Lassalle G, André I, Duarte J, Gauthier V, Lucante N, Marty A (2013) High-density SNP-based genetic map development and linkage disequilibrium assessment in Brassica napus L. BMC genomics 14: 1-18

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43: 491-498

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15-21

Dong Y, Burch-Smith TM, Liu Y, Mamillapalli P, Dinesh-Kumar SP (2007) A ligation-independent cloning tobacco rattle virus vector for high-throughput virus-induced gene silencing identifies roles for *NbMADS4-1* and *-2* in floral development. Plant physiology 145: 1161-1170

Ezquer I, Salameh I, Colombo L, Kalaitzis P (2020) Plant cell walls tackling climate change: Insights into plant cell wall remodeling, its regulation, and biotechnological strategies to improve crop adaptations and photosynthesis in response to global warming. Plants 9: 1-27

Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Tecle IY, Strickler SR, Bombarely A, Fisher-York T, Pujar A, Foerster H (2015) The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. Nucleic acids research 43: D1036-D1041

Fleming AJ (2005) Formation of primordia and phyllotaxy. Current opinion in plant biology 8: 53-58

Fleming AJ, McQueen-Mason S, Mandel T, Kuhlemeier C (1997) Induction of leaf primordia by the cell wall protein expansin. Science 276: 1415-1418

Forestan C, Varotto S (2012) The role of PIN auxin efflux carriers in polar auxin transport and accumulation and their effect on shaping maize development. Molecular plant 5: 787-798

Fu D-Q, Zhu B-Z, Zhu H-L, Zhang H-X, Xie Y-H, Jiang W-B, Zhao X-D, Luo Y-B (2006) Enhancement of virus-induced gene silencing in tomato by low temperature and low humidity. Molecules and Cells 21: 153-160

Gelvin S (2012) Plant molecular biology manual. Springer Science & Business Media

Gerats T, Strommer J (2008) *Petunia*: evolutionary, developmental and physiological genetics. Springer Science & Business Media

Gholizadeh A (2020) Pectin methylesterase activity of plant *DUF538* protein superfamily. Physiology and molecular biology of plants 26: 829-839

Giulini A, Wang J, Jackson D (2004) Control of phyllotaxy by the cytokinin-inducible response regulator homologue *ABPHYL1*. Nature 430: 1031-1034

Goettemoeller J, Ching A (1999) Seed germination in *Stevia rebaudiana*. Perspectives on new crops and new uses. ASHS Press, Alexandria, VA: 510-511

Gou J-Y, Miller LM, Hou G, Yu X-H, Chen X-Y, Liu C-J (2012) Acetylesterase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. The Plant Cell 24: 50-65

Goyal S, Samsher GR, Goyal R (2010) Stevia (*Stevia rebaudiana*) a bio-sweetener: a review. Int J Food Sci Nutr 61: 1-10

Green PB, Steele C, Rennich S (1996) Phyllotactic patterns: a biophysical mechanism for their origin. Annals of Botany 77: 515-528

Gregis V, Andrés F, Sessa A, Guerra RF, Simonini S, Mateos JL, Torti S, Zambelli F, Prazzoli GM, Bjerkan KN (2013) Identification of pathways directly regulated by *SHORT*

VEGETATIVE PHASE during vegetative and reproductive development in *Arabidopsis*. Genome biology 14: 1-26

Griffiths J, Barrero JM, Taylor J, Helliwell CA, Gubler F (2011) *ALTERED MERISTEM PROGRAM 1* is involved in development of seed dormancy in *Arabidopsis*. PLoS One 6: e20408

Grigg SP, Canales C, Hay A, Tsiantis M (2005) *SERRATE* coordinates shoot meristem function and leaf axial patterning in *Arabidopsis*. Nature 437: 1022-1026

Groszmann M, Bylstra Y, Lampugnani ER, Smyth DR (2010) Regulation of tissue-specific expression of *SPATULA*, a bHLH gene involved in carpel development, seedling germination, and lateral organ growth in *Arabidopsis*. Journal of experimental botany 61: 1495-1508

Guo Y, Lin W-K, Chen Q, Vallejo VA, Warner RM (2017) Genetic determinants of crop timing and quality traits in two interspecific petunia recombinant inbred line populations. Scientific Reports 7: 3200

Guo Y, Wiegert-Rininger KE, Vallejo VA, Barry CS, Warner RM (2015) Transcriptome-enabled marker discovery and mapping of plastochron-related genes in *Petunia spp*. BMC genomics 16: 1-19

Gupta E, Purwar S, Sundaram S, Rai G (2013) Nutritional and therapeutic values of *Stevia rebaudiana*: A review. Journal of Medicinal Plants Research 7: 3343-3353

Hao Y, Zong X, Ren P, Qian Y, Fu A (2021) Basic helix-loop-helix (bHLH) transcription factors regulate a wide range of functions in *Arabidopsis.* International Journal of Molecular Sciences 22: 7152

Hasegawa M, Kishino H, Yano T-a (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution 22: 160-174

Hastoy C, Cosson P, Cavaignac S, Boutie P, Waffo-Teguo P, Rolin D, Schurdi-Levraud V (2019) Deciphering performances of fifteen genotypes of *Stevia rebaudiana* in southwestern France through dry biomass and steviol glycoside evaluation. Industrial Crops and Products 128: 607-619

Hastoy C, Le Bihan Z, Gaudin J, Cosson P, Rolin D, Schurdi-Levraud V (2019) First report of *Septoria* sp. infecting *Stevia rebaudiana* in France and screening of *Stevia rebaudiana* genotypes for host resistance. Plant disease 103: 1544-1550

Heikal AH, Badawy O, Hafez AM (2008) Genetic relationships among some Stevia (*Stevia rebaudiana* Bertoni) accessions based on ISSR analysis. Research Journal of Cell and Molecular Biology 2: 1-5

Hein I, Barciszewska-Pacak M, Hrubikova K, Williamson S, Dinesen M, Soenderby IE, Sundar S, Jarmolowski A, Shirasu K, Lacomme C (2005) Virus-induced gene silencing-based

functional characterization of genes associated with powdery mildew resistance in barley. Plant Physiology 138: 2155-2164

Heisler MG, Atkinson A, Bylstra YH, Walsh R, Smyth DR (2001) *SPATULA*, a gene that controls development of carpel margin tissues in *Arabidopsis*, encodes a bHLH protein. Development 128: 1089-1098

Helliwell CA, Chin-Atkins AN, Wilson IW, Chapple R, Dennis ES, Chaudhury A (2001) The *Arabidopsis AMP1* gene encodes a putative glutamate carboxypeptidase. The Plant Cell 13: 2115-2125

Hibara K-i, Miya M, Benvenuto SA, Hibara-Matsuo N, Mimura M, Yoshikawa T, Suzuki M, Kusaba M, Taketa S, Itoh J-i (2021) Regulation of the plastochron by three many-noded dwarf genes in barley. PLoS genetics 17: e1009292

Hirayama T, Ishida C, Kuromori T, Obata S, Shimoda C, Yamamoto M, Shinozaki K, Ohto C (1997) Functional cloning of a cDNA encoding *Mei2-like protein from Arabidopsis thaliana* using a fission yeast pheromone receptor deficient mutant. FEBS letters 413: 16-20

Holen T, Amarzguioui M, Wiiger MT, Babaie E, Prydz H (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. Nucleic acids research 30: 1757-1766

Houston K, Tucker MR, Chowdhury J, Shirley N, Little A (2016) The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. Frontiers in plant science 7: 205693

Hu X, Hipolito S, Lynn R, Abraham V, Ramos S, Wong-Staal F (2004) Relative gene-silencing efficiencies of small interfering RNAs targeting sense and antisense transcripts from the same genetic locus. Nucleic Acids Research 32: 4609-4617

Huber BM, Wehner TC (2023) Heritability and genetic variance estimates for agronomic traits and glycoside yield in four elite stevia breeding populations. Crop Science

Huq E, Quail PH (2002) *PIF4*, a phytochrome-interacting bHLH factor, functions as a negative regulator of phytochrome B signaling in *Arabidopsis*. The EMBO journal

Ijaz M, Pirzada AM, Saqib M, Latif M (2015) *Stevia rebaudiana*: An alternative sugar crop in Pakistan–a review. Erling Verl. GmbH Co. KG 20: 88-96

Jackson D, Hake S (1999) Control of phyllotaxy in maize by the *abphyl1* gene. Development 126: 315-323

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

Jean RV, Barab D (1998) Symmetry in plants, Vol 4. World Scientific

Jeffares DC, Phillips MJ, Moore S, Veit B (2004) A description of the *Mei2-like protein* family; structure, phylogenetic distribution and biological context. Development Genes and Evolution 214: 149-158

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X (2010) Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. Nature genetics 42: 541-544

Jing Y, Lin R (2020) Transcriptional regulatory network of the light signaling pathways. New Phytologist 227: 683-697

Kaur J, Sebastian J, Siddiqi I (2006) The *Arabidopsis-mei2-like* genes play a role in meiosis and vegetative growth in *Arabidopsis*. The Plant Cell 18: 545-559

Kaur R, Sharma N, Raina R (2015) Identification and functional annotation of expressed sequencetags based SSR markers of *Stevia rebaudiana*. Turkish Journal of Agriculture and Forestry 39: 439-450

Kaur Y, Das N (2023) Gibberellin 2-Oxidases in Potato (Solanum tuberosum L.): Cloning, Characterization, In Silico Analysis and Molecular Docking. Molecular Biotechnology: 1-16

Kawakatsu T, Itoh J-I, Miyoshi K, Kurata N, Alvarez N, Veit B, Nagato Y (2006) *PLASTOCHRON2* regulates leaf initiation and maturation in rice. The Plant Cell 18: 612-625

Kawakatsu T, Taramino G, Itoh JI, Allen J, Sato Y, Hong SK, Yule R, Nagasawa N, Kojima M, Kusaba M (2009) *PLASTOCHRON3/GOLIATH* encodes a glutamate carboxypeptidase required for proper development in rice. The Plant Journal 58: 1028-1040

Kessler S, Sinha N (2004) Shaping up: the genetic control of leaf shape. Current Opinion in Plant Biology 7: 65-72

Kho KH, Sukhan ZP, Hossen S, Cho Y, Kim SC, Sharker MR, Jung H-J, Nou I-S (2021) Construction of a genetic linkage map based on SNP markers, QTL mapping and detection of candidate genes of growth-related traits in Pacific abalone using genotyping-by-sequencing. Frontiers in Marine Science 8: 713783

Khuluq AD, Widaryanto E, Ariffin A, Nihayati E (2022) Adaptive strategy of *Stevia rebaudiana* to environmental change in tropical climate based on anatomy and physiology characteristics. Biodiversitas Journal of Biological Diversity 23

Kosambi DD (2016) The estimation of map distances from recombination values. DD Kosambi: selected works in mathematics and statistics: 125-130

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution 35: 1547

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics 9: 1-13

Lee B-h, Yu S-i, Jackson D (2009) Control of plant architecture: the role of phyllotaxy and plastochron. Journal of Plant Biology 52: 277-282

Li C, Zheng L, Wang X, Hu Z, Zheng Y, Chen Q, Hao X, Xiao X, Wang X, Wang G (2019) Comprehensive expression analysis of *Arabidopsis GA2-oxidase* genes and their functional insights. Plant Science 285: 1-13

Li D, Fu X, Guo L, Huang Z, Li Y, Liu Y, He Z, Cao X, Ma X, Zhao M (2016) *FAR-RED ELONGATED HYPOCOTYL3* activates *SEPALLATA2* but inhibits *CLAVATA3* to regulate meristem determinacy and maintenance in *Arabidopsis*. Proceedings of the National Academy of Sciences 113: 9375-9380

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997

Li J, Li G, Wang H, Deng XW (2011) Phytochrome signaling mechanisms. The *Arabidopsis* Book/American Society of Plant Biologists 9

Li X, Gu W, Sun S, Chen Z, Chen J, Song W, Zhao H, Lai J (2018) *Defective Kernel 39* encodes a PPR protein required for seed development in maize. Journal of integrative plant biology 60: 45-64

Lian X-Y, Gao H-N, Jiang H, Liu C, Li Y-Y (2021) *MdKCS2* increased plant drought resistance by regulating wax biosynthesis. Plant Cell Reports 40: 2357-2368

Libik-Konieczny M, Capecka E, Tuleja M, Konieczny R (2021) Synthesis and production of steviol glycosides: recent research trends and perspectives. Applied Microbiology and Biotechnology 105: 3883-3900

Lin R, Ding L, Casola C, Ripoll DR, Feschotte C, Wang H (2007) Transposase-derived transcription factors regulate light signaling in *Arabidopsis*. Science 318: 1302-1305

Lin Y, Jones ML (2022) CRISPR/Cas9-mediated editing of *Autophagy Gene 6* in petunia decreases flower longevity, seed yield, and phosphorus remobilization by accelerating ethylene production and senescence-related gene expression. Frontiers in Plant Science 13: 840218

Liu Y, He J, Chen Z, Ren X, Hong X, Gong Z (2010) *ABA overly-sensitive 5* (*ABO5*), encoding a pentatricopeptide repeat protein required for cis-splicing of mitochondrial *nad2* intron 3, is involved in the abscisic acid response in *Arabidopsis*. The Plant Journal 63: 749-765

Liu Y, Schiff M, Marathe R, Dinesh-Kumar S (2002) Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for N-mediated resistance to tobacco mosaic virus. The Plant Journal 30: 415-429

Liu Y-J, Xiu Z-H, Meeley R, Tan B-C (2013) *Empty pericarp5* encodes a pentatricopeptide repeat protein that is required for mitochondrial RNA editing and seed development in maize. The Plant Cell 25: 868-883

Livak KJ, Wills QF, Tipping AJ, Datta K, Mittal R, Goldson AJ, Sexton DW, Holmes CC (2013) Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. Methods 59: 71-79

Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J (2006) *SERRATE*: a new player on the plant microRNA scene. EMBO reports 7: 1052-1058

Lohmann D, Stacey N, Breuninger H, Jikumaru Y, Müller D, Sicard A, Leyser O, Yamaguchi S, Lenhard M (2010) *SLOW MOTION* is required for within-plant auxin homeostasis and normal timing of lateral organ initiation at the shoot meristem in *Arabidopsis.* The Plant Cell 22: 335-348

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15: 1-21

Lu Y, Li C, Wang H, Chen H, Berg H, Xia Y (2011) *AtPPR2*, an *Arabidopsis* pentatricopeptide repeat protein, binds to plastid 23S rRNA and plays an important role in the first mitotic division during gametogenesis and in cell proliferation during embryogenesis. The Plant Journal 67: 13-25

Lurin C, Andreés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. The Plant Cell 16: 2089-2103

Ma L, Li G (2018) FAR1-related sequence (FRS) and FRS-related factor (FRF) family proteins in *Arabidopsis* growth and development. Frontiers in Plant Science 9: 355778

Madan S, Ahmad S, Singh G, Kohli K, Kumar Y, Singh R, Garg M (2010) *Stevia rebaudiana* (Bert.) Bertoni-a review.

Mallona I, Lischewski S, Weiss J, Hause B, Egea-Cortines M (2010) Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. BMC plant biology 10: 1-11

Markham JE, Molino D, Gissot L, Bellec Y, Hématy K, Marion J, Belcram K, Palauqui J-C, Satiat-JeuneMaître B, Faure J-D (2011) Sphingolipids containing very-long-chain fatty acids define a secretory pathway for specific polar plasma membrane protein targeting in *Arabidopsis*. The Plant Cell 23: 2362-2378

Matsushima N, Takatsuka S, Miyashita H, Kretsinger RH (2019) Leucine rich repeat proteins: sequences, mutations, structures and diseases. Protein and peptide letters 26: 108-131

Mercier R, Grelon M (2008) Meiosis in plants: ten years of gene discovery. Cytogenetic and genome research 120: 281-290

Mimura M, Itoh J-I (2014) Genetic interaction between rice *PLASTOCHRON* genes and the gibberellin pathway in leaf development. Rice 7: 25

Mimura M, Nagato Y, Itoh J-I (2012) Rice *PLASTOCHRON* genes regulate leaf maturation downstream of the gibberellin signal transduction pathway. Planta 235: 1081-1089

Mindrebo JT, Nartey CM, Seto Y, Burkart MD, Noel JP (2016) Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom. Current opinion in structural biology 41: 233-246

Miyoshi K, Ahn B-O, Kawakatsu T, Ito Y, Itoh J-I, Nagato Y, Kurata N (2004) *PLASTOCHRON1*, a timekeeper of leaf initiation in rice, encodes cytochrome P450. Proceedings of the National Academy of Sciences 101: 875-880

Mizutani K, Tanaka O (2001) Use of *Stevia rebaudiana* sweeteners in Japan. *In* Stevia. CRC Press, pp 191-209

Moraes RM, Donega MA, Cantrell CL, Mello SC, McChesney JD (2013) Effect of harvest timing on leaf production and yield of diterpene glycosides in *Stevia rebaudiana* Bert: A specialty perennial crop for Mississippi. Industrial Crops and Products 51: 385-389

Muruganantham M, Moskovitz Y, Haviv S, Horesh T, Fenigstein A, du Preez J, Stephan D, Burger JT, Mawassi M (2009) Grapevine virus A-mediated gene silencing in *Nicotiana benthamiana* and *Vitis vinifera*. Journal of virological methods 155: 167-174

Noman A, Aqeel M, He S (2016) CRISPR-Cas9: tool for qualitative and quantitative plant genome editing. Frontiers in plant science 7: 216074

Noor SH, Ushijima K, Murata A, Yoshida K, Tanabe M, Tanigawa T, Kubo Y, Nakano R (2014) Double flower formation induced by silencing of C-class *MADS-box* genes and its variation among petunia cultivars. Scientia Horticulturae 178: 1-7

Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S (2003) Gibberellin biosynthesis and response during *Arabidopsis* seed germination. The Plant Cell 15: 1591-1604

Oliveros J (2016) Venny 2.1. 0. Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams.(2007-2015). Available online at: http://bioinfogp. cnb. csic. es/tools/venny/(Accessed February 15, 2016)

Ooijen V (2009) MapQTL6, Software for the mapping of quantitative trait loci in experimental population of diploid species. Wageningen, Netherlands, Kyazma BV

Ooijen V (2018) JoinMap® 5, Software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma BV, Wageningen, Netherlands.

Pei G, Chen L, Zhang W (2017) WGCNA application to proteomic and metabolomic data analysis. Methods in enzymology, Vol 585. Elsevier, pp 135-158

Peng J, Harberd NP (2002) The role of GA-mediated signalling in the control of seed germination. Current opinion in plant biology 5: 376-381

Pien S, Wyrzykowska J, McQueen-Mason S, Smart C, Fleming A (2001) Local expression of expansin induces the entire process of leaf development and modifies leaf shape. Proceedings of the National Academy of Sciences 98: 11812-11817

Pineda-Hernández E, Cruz-Valderrama JE, Gómez-Maqueo X, Martínez-Barajas E, Gamboa-deBuen A (2022) BIIDXI, a DUF642 cell wall protein that regulates pectin methyl esterase activity, is involved in thermotolerance processes in *Arabidopsis thaliana*. Plants 11: 3049

Prakash I, Markosyan A, Bunders C (2014) Development of next generation stevia sweetener: Rebaudioside M. Foods 3: 162-175

Preston JC, Jorgensen SA, Orozco R, Hileman LC (2016) Paralogous *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes differentially regulate leaf initiation and reproductive phase change in petunia. Planta 243: 429-440

Prigge MJ, Wagner DR (2001) The *Arabidopsis SERRATE* gene encodes a zinc-finger protein required for normal shoot development. The Plant Cell 13: 1263-1280

Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F (2022) Analysing high-throughput sequencing data in Python with HTSeq 2.0. Bioinformatics 38: 2943-2945

Ramesh K, Singh V, Megeji NW (2006) Cultivation of stevia [*Stevia rebaudiana* (Bert.) Bertoni]: A comprehensive review. Advances in agronomy 89: 137-177

Reid M, Chen J-C, Jiang C-Z (2009) Virus-Induced Gene Silencing for functional characterization of genes in petunia. *In* T Gerats, J Strommer, eds, Petunia: Evolutionary, Developmental and Physiological Genetics. Springer New York, New York, NY, pp 381-394

Reinhardt D, Kuhlemeier C (2002) Plant architecture. EMBO reports 3: 846-851

Reinhardt D, Mandel T, Kuhlemeier C (2000) Auxin regulates the initiation and radial position of plant lateral organs. The Plant Cell 12: 507-518

Reinhardt D, Pesce E-R, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C (2003) Regulation of phyllotaxis by polar auxin transport. Nature 426: 255-260

Reinhardt D, Wittwer F, Mandel T, Kuhlemeier C (1998) Localized upregulation of a new expansin gene predicts the site of leaf formation in the tomato meristem. The Plant Cell 10: 1427-1437

Rossi ML, Souza EH, Graner EM, Almeida MD, Martinelli AP (2018) Post-seminal development and morphoanatomy of vegetative and reproductive organs in Stevia rebaudiana (Bert.) Bertoni (Asteraceae). Anais da Academia Brasileira de Ciências 90: 2167-2177

Rutishauser R, Peisl P (2001) Phyllotaxy. e LS

Sairkar P, Chandravanshi M, Shukla N, Mehrotra N (2009) Mass production of an economically important medicinal plant *Stevia rebaudiana* using in vitro propagation techniques. Journal of Medicinal Plants Research 3: 266-270

Salazar-Iribe A, Agredano-Moreno LT, Zúñiga-Sánchez E, Jiménez-Garcia LF, Gamboa-deBuen A (2016) The cell wall DUF642 At2g41800 (TEB) protein is involved in hypocotyl cell elongation. Plant science 253: 206-214

Santos KM, Fisher PR, Yeager T, Simonne EH, Carter HS, Argo WR (2011) Timing of macronutrient supply during cutting propagation of Petunia. HortScience 46: 475-480

Sanyal SK, Rajasheker G, Kishor PK, Kumar SA, Kumari PH, Saritha K, Rathnagiri P, Pandey GK (2020) Role of protein phosphatases in signaling, potassium transport, and abiotic stress responses. Protein phosphatases and stress management in plants: Functional genomic perspective: 203-232

Schmidt P, Hartung J, Rath J, Piepho H-P (2019) Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. Crop Science 59: 525-536

Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific Effects of MicroRNAs on the Plant Transcriptome. Developmental Cell 8: 517-527

Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P (2008) The microRNA regulated SBP-box genes *SPL9* and *SPL15* control shoot maturation in *Arabidopsis*. Plant Molecular Biology 67: 183-195

Selker JM, Steucek GL, Green PB (1992) Biophysical mechanisms for morphogenetic progressions at the shoot apex. Developmental Biology 153: 29-43

Sharma S, Walia S, Singh B, Kumar R (2016) Comprehensive review on agro technologies of low-calorie natural sweetener stevia (*Stevia rebaudiana* Bertoni): a boon to diabetic patients. Journal of the Science of Food and Agriculture 96: 1867-1879

Shivanna N, Naika M, Khanum F, Kaul VK (2013) Antioxidant, anti-diabetic and renal protective properties of *Stevia rebaudiana*. Journal of Diabetes and its Complications 27: 103-113

Shock C (1982) Rebaudi's stevia: natural noncaloric sweeteners. California Agriculture 36: 4-5

Simlat M, Ptak A, Skrzypek E, Warchoł M, Morańska E, Piórkowska E (2018) Melatonin significantly influences seed germination and seedling growth of *Stevia rebaudiana* Bertoni. PeerJ 6: e5009

Singh AK, Ghosh D, Chakraborty S (2022) Optimization of Tobacco Rattle Virus (TRV)-based virus-induced gene silencing (VIGS) in tomato. *In* Plant Gene Silencing: Methods and Protocols. Springer, pp 133-145

Skinner AC, Ravanbakht SN, Skelton JA, Perrin EM, Armstrong SC (2018) Prevalence of obesity and severe obesity in US children, 1999–2016. Pediatrics 141

Smirnoff N, Wheeler GL (2000) Ascorbic acid in plants: biosynthesis and function. Critical reviews in plant sciences 19: 267-290

Smitha G, Umesha K (2012) Vegetative propagation of Stevia [*Stevia rebaudiana* (Bertoni) Hemsl.] through stem cuttings. Journal of Tropical Agriculture 50: 72-75

Snow M, Snow GRS (1932) I. Experiments on Phyllotaxis. I.-The effect of isolating a Primordium. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character 221: 1-43

Stecher G, Tamura K, Kumar S (2020) Molecular evolutionary genetics analysis (MEGA) for macOS. Molecular biology and evolution 37: 1237-1239

Stehmann JR, Lorenz-Lemke AP, Freitas LB, Semir J (2009) The Genus *Petunia*. T Gerats, J Strommer, eds, Petunia: Evolutionary, Developmental and Physiological Genetics. Springer New York, New York, NY, pp 1-28

Stevens R, Truffault V, Baldet P, Gautier H (2017) Ascorbate oxidase in plant growth, development, and stress tolerance. Ascorbic acid in plant growth, development and stress tolerance: 273-295

Stuurman J, Jäggi F, Kuhlemeier C (2002) Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. Genes & development 16: 2213-2218

Su C, Lu W, Zhao T, Gai J (2010) Verification and fine-mapping of QTLs conferring days to flowering in soybean using residual heterozygous lines. Chinese Science Bulletin 55: 499-508

Subburaj S, Chung SJ, Lee C, Ryu S-M, Kim DH, Kim J-S, Bae S, Lee G-J (2016) Site-directed mutagenesis in *Petunia× hybrida* protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. Plant cell reports 35: 1535-1544

Suzuki M, Sato Y, Wu S, Kang B-H, McCarty DR (2015) Conserved functions of the MATE transporter *BIG EMBRYO 1* in regulation of lateral organ size and initiation rate. The Plant Cell 27: 2288-2300

Tam NT, Dwiyanti MS, Koide Y, Nagano AJ, Ky H, Tin HQ, Hien NL, Dung LV, Kishima Y (2019) Profiling SNP and nucleotide diversity to characterize Mekong Delta rice landraces in southeast Asian populations. The Plant Genome 12: 190042

Tavarini S, Passera B, Angelini LG (2018) Crop and steviol glycoside improvement in stevia by breeding.

Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M (2010) A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. Methods 50: S1-S5

Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G (2020) PlantRegMap: charting functional regulatory maps in plants. Nucleic acids research 48: D1104-D1113

Toma F, Vasilescu T, Sorina P, Zamfir-Vasca D (2011) Research concerning the propagation by cutting of some new cultivars of Petunia.

Tomar M, Singh B, Bhardwaj V, Sood S, Singh B, Salaria N, Thakur K, Kumar A, Sharma N, Goutam U (2021) Validation of molecular response of tuberization in response to elevated temperature by using a transient Virus Induced Gene Silencing (VIGS) in potato. Functional & Integrative Genomics 21: 215-229

Tsukaya H (2005) Leaf shape: genetic controls and environmental factors. The International journal of developmental biology 49: 547-555

Tsykun T, Rellstab C, Dutech C, Sipos G, Prospero S (2017) Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus *Armillaria cepistipes*. Heredity 119: 371-380

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, McElver J, Aux G, Patton D (2004) Identification of genes required for embryo development in *Arabidopsis*. Plant physiology 135: 1206-1220

Ucar E, Özyiğit Y, Turgut K (2016) The effects of light and temperature on germination of Stevia (*Stevia rebaudiana* Bert.) seeds. Turk J. Agric Res (2016) 3: 37-40

Ueguchi-Tanaka M, Ashikari M, Nakajima M, Itoh H, Katoh E, Kobayashi M, Chow T-y, Hsing Y-iC, Kitano H, Yamaguchi I (2005) *GIBBERELLIN INSENSITIVE DWARF1* encodes a soluble receptor for gibberellin. Nature 437: 693-698

Unver T, Budak H (2009) Virus-Induced Gene Silencing, a post transcriptional gene silencing method. International Journal of Plant Genomics 2009

Upadhyay N, Kar D, Deepak Mahajan B, Nanda S, Rahiman R, Panchakshari N, Bhagavatula L, Datta S (2019) The multitasking abilities of MATE transporters in plants. Journal of experimental botany 70: 4643-4656

Urso S, Zottini M, Ruberti C, Lo Schiavo F, Stanca AM, Cattivelli L, Valè G (2013) An *Agrobacterium tumefaciens*-mediated gene silencing system for functional analysis in grapevine. Plant Cell, Tissue and Organ Culture (PCTOC) 114: 49-60

Vallejo VA, Tychonievich J, Lin W-K, Wangchu L, Barry CS, Warner RM (2015) Identification of QTL for crop timing and quality traits in an interspecific *Petunia* population. Molecular Breeding 35: 2

Vallejo VA, Warner RM (2021) Identifying quantitative trait loci for steviol glycoside production in *Stevia rebaudiana* using transcriptome-derived SSRs. Industrial Crops and Products 161: 113176

Vandenbussche M, Chambrier P, Rodrigues Bento S, Morel P (2016) Petunia, your next supermodel? Frontiers in plant science 7: 72

Veit B, Briggs SP, Schmidt RJ, Yanofsky MF, Hake S (1998) Regulation of leaf initiation by the *terminal ear 1* gene of maize. Nature 393: 166-168

Velásquez AC, Chakravarthy S, Martin GB (2009) Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. JoVE (Journal of Visualized Experiments): e1292

Venkataraman K, Futerman AH (2002) Do longevity assurance genes containing Hox domains regulate cell development via ceramide synthesis? FEBS letters 528: 3-4

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of heredity 93: 77-78

Wang F, Yu Z, Zhang M, Wang M, Lu X, Liu X, Li Y, Zhang X, Tan Bc, Li C (2022) *ZmTE1* promotes plant height by regulating intercalary meristem formation and internode cell elongation in maize. Plant biotechnology journal 20: 526-537

Wang H, Deng XW (2002) *Arabidopsis FHY3* defines a key phytochrome A signaling component directly interacting with its homologous partner FAR1. The EMBO Journal

Wang H, Kong F, Zhou C (2021) From genes to networks: The genetic control of leaf development. Journal of Integrative Plant Biology 63: 1181-1196

Wang J-W, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*. The Plant Cell 20: 1231-1243

Wang L, Sun S, Jin J, Fu D, Yang X, Weng X, Xu C, Li X, Xiao J, Zhang Q (2015) Coordinated regulation of vegetative and reproductive branching in rice. Proceedings of the National Academy of Sciences 112: 15504-15509

Wang M, Wang G, Ji J, Wang J (2009) The effect of *PDS* gene silencing on chloroplast pigment composition, thylakoid membrane structure and photosynthesis efficiency in tobacco plants. Plant Science 177: 222-226

Warner RM (2009) Determination of photoperiod-sensitive stages of development of the short-day plant Celosia. HortScience 44: 328-333

Warner RM (2010) Temperature and Photoperiod Influence Flowering and Morphology of Four Petunia spp. HortScience 45: 365-368

Warner RM, Walworth AE (2010) Quantitative inheritance of crop timing traits in interspecific hybrid *Petunia* populations and interactions with crop quality parameters. Journal of Heredity 101: 308-316

Wei D, Liu M, Chen H, Zheng Y, Liu Y, Wang X, Yang S, Zhou M, Lin J (2018) *INDUCER OF CBF EXPRESSION 1* is a male fertility regulator impacting anther dehydration in *Arabidopsis*. PLoS genetics 14: e1007695

Werner T, Motyka V, Strnad M, Schmülling T (2001) Regulation of plant growth by cytokinin. Proceedings of the National Academy of Sciences 98: 10487-10492

Werner Ts, Motyka V, Laucou V, Smets R, Van Onckelen H, Schmülling T (2003) Cytokinin-deficient transgenic *Arabidopsis* plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. The Plant Cell 15: 2532-2550

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

Wu H-C, Bulgakov VP, Jinn T-L (2018) Pectin methylesterases: cell wall remodeling proteins are required for plant response to heat stress. Frontiers in plant science 9: 412022

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

Xie K, Wu C, Xiong L (2006) Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. Plant physiology 142: 280-293

Xu J, Naing AH, Bunch H, Jeong J, Kim H, Kim CK (2021) Enhancement of the flower longevity of petunia by CRISPR/Cas9-mediated targeted editing of ethylene biosynthesis genes. Postharvest Biology and Technology 174: 111460

Xu J, Naing AH, Kang H, Lee SY, Li W, Chung MY, Kim CK (2023) CRISPR/Cas9-mediated editing of *PhMLO1* confers powdery mildew resistance in petunia. Plant Biotechnology Reports 17: 767-775

Xu W, Huang W (2017) Calcium-dependent protein kinases in phytohormone signaling pathways. International journal of molecular sciences 18: 2436

Xu X, Yuan H, Yu X, Huang S, Sun Y, Zhang T, Liu Q, Tong H, Zhang Y, Wang Y (2021) The chromosome-level Stevia genome provides insights into steviol glycoside biosynthesis. Horticulture research 8

Yadav AK, Singh S, Rajeev R (2014) Self-incompatibility evidenced through scanning electron microscopy and pollination behaviour in *Stevia rebaudiana*. Indian Journal of Agricultural Sciences 84: 93-100

Yadav SK, Guleria P (2012) Steviol glycosides from Stevia: Biosynthesis Pathway Review and their Application in Foods and Medicine. Critical Reviews in Food Science and Nutrition 52: 988-998

Yang J, Guo Z, Luo L, Gao Q, Xiao W, Wang J, Wang H, Chen Z, Guo T (2021) Identification of QTL and candidate genes involved in early seedling growth in rice via high-density genetic mapping and RNA-seq. The Crop Journal 9: 360-371

Yang W, Schuster C, Beahan CT, Charoensawan V, Peaucelle A, Bacic A, Doblin MS, Wightman R, Meyerowitz EM (2016) Regulation of meristem morphogenesis by cell wall synthases in *Arabidopsis*. Current Biology 26: 1404-1415

Yao Y, Ban M, Brandle J (1999) A genetic linkage map for *Stevia rebaudiana*. Genome 42: 657-661

Yoshikawa T, Ozawa S, Sentoku N, Itoh J-I, Nagato Y, Yokoi S (2013) Change of shoot architecture during juvenile-to-adult phase transition in soybean. Planta 238: 229-237

Zhang H, Wang X, Pan Q, Li P, Liu Y, Lu X, Zhong W, Li M, Han L, Li J (2019) QTG-Seq accelerates QTL fine mapping through QTL partitioning and whole-genome sequencing of bulked segregant samples. Molecular plant 12: 426-437

Zhang M, Zhang S (2022) Mitogen-activated protein kinase cascades in plant signaling. Journal of integrative plant biology 64: 301-341

Zhang S, Yang C, Peng J, Sun S, Wang X (2009) *GASA5*, a regulator of flowering time and stem growth in *Arabidopsis thaliana*. Plant molecular biology 69: 745-759

Zhou Y, Kusmec A, Schnable PS (2024) Genetic regulation of self-organizing azimuthal canopy orientations and their impacts on light interception in maize. The Plant Cell: koae007

Zulfiqar S, Farooq MA, Zhao T, Wang P, Tabusam J, Wang Y, Xuan S, Zhao J, Chen X, Shen S (2023) Virus-induced gene silencing (VIGS): a powerful tool for crop improvement and its advancement towards epigenetics. International Journal of Molecular Sciences 24: 5608

APPENDIX

Table APP-3-1: Summary of hub genes of robust WGCNA modules.

GeneID	Functional description	Module
Peaxi162Scf00274g00850.1	conserved hypothetical protein [Ricinus communis] gb|EEF44747.1| conserved hypothetical protein [Ricinus communis]	black
Peaxi162Scf00120g00723.1	Phosphoribosylaminoimidazole carboxylase	black
Peaxi162Scf00192g01018.1	cleavage and polyadenylation specificity factor 160	black
Peaxi162Scf00100g00717.1	RNA-binding KH domain-containing protein	black
Peaxi162Scf01115g00013.1	Zinc finger CCCH domain-containing protein 23	black
Peaxi162Scf00020g00421.1	U-box domain-containing protein kinase family protein	black
Peaxi162Scf00129g00828.1	CRM family member 2	black
Peaxi162Scf00372g00127.1	S-norcoclaurine synthase [Morus notabilis]	black
Peaxi162Scf00777g00213.1	Elongation factor Tu	black
Peaxi162Scf01204g00001.1	Unknown protein	black
Peaxi162Scf00362g00145.1	tonoplast monosaccharide transporter2	black
Peaxi162Scf01123g00025.1	multidrug resistance-associated protein 10	black
Peaxi162Scf00065g01016.1	transducin family protein / WD-40 repeat family protein	black
Peaxi162Scf00632g00022.1	binding	black
Peaxi162Scf00071g00020.1	GTPase Der	black
Peaxi162Scf00539g00210.1	novel interactor of JAZ	black
Peaxi162Scf00320g00212.1	**auxin response factor 19**	black
Peaxi162Scf00394g00815.1	Chloroplastic group IIA intron splicing facilitator CRS1, chloroplastic	black
Peaxi162Scf01622g00017.1	Protein translocase subunit SecY	black
Peaxi162Scf00128g00134.1	Cyclin-related, putative isoform 2 [Theobroma cacao] gb|EOY00626.1| Cyclin-related, putative isoform 2 [Theobroma cacao]	black
Peaxi162Scf00081g02729.1	Transducin family protein / WD-40 repeat family protein	black
Peaxi162Scf00548g00537.1	Pentatricopeptide repeat-containing protein	black
Peaxi162Scf00275g00114.1	Pentatricopeptide repeat (PPR) superfamily protein	black
Peaxi162Scf00214g00068.1	**MEI2-like protein 5**	black
Peaxi162Scf00139g01315.1	Poly(A)-specific ribonuclease PARN	black
Peaxi162Scf00371g00319.1	Transducin/WD40 repeat-like superfamily protein	black
Peaxi162Scf00033g00415.1	SEUSS-like 2	black
Peaxi162Scf00301g01111.1	Decapping nuclease rai1	black
Peaxi162Scf00069g00185.1	Pentatricopeptide repeat-containing protein	black
Peaxi162Scf00991g00016.1	**Haloacid dehalogenase-like hydrolase (HAD) superfamily protein**	black
Peaxi162Scf00389g00932.1	RNAligase	black
Peaxi162Scf00177g00619.1	ARID/BRIGHT DNA-binding domain-containing protein	black
Peaxi162Scf00066g01518.1	ATP-dependent zinc metalloprotease FtsH	black

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00161g00018.1	N6-adenosine-methyltransferase subunit METTL14	black
Peaxi162Scf00074g01827.1	anthranilate synthase 2	black
Peaxi162Scf00732g00248.1	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein	black
Peaxi162Scf00486g00096.1	Copper-exporting P-type ATPase A	black
Peaxi162Scf00929g00519.1	Acylamino-acid-releasing enzyme, putative isoform 1 [Theobroma cacao] ref|XP_007017351.1| Acylamino-acid-releasing enzyme, putative isoform 1 [Theobroma cacao] gb|EOY14575.1| Acylamino-acid-releasing enzyme, putative isoform 1 [Theobroma cacao] gb|EOY14576.1| Acylamino-acid-releasing enzyme, putative isoform 1 [Theobroma cacao]	black
Peaxi162Scf00038g01649.1	FAR1-related sequence 4	black
Peaxi162Scf00072g00510.1	Carbamoyl-phosphate synthase large chain	black
Peaxi162Scf00119g00520.1	Serine/threonine-protein kinase Rio1	black
Peaxi162Scf00527g00065.1	TatD related DNase	black
Peaxi162Scf00008g03517.1	ATPase family AAA domain-containing protein 3	black
Peaxi162Scf00740g00424.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	black
Peaxi162Scf00526g00520.1	Duplicated homeodomain-like superfamily protein	black
Peaxi162Scf00166g01048.1	ERD (early-responsive to dehydration stress) family protein	black
Peaxi162Scf00045g00731.1	Unknown protein	black
Peaxi162Scf01061g00130.1	PENTATRICOPEPTIDE REPEAT 596	black
Peaxi162Scf01006g00315.1	Protein CPR-5	black
Peaxi162Scf00503g00220.1	Protein kinase superfamily protein	black
Peaxi162Scf00332g00222.1	Major facilitator superfamily protein	black
Peaxi162Scf01015g00113.1	Transcription initiation factor TFIID subunit 12	black
Peaxi162Scf00078g00548.1	KH domain-containing protein	black
Peaxi162Scf00071g00022.1	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein	black
Peaxi162Scf00351g00625.1	DEAD-box ATP-dependent RNA helicase 53	black
Peaxi162Scf00877g00127.1	Unknown protein	black
Peaxi162Scf00357g00734.1	Thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 2	black
Peaxi162Scf00005g04810.1	Lariat debranching enzyme	black
Peaxi162Scf00037g00123.1	Heavy metal transport/detoxification superfamily protein	black
Peaxi162Scf00113g00317.1	geranylgeranyl pyrophosphate synthase 1	black
Peaxi162Scf00685g00122.1	actin-related protein 5	black
Peaxi162Scf00045g00223.1	RNA binding	black
Peaxi162Scf00106g00514.1	Unknown protein	black
Peaxi162Scf00171g00246.1	Plastid division protein CDP1, chloroplastic	black
Peaxi162Scf00276g00133.1	Nucleotide-diphospho-sugar transferases superfamily protein	black
Peaxi162Scf00128g01776.1	Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 3	black
Peaxi162Scf01276g00039.1	**2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein**	black

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00660g00118.1	UPF0505 protein C16orf62 homolog	black
Peaxi162Scf00015g00521.1	ADP,ATP carrier protein 1	black
Peaxi162Scf00437g00632.1	alpha/beta-Hydrolases superfamily protein	black
Peaxi162Scf00084g00818.1	zinc finger family protein [Populus trichocarpa] gb|ERP50285.1| zinc finger family protein [Populus trichocarpa]	black
Peaxi162Scf00225g00014.1	cysteine synthase D1	black
Peaxi162Scf00650g00229.1	Unknown protein	black
Peaxi162Scf00062g00128.1	eukaryotic translation initiation factor 4G	black
Peaxi162Scf00945g00216.1	heavy metal atpase 1	black
Peaxi162Scf01123g00232.1	multidrug resistance-associated protein 4	black
Peaxi162Scf00071g00942.1	alpha/beta-Hydrolases superfamily protein	black
Peaxi162Scf00074g00928.1	Mitochondrial inner membrane protein OXA1	black
Peaxi162Scf00189g00020.1	prohibitin 1	black
Peaxi162Scf00102g01053.1	Phosphatidylinositol N-acetylglucosaminyltransferase subunit P	black
Peaxi162Scf00444g00851.1	VASCULAR-RELATED NAC-DOMAIN 6	black
Peaxi162Scf00081g00218.1	Emsy N Terminus (ENT)/ plant Tudor-like domains-containing protein	black
Peaxi162Scf00007g00122.1	allene oxide synthase	black
Peaxi162Scf00413g00085.1	LL-diaminopimelate aminotransferase	black
Peaxi162Scf00078g00736.1	Bifunctional aspartokinase/homoserine dehydrogenase 1	black
Peaxi162Scf00878g00327.1	DNA-binding bromodomain-containing protein	black
Peaxi162Scf00875g00226.1	50S ribosomal protein L14	mistyrose2
Peaxi162Scf00598g00071.1	Unknown protein	mistyrose2
Peaxi162Scf00257g01624.1	S-locus lectin protein kinase family protein	mistyrose2
Peaxi162Scf00074g01728.1	Unknown protein	mistyrose2
Peaxi162Scf00038g02668.1	Tetratricopeptide repeat (TPR)-like superfamily protein	mistyrose2
Peaxi162Scf00003g04426.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	mistyrose2
Peaxi162Scf00854g00219.1	Protein of unknown function, DUF647	mistyrose2
Peaxi162Scf00110g00912.1	Threonine dehydratase biosynthetic	mistyrose2
Peaxi162Scf00193g00726.1	nudix hydrolase homolog 3	mistyrose2
Peaxi162Scf00714g00217.1	Proteinase inhibitor I-B	mistyrose2
Peaxi162Scf00714g00528.1	GPI transamidase subunit PIG-U	mistyrose2
Peaxi162Scf00595g00039.1	Mediator of RNA polymerase II transcription subunit 21	mistyrose2
Peaxi162Scf00003g01335.1	MATE efflux family protein	mistyrose2
Peaxi162Scf00009g01231.1	Kelch-like protein 8	mistyrose2
Peaxi162Scf00170g00624.1	Signal recognition particle 43 kDa protein, chloroplastic	mistyrose2
Peaxi162Scf00036g00320.1	Protein kinase superfamily protein	mistyrose2
Peaxi162Scf00014g02119.1	transposase-like protein [Arabidopsis thaliana]	mistyrose2
Peaxi162Scf00488g00925.1	Cytosolic Fe-S cluster assembly factor NARFL	mistyrose2
Peaxi162Scf00331g00003.1	ATP-dependent RNA helicase ddx23	mistyrose2

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00129g01648.1	Protein kinase superfamily protein	mistyrose2
Peaxi162Scf00031g00054.1	--	mistyrose2
Peaxi162Scf00110g01622.1	26S proteasome non-ATPase regulatory subunit 13 homolog B	mistyrose2
Peaxi162Scf00827g00010.1	phospholipase D alpha 1	mistyrose2
Peaxi162Scf00408g00833.1	50S ribosomal protein L24	mistyrose2
Peaxi162Scf00064g00421.1	--	mistyrose2
Peaxi162Scf00264g00857.1	Unknown protein	mistyrose2
Peaxi162Scf00003g05025.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	mistyrose2
Peaxi162Scf00014g02020.1	AC transp product [Oryza sativa Japonica Group]	mistyrose2
Peaxi162Scf00959g00124.1	tRNA N6-adenosine threonylcarbamoyltransferase	mistyrose2
Peaxi162Scf00194g00931.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	mistyrose2
Peaxi162Scf00153g01323.1	--	mistyrose2
Peaxi162Scf00153g01322.1	protein MATERNAL EFFECT EMBRYO ARREST 50 [Arabidopsis thaliana] gb|AAC19301.1|	mistyrose2
Peaxi162Scf00393g00052.1	far-red elongated hypocotyls 3	mistyrose2
Peaxi162Scf00655g00328.1	Protein NRT1/ PTR FAMILY 8.3	mistyrose2
Peaxi162Scf00285g00428.1	Cytochrome P450 superfamily protein	mistyrose2
Peaxi162Scf00165g01722.1	Dof-type zinc finger DNA-binding family protein	mistyrose2
Peaxi162Scf00096g01719.1	Branched-chain-amino-acid aminotransferase 6	mistyrose2
Peaxi162Scf00263g01019.1	Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain	mistyrose2
Peaxi162Scf00009g02530.1	F-box family protein	mistyrose2
Peaxi162Scf00074g01332.1	Mitochondrial ribosomal protein L27	mistyrose2
Peaxi162Scf00618g00115.1	6-phosphogluconolactonase 5	mistyrose2
Peaxi162Scf00765g00049.1	exocyst complex component sec5	mistyrose2
Peaxi162Scf00227g00611.1	RNA recognition motif (RRM)-containing protein	mistyrose2
Peaxi162Scf00312g00017.1	ent-kaurenoic acid hydroxylase 2	mistyrose2
Peaxi162Scf00110g01617.1	Unknown protein	mistyrose2
Peaxi162Scf00174g00026.1	Pectin lyase-like superfamily protein	mistyrose2
Peaxi162Scf00407g00632.1	ATP binding microtubule motor family protein isoform 1 [Theobroma cacao] gb|EOY07615.1| ATP binding microtubule motor family protein isoform 1 [Theobroma cacao]	mistyrose2
Peaxi162Scf00071g00837.1	Chlorophyll(ide) b reductase NOL, chloroplastic	mistyrose2
Peaxi162Scf00152g01537.1	FAD/NAD(P)-binding oxidoreductase family protein	mistyrose2
Peaxi162Scf00014g00082.1	conserved hypothetical protein [Ricinus communis] gb|EEF32658.1| conserved hypothetical protein [Ricinus communis]	mistyrose2
Peaxi162Scf00342g00073.1	Ethylene insensitive 3 family protein	mistyrose2
Peaxi162Scf00073g00057.1	Elongation factor 2	mistyrose2
Peaxi162Scf00684g00224.1	Diaminopimelate decarboxylase	mistyrose2

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00476g00222.1	translocon at the inner envelope membrane of chloroplasts 20	mistyrose2
Peaxi162Scf00059g00128.1	methyltransferase [Arabidopsis thaliana] dbj|BAB02862.1| unnamed protein product [Arabidopsis thaliana] gb|AAL67056.1| unknown protein [Arabidopsis thaliana] gb|AAN13150.1| unknown protein [Arabidopsis thaliana] gb|AEE77448.1| methyltransferase [Arabidopsis thaliana]	mistyrose2
Peaxi162Scf00272g00224.1	MLO-like protein 6	mistyrose2
Peaxi162Scf00152g00329.1	Protein of unknown function (DUF620)	mistyrose2
Peaxi162Scf00016g01745.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	mistyrose2
Peaxi162Scf00130g00620.1	UDP-Glycosyltransferase superfamily protein	mistyrose2
Peaxi162Scf00102g01433.1	Protein kinase superfamily protein	mistyrose2
Peaxi162Scf00311g01419.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein	mistyrose2
Peaxi162Scf00038g00043.1	S-adenosylmethionine decarboxylase proenzyme	mistyrose2
Peaxi162Scf00078g00824.1	Protein kinase family protein with ARM repeat domain	mistyrose2
Peaxi162Scf00272g00828.1	rhodanese-like domain-containing protein / PPIC-type PPIASE domain-containing protein	mistyrose2
Peaxi162Scf00003g02311.1	Origin recognition complex subunit 2	mistyrose2
Peaxi162Scf00009g01923.1	8-amino-7-oxononanoate synthase [Theobroma cacao] gb|EOY07291.1| 8-amino-7-oxononanoate synthase [Theobroma cacao]	mistyrose2
Peaxi162Scf00431g00311.1	Protein transport protein Sec24-like	mistyrose2
Peaxi162Scf00248g01414.1	NAC domain containing protein 73	mistyrose2
Peaxi162Scf00257g00111.1	formin homology5	mistyrose2
Peaxi162Scf00140g01221.1	beta glucosidase 11	mistyrose2
Peaxi162Scf00025g03227.1	DHHC-type zinc finger family protein	mistyrose2
Peaxi162Scf00673g00002.1	trigalactosyldiacylglycerol 1	mistyrose2
Peaxi162Scf00789g00035.1	lysine histidine transporter 1	mistyrose2
Peaxi162Scf00822g00313.1	Plastid-lipid associated protein PAP / fibrillin family protein	mistyrose2
Peaxi162Scf00658g00065.1	expressed protein [Oryza sativa Japonica Group]	mistyrose2
Peaxi162Scf00954g00314.1	**auxin response factor 19**	sienna2
Peaxi162Scf00270g00419.1	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	sienna2
Peaxi162Scf00188g00545.1	Uridine kinase	sienna2
Peaxi162Scf00078g00424.1	WD repeat-containing protein 74	sienna2
Peaxi162Scf00458g00318.1	Remorin family protein	sienna2
Peaxi162Scf00380g00098.1	NRC1 [Solanum lycopersicum] gb|ABC26878.1| NRC1 [Solanum lycopersicum]	sienna2
Peaxi162Scf00264g00004.1	Phototropic-responsive NPH3 family protein	sienna2
Peaxi162Scf00818g00118.1	uracil dna glycosylase	sienna2
Peaxi162Scf00363g00082.1	zinc finger protein 8	sienna2
Peaxi162Scf00015g90043.1	Protein of unknown function (DUF630 and DUF632)	sienna2
Peaxi162Scf00271g00518.1	Cysteine protease ATG4A	sienna2

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00450g00317.1	Remorin family protein	sienna2
Peaxi162Scf00189g00610.1	Leucine-rich repeat receptor-like protein kinase family protein	sienna2
Peaxi162Scf00441g00624.1	Histone-lysine N-methyltransferase, H3 lysine-36 specific	sienna2
Peaxi162Scf00241g00059.1	Plant invertase/pectin methylesterase inhibitor superfamily protein	sienna2
Peaxi162Scf00269g01519.1	Omega-amidase NIT2	sienna2
Peaxi162Scf00044g00112.1	tRNA (cytosine(34)-C(5))-methyltransferase	sienna2
Peaxi162Scf00351g00526.1	Unknown protein	sienna2
Peaxi162Scf00140g01136.1	beta glucosidase 11	sienna2
Peaxi162Scf00103g01617.1	Ras-related protein Rab-6.1	sienna2
Peaxi162Scf00351g00412.1	Elongation factor 1-alpha	sienna2
Peaxi162Scf00349g00059.1	--	sienna2
Peaxi162Scf00008g04010.1	Splicing factor U2af large subunit B	sienna2
Peaxi162Scf01003g00017.1	Protein kinase superfamily protein	sienna2
Peaxi162Scf00341g00714.1	Polyadenylation factor subunit 2	sienna2
Peaxi162Scf00493g00012.1	Disease resistance protein (CC-NBS-LRR class) family	sienna2
Peaxi162Scf01276g00047.1	**Cytochrome P450 superfamily protein**	sienna2
Peaxi162Scf01159g00023.1	**Auxin-induced protein-like protein [Medicago truncatula] gb|AES85188.1| Auxin-induced protein-like protein [Medicago truncatula]**	sienna2
Peaxi162Scf00892g00120.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	sienna2
Peaxi162Scf00377g00029.1	Basic helix-loop-helix DNA-binding superfamily protein isoform 1 [Theobroma cacao] gb|EOX96336.1|	sienna2
Peaxi162Scf00531g00610.1	DNA mismatch repair protein MutS	sienna2
Peaxi162Scf01159g00226.1	ABC-2 type transporter family protein	sienna2
Peaxi162Scf00525g00079.1	4-hydroxy-tetrahydrodipicolinate synthase	sienna2
Peaxi162Scf00695g00009.1	cytochrome P450, family 88, subfamily A, polypeptide 3	sienna2
Peaxi162Scf00258g00925.1	F-box/LRR-repeat protein 20	sienna2
Peaxi162Scf00039g00093.1	diaminopimelate epimerase family protein	sienna2
Peaxi162Scf00650g00226.1	exostosin family protein	sienna2
Peaxi162Scf00130g00071.1	UDP-Glycosyltransferase superfamily protein	sienna2
Peaxi162Scf00180g00519.1	DNA double-strand break repair rad50 ATPase, putative isoform 3 [Theobroma cacao] gb|EOY04772.1|	sienna2
Peaxi162Scf00915g00249.1	F-box family protein	sienna2
Peaxi162Scf00258g00314.1	DNA-binding bromodomain-containing protein	sienna2
Peaxi162Scf00825g00416.1	Double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases superfamily protein	sienna2
Peaxi162Scf00140g01352.1	Protein of unknown function (DUF1624)	sienna2
Peaxi162Scf00038g02247.1	CCAAT/enhancer-binding protein zeta	sienna2
Peaxi162Scf00403g00113.1	Enoyl-[acyl-carrier-protein] reductase [NADH], chloroplastic	sienna2
Peaxi162Scf00130g00932.1	Transmembrane proteins 14C	sienna2

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00071g00119.1	PAR1 protein	sienna2
Peaxi162Scf01336g00122.1	OB-fold-like isoform 1 [Theobroma cacao] gb|EOY05603.1| OB-fold-like isoform 1 [Theobroma cacao]	sienna2
Peaxi162Scf01010g00217.1	Uridine 5'-monophosphate synthase	sienna2
Peaxi162Scf00608g00317.1	embryo defective 1923	sienna2
Peaxi162Scf01139g00002.1	Serine/threonine-protein kinase irlC isoform 1 [Theobroma cacao] gb|EOY08470.1|	sienna2
Peaxi162Scf00004g02625.1	sulfite reductase	sienna2
Peaxi162Scf00915g00250.1	Protein of unknown function (DUF2930)	sienna2
Peaxi162Scf00175g00525.1	**Squamosa promoter-binding-like protein 6**	sienna2
Peaxi162Scf00444g00034.1	NADH-ubiquinone oxidoreductase-related	sienna2
Peaxi162Scf00097g00108.1	Eukaryotic translation initiation factor 5A-4	sienna2
Peaxi162Scf00073g00115.1	Plasma membrane, myosin-like, Tubulin/FtsZ, N-terminal, putative isoform 4 [Theobroma cacao] gb|EOX91323.1|	sienna2
Peaxi162Scf00254g00215.1	Acetolactate synthase small subunit	sienna2
Peaxi162Scf01333g00136.1	BTB/POZ domain-containing protein	tan2
Peaxi162Scf00408g00320.1	COBW domain-containing protein 1	tan2
Peaxi162Scf00451g00723.1	Transmembrane amino acid transporter family protein	bisque4
Peaxi162Scf00911g00028.1	cytochrome P450, family 704, subfamily B, polypeptide 1	bisque4
Peaxi162Scf00042g02412.1	SBP (S-ribonuclease binding protein) family protein	bisque4
Peaxi162Scf00052g00210.1	xyloglucan endotransglucosylase/hydrolase 30	bisque4
Peaxi162Scf00253g01317.1	Unknown protein	bisque4
Peaxi162Scf00288g00815.1	Galactosyltransferase family protein	bisque4
Peaxi162Scf00022g00098.1	MADS-box transcription factor 3	bisque4
Peaxi162Scf00344g01720.1	YELLOW STRIPE like 7	bisque4
Peaxi162Scf01039g00134.1	dehydroquinate dehydratase, putative / shikimate dehydrogenase, putative	bisque4
Peaxi162Scf00038g01035.1	pleiotropic drug resistance 7	bisque4
Peaxi162Scf00152g00612.1	squalene monooxygenase 2	bisque4
Peaxi162Scf00166g00529.1	4-coumarate--CoA ligase-like 1	bisque4
Peaxi162Scf01056g90029.1	Endo-1 3(4)-beta-glucanase 1	bisque4
Peaxi162Scf00683g00448.1	Calcium-dependent lipid-binding (CaLB domain) family protein	bisque4
Peaxi162Scf00516g00671.1	NAD(P)-binding Rossmann-fold superfamily protein	bisque4
Peaxi162Scf00139g00015.1	Ycf1 protein	bisque4
Peaxi162Scf00717g00215.1	Calcium-dependent phosphotriesterase superfamily protein	bisque4
Peaxi162Scf01022g00329.1	C2H2-like zinc finger protein	bisque4
Peaxi162Scf00658g00634.1	Pectin lyase-like superfamily protein	bisque4
Peaxi162Scf00019g03125.1	magnesium transporter 9	bisque4
Peaxi162Scf00002g02826.1	Transcription initiation factor IIB-2	bisque4
Peaxi162Scf00241g00053.1	Plant invertase/pectin methylesterase inhibitor superfamily protein	bisque4
Peaxi162Scf00153g00416.1	zinc induced facilitator-like 1	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00128g01233.1	HORMA domain-containing protein 1	bisque4
Peaxi162Scf01240g00005.1	Pectin lyase-like superfamily protein	bisque4
Peaxi162Scf00131g00027.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	bisque4
Peaxi162Scf00326g00711.1	Leucine-rich repeat receptor-like protein kinase family protein	bisque4
Peaxi162Scf00200g00418.1	RNA binding protein, putative [Ricinus communis] gb|EEF38985.1| RNA binding protein, putative [Ricinus communis]	bisque4
Peaxi162Scf00422g00413.1	Acyl-coenzyme A oxidase 2, peroxisomal	bisque4
Peaxi162Scf00003g02039.1	conserved hypothetical protein [Ricinus communis] gb|EEF43357.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00074g01411.1	Cation/H(+) antiporter 18	bisque4
Peaxi162Scf00073g00173.1	Unknown protein	bisque4
Peaxi162Scf00177g00415.1	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein	bisque4
Peaxi162Scf00922g00002.1	E3 ubiquitin-protein ligase AIP2	bisque4
Peaxi162Scf00104g00017.1	ERD (early-responsive to dehydration stress) family protein	bisque4
Peaxi162Scf01108g00332.1	conserved hypothetical protein 16 [Hevea brasiliensis]	bisque4
Peaxi162Scf00132g01516.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	bisque4
Peaxi162Scf00069g00810.1	beta-fructofuranosidase, insoluble isoenzyme 1-like [Solanum tuberosum] gb|AEV46310.1| apoplastic invertase [Solanum tuberosum]	bisque4
Peaxi162Scf00003g05227.1	nodulin MtN21 /EamA-like transporter family protein	bisque4
Peaxi162Scf00174g00101.1	phospholipase D alpha 1	bisque4
Peaxi162Scf00149g00116.1	Pyrroline-5-carboxylate reductase	bisque4
Peaxi162Scf00449g00512.1	squalene monooxygenase 2	bisque4
Peaxi162Scf00029g02515.1	early nodulin-like protein 18	bisque4
Peaxi162Scf00362g00634.1	SBP family protein, putative [Theobroma cacao] gb|EOX97652.1| SBP family protein, putative [Theobroma cacao]	bisque4
Peaxi162Scf00329g00213.1	dihydroflavonol 4-reductase-like1	bisque4
Peaxi162Scf00931g00032.1	conserved hypothetical protein [Ricinus communis] gb|EEF43976.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00943g00003.1	conserved hypothetical protein [Ricinus communis] gb|EEF30394.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00560g00223.1	Aldehyde dehydrogenase family 2 member C4	bisque4
Peaxi162Scf00459g00839.1	RING/U-box superfamily protein	bisque4
Peaxi162Scf00525g00614.1	Serine/threonine-protein kinase SAPK7	bisque4
Peaxi162Scf00199g01019.1	Pectin lyase-like superfamily protein	bisque4
Peaxi162Scf00004g04219.1	Ribosome production factor 1	bisque4
Peaxi162Scf00680g00410.1	RING/U-box superfamily protein	bisque4
Peaxi162Scf00945g00013.1	Auxin-responsive GH3 family protein	bisque4
Peaxi162Scf00155g00096.1	cationic amino acid transporter 5	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00032g00821.1	NAD kinase 1	bisque4
Peaxi162Scf01160g00149.1	poly(A) polymerase 3	bisque4
Peaxi162Scf02085g00005.1	DNAse I-like superfamily protein	bisque4
Peaxi162Scf01290g00241.1	actin 4	bisque4
Peaxi162Scf00715g00117.1	DNA repair and recombination protein RAD54-like	bisque4
Peaxi162Scf00091g00176.1	Cytochrome P450 superfamily protein	bisque4
Peaxi162Scf00002g00332.1	calcium ATPase 2	bisque4
Peaxi162Scf00129g00085.1	zinc finger (AN1-like) family protein	bisque4
Peaxi162Scf00177g01228.1	Tetraspanin family protein	bisque4
Peaxi162Scf00069g01326.1	Unknown protein	bisque4
Peaxi162Scf00045g01440.1	Keratin-associated protein 10-6 isoform 1 [Theobroma cacao] ref|XP_007011486.1| Keratin-associated protein 10-6 isoform 1 [Theobroma cacao] gb|EOY29104.1| Keratin-associated protein 10-6 isoform 1 [Theobroma cacao] gb|EOY29105.1| Keratin-associated protein 10-6 isoform 1 [Theobroma cacao]	bisque4
Peaxi162Scf00107g00910.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	bisque4
Peaxi162Scf01204g00123.1	Unknown protein	bisque4
Peaxi162Scf00128g00314.1	Protein of unknown function (DUF707)	bisque4
Peaxi162Scf00284g00022.1	Peroxidase superfamily protein	bisque4
Peaxi162Scf00570g00049.1	PIG93, partial [Petunia x hybrida]	bisque4
Peaxi162Scf00535g00320.1	DA1	bisque4
Peaxi162Scf00186g00133.1	Unknown protein	bisque4
Peaxi162Scf00186g01211.1	Pyruvate kinase family protein	bisque4
Peaxi162Scf00005g00017.1	--	bisque4
Peaxi162Scf01343g00154.1	N-alpha-acetyltransferase 50	bisque4
Peaxi162Scf00103g00071.1	Blue copper protein	bisque4
Peaxi162Scf00396g00134.1	6-phosphogluconate dehydrogenase, decarboxylating 3	bisque4
Peaxi162Scf00046g00168.1	Zinc-binding dehydrogenase family protein	bisque4
Peaxi162Scf00486g00310.1	alpha/beta-Hydrolases superfamily protein	bisque4
Peaxi162Scf00983g00045.1	cinnamyl alcohol dehydrogenase 9	bisque4
Peaxi162Scf00134g02026.1	Unknown protein	bisque4
Peaxi162Scf00038g00519.1	Oxysterol-binding protein-related protein 1C	bisque4
Peaxi162Scf00075g01418.1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	bisque4
Peaxi162Scf00045g00142.1	NAD(P)-binding Rossmann-fold superfamily protein	bisque4
Peaxi162Scf00913g00019.1	cinnamyl alcohol dehydrogenase 6	bisque4
Peaxi162Scf00035g00412.1	Nuclear pore complex protein Nup96 homolog	bisque4
Peaxi162Scf00592g00446.1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	bisque4
Peaxi162Scf00117g01515.1	Cation/H(+) antiporter 18	bisque4
Peaxi162Scf00498g00523.1	serine/threonine protein kinase 2	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00953g00514.1	conserved hypothetical protein [Ricinus communis] gb|EEF44617.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00089g01857.1	UDP-N-acetylglucosamine--N-acetylmuramyl-pyrophosphoryl-undecaprenol N-acetylglucosamine transferase isoform 1 [Theobroma cacao]	bisque4
Peaxi162Scf00372g01015.1	serine carboxypeptidase-like 48	bisque4
Peaxi162Scf00268g00613.1	conserved hypothetical protein [Ricinus communis] ref|XP_004144901.1| PREDICTED: uncharacterized protein LOC101221471 [Cucumis sativus] ref|XP_004153271.1| PREDICTED: uncharacterized protein LOC101206100 [Cucumis sativus] ref|XP_004161015.1| PREDICTED: uncharacterized protein LOC101226661 [Cucumis sativus] gb|EEF48006.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00137g00107.1	ARF-GAP domain 5	bisque4
Peaxi162Scf00011g00157.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00523g00015.1	aspartate aminotransferase 3	bisque4
Peaxi162Scf00415g00526.1	RNA-binding (RRM/RBD/RNP motifs) family protein	bisque4
Peaxi162Scf00763g00428.1	Copper amine oxidase family protein	bisque4
Peaxi162Scf00086g00615.1	conserved hypothetical protein [Ricinus communis] gb|EEF33266.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00179g00097.1	copper transporter 1	bisque4
Peaxi162Scf01223g00009.1	RHOMBOID-like protein 12	bisque4
Peaxi162Scf00026g02614.1	Isocitrate dehydrogenase [NAD] subunit 2, mitochondrial	bisque4
Peaxi162Scf00638g00210.1	Heavy metal transport/detoxification superfamily protein	bisque4
Peaxi162Scf00215g00053.1	Alpha-1,4-glucan-protein synthase family protein	bisque4
Peaxi162Scf00157g00417.1	PHD type transcription factor with transmembrane domain protein [Arabidopsis thaliana] gb|AED93945.1| DNA binding and zinc-finger domain-containing protein [Arabidopsis thaliana]	bisque4
Peaxi162Scf00083g01118.1	Lactoylglutathione lyase / glyoxalase I family protein	bisque4
Peaxi162Scf00253g00321.1	NAC domain protein,	bisque4
Peaxi162Scf00745g00839.1	Major facilitator superfamily protein	bisque4
Peaxi162Scf00748g00111.1	COP1-interacting protein 7	bisque4
Peaxi162Scf00102g00108.1	C2H2-like zinc finger protein	bisque4
Peaxi162Scf00198g00148.1	MLP-like protein 28	bisque4
Peaxi162Scf00053g00524.1	Alternative oxidase 3, mitochondrial	bisque4
Peaxi162Scf00028g00206.1	Peroxidase superfamily protein	bisque4
Peaxi162Scf02113g00007.1	Calcium-binding EF-hand family protein	bisque4
Peaxi162Scf00406g00134.1	Glutelin type-A 1 [Morus notabilis]	bisque4
Peaxi162Scf00765g00133.1	actin-11	bisque4
Peaxi162Scf00102g01416.1	IQ-domain 22	bisque4
Peaxi162Scf00373g00139.1	Plant protein of unknown function (DUF946)	bisque4
Peaxi162Scf00212g00310.1	Peroxidase superfamily protein	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00039g00429.1	Lung seven transmembrane receptor family protein	bisque4
Peaxi162Scf00527g00516.1	Coatomer, alpha subunit	bisque4
Peaxi162Scf00002g01611.1	conserved hypothetical protein [Ricinus communis] gb|EEF33467.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00078g01015.1	Sucrose-phosphate synthase family protein	bisque4
Peaxi162Scf00668g00445.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00959g00046.1	sulfotransferase 16	bisque4
Peaxi162Scf00015g00327.1	sulfate transporter 3	bisque4
Peaxi162Scf00153g01319.1	Zinc finger, C3HC4 type (RING finger) family protein	bisque4
Peaxi162Scf00084g00027.1	SBP family protein [Theobroma cacao] gb|EOY24121.1| SBP family protein [Theobroma cacao]	bisque4
Peaxi162Scf00270g00009.1	Protein of unknown function (DUF1666)	bisque4
Peaxi162Scf00141g00334.1	Yippee family putative zinc-binding protein	bisque4
Peaxi162Scf00619g00113.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00342g00113.1	myb domain protein 33	bisque4
Peaxi162Scf00209g00122.1	Oxidoreductase family protein	bisque4
Peaxi162Scf00175g00122.1	ureide permease 2	bisque4
Peaxi162Scf00222g00823.1	RING/FYVE/PHD zinc finger superfamily protein	bisque4
Peaxi162Scf01068g00019.1	Late cornified envelope protein 1E [Theobroma cacao] gb|EOX96360.1| Late cornified envelope protein 1E [Theobroma cacao]	bisque4
Peaxi162Scf00126g01127.1	Calmodulin-binding transcription activator 2	bisque4
Peaxi162Scf00658g00419.1	Mannose-1-phosphate guanyltransferase	bisque4
Peaxi162Scf00579g00049.1	Protein SUPPRESSOR OF GENE SILENCING 3	bisque4
Peaxi162Scf00406g00236.1	UDP-glucuronic acid decarboxylase 3	bisque4
Peaxi162Scf00944g00145.1	Inositol oxygenase 2	bisque4
Peaxi162Scf00129g00823.1	GPI ethanolamine phosphate transferase 1	bisque4
Peaxi162Scf00235g00085.1	Unknown protein	bisque4
Peaxi162Scf00232g00517.1	ubiquitin-conjugating enzyme 28	bisque4
Peaxi162Scf00033g01724.1	Acyl carrier protein 1, chloroplastic	bisque4
Peaxi162Scf00015g00623.1	Binding-like protein isoform 4 [Theobroma cacao] gb|EOY20913.1| Binding-like protein isoform 4 [Theobroma cacao]	bisque4
Peaxi162Scf00349g00711.1	calmodulin-binding family protein	bisque4
Peaxi162Scf00004g04126.1	Hydroxymethylglutaryl-CoA lyase, mitochondrial	bisque4
Peaxi162Scf00833g00531.1	Cysteine proteinases superfamily protein	bisque4
Peaxi162Scf00485g00037.1	Developmental regulator, ULTRAPETALA	bisque4
Peaxi162Scf00517g90028.1	Protein of Unknown Function (DUF239)	bisque4
Peaxi162Scf00493g00219.1	conserved hypothetical protein [Ricinus communis] gb|EEF44617.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00076g01055.1	galactose-1-phosphate uridyl transferase-like protein [Arabidopsis thaliana]	bisque4
Peaxi162Scf00016g02427.1	1,2-alpha-L-fucosidases	bisque4
Peaxi162Scf00111g00125.1	NADH-ubiquinone oxidoreductase chain 5	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00479g00005.1	Transcription factor CYCLOIDEA	bisque4
Peaxi162Scf00223g01515.1	LisH dimerisation motif	bisque4
Peaxi162Scf00152g01526.1	Lactoylglutathione lyase / glyoxalase I family protein	bisque4
Peaxi162Scf00122g02417.1	RB1-inducible coiled-coil protein 1, putative isoform 2 [Theobroma cacao] gb|EOX90658.1| RB1-inducible coiled-coil protein 1, putative isoform 2 [Theobroma cacao]	bisque4
Peaxi162Scf00945g00123.1	Maternal effect embryo arrest 59, putative isoform 1 [Theobroma cacao] gb|EOX91570.1| Maternal effect embryo arrest 59, putative isoform 1 [Theobroma cacao]	bisque4
Peaxi162Scf00119g00722.1	Unknown protein	bisque4
Peaxi162Scf00081g01411.1	LOB domain-containing protein 38	bisque4
Peaxi162Scf00006g00529.1	RING/U-box superfamily protein	bisque4
Peaxi162Scf00899g00310.1	Nodulin MtN3 family protein	bisque4
Peaxi162Scf01650g00038.1	conserved hypothetical protein [Ricinus communis] gb|EEF42326.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00234g00713.1	SPFH/Band 7/PHB domain-containing membrane-associated protein family	bisque4
Peaxi162Scf01650g00057.1	hydrogen peroxide-induced 1 [Nicotiana tabacum]	bisque4
Peaxi162Scf00159g00912.1	nudix hydrolase homolog 25	bisque4
Peaxi162Scf00235g00842.1	Actin binding Calponin homology (CH) domain-containing protein	bisque4
Peaxi162Scf00006g00811.1	ubiquitin-conjugating enzyme 28	bisque4
Peaxi162Scf00746g00039.1	--	bisque4
Peaxi162Scf00003g05367.1	Basic-leucine zipper (bZIP) transcription factor family protein	bisque4
Peaxi162Scf00258g00045.1	phosphatidylinositol-4-phosphate 5-kinase family protein	bisque4
Peaxi162Scf00001g00481.1	SPX (SYG1/Pho81/XPR1) domain-containing protein	bisque4
Peaxi162Scf00084g00923.1	Vacuolar protein sorting 55 (VPS55) family protein	bisque4
Peaxi162Scf00089g01235.1	Quinolinate synthase, chloroplastic	bisque4
Peaxi162Scf00739g00426.1	response regulator 17	bisque4
Peaxi162Scf00284g00317.1	--	bisque4
Peaxi162Scf00146g00319.1	Polyadenylate-binding protein 7	bisque4
Peaxi162Scf00574g00117.1	ABC transporter A family member 7	bisque4
Peaxi162Scf00944g00140.1	myb domain protein 68	bisque4
Peaxi162Scf00873g00113.1	D-aminoacyl-tRNA deacylases	bisque4
Peaxi162Scf00006g00534.1	Nucleotide/sugar transporter family protein	bisque4
Peaxi162Scf00175g00523.1	cytochrome B5 isoform A	bisque4
Peaxi162Scf00907g00120.1	Major facilitator superfamily protein	bisque4
Peaxi162Scf00013g01413.1	TRICHOME BIREFRINGENCE-LIKE 8	bisque4
Peaxi162Scf00915g00120.1	RHOMBOID-like protein 3	bisque4
Peaxi162Scf00635g00021.1	armadillo repeat only 2	bisque4
Peaxi162Scf00133g01414.1	Defence response isoform 1 [Theobroma cacao] gb|EOX92285.1| Defence response isoform 1 [Theobroma cacao]	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00450g00124.1	Basic helix-loop-helix DNA-binding superfamily protein, putative isoform 1 [Theobroma cacao] gb|EOX91461.1| Basic helix-loop-helix DNA-binding superfamily protein, putative isoform 1 [Theobroma cacao]	bisque4
Peaxi162Scf00117g00315.1	Inorganic pyrophosphatase	bisque4
Peaxi162Scf00105g01426.1	conserved hypothetical protein [Ricinus communis] gb|EEF45949.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00218g00312.1	--	bisque4
Peaxi162Scf00037g00424.1	long-chain acyl-CoA synthetase 7	bisque4
Peaxi162Scf00156g01314.1	auxin response factor 11	bisque4
Peaxi162Scf00341g00036.1	Eukaryotic aspartyl protease family protein	bisque4
Peaxi162Scf00885g00320.1	--	bisque4
Peaxi162Scf01002g00114.1	Pollen Ole e 1 allergen and extensin family protein [Theobroma cacao] gb|EOY03810.1| Pollen Ole e 1 allergen and extensin family protein [Theobroma cacao]	bisque4
Peaxi162Scf00740g00554.1	SAUR-like auxin-responsive protein family	bisque4
Peaxi162Scf00091g00621.1	ferredoxin 3	bisque4
Peaxi162Scf00067g00626.1	sodium hydrogen exchanger 2	bisque4
Peaxi162Scf00081g00089.1	RNA binding protein, putative [Ricinus communis] gb|EEF52315.1| RNA binding protein, putative [Ricinus communis]	bisque4
Peaxi162Scf00166g01042.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00753g00333.1	ubiquitin-associated (UBA)/TS-N domain-containing protein	bisque4
Peaxi162Scf00481g00623.1	conserved hypothetical protein [Ricinus communis] gb|EEF41898.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00228g00514.1	Peroxidase superfamily protein	bisque4
Peaxi162Scf00344g00159.1	Unknown protein	bisque4
Peaxi162Scf01112g00011.1	Inositol monophosphatase - like protein [Arabidopsis thaliana]	bisque4
Peaxi162Scf00016g03013.1	Unknown protein	bisque4
Peaxi162Scf00100g00094.1	Unknown protein	bisque4
Peaxi162Scf00017g02910.1	conserved hypothetical protein [Ricinus communis] gb|EEF52000.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00553g00216.1	TFIIB zinc-binding protein	bisque4
Peaxi162Scf00595g00425.1	ATP-dependent Clp protease proteolytic subunit 1	bisque4
Peaxi162Scf00269g01526.1	methyl esterase 12	bisque4
Peaxi162Scf00316g00340.1	aspartic proteinase A1	bisque4
Peaxi162Scf00579g00211.1	spermidine synthase 1	bisque4
Peaxi162Scf00015g00049.1	Regulator of Vps4 activity in the MVB pathway protein	bisque4
Peaxi162Scf00345g01028.1	F-box family protein	bisque4
Peaxi162Scf00020g01714.1	Bifunctional pinoresinol-lariciresinol reductase	bisque4
Peaxi162Scf00620g00621.1	Protein STAY-GREEN, chloroplastic	bisque4
Peaxi162Scf00421g00323.1	nodulin MtN21 /EamA-like transporter family protein	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00297g00712.1	Fatty acid hydroxylase superfamily	bisque4
Peaxi162Scf00550g00448.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00819g00027.1	Peptidyl-prolyl cis-trans isomerase G isoform 3 [Theobroma cacao] gb|EOY14010.1	bisque4
Peaxi162Scf00585g01018.1	Eukaryotic translation initiation factor 1A	bisque4
Peaxi162Scf00071g00528.1	translationally controlled tumor protein	bisque4
Peaxi162Scf00620g00814.1	Glucose-methanol-choline (GMC) oxidoreductase family protein	bisque4
Peaxi162Scf00332g00444.1	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	bisque4
Peaxi162Scf01011g00005.1	phosphoglucosamine mutase family protein	bisque4
Peaxi162Scf00548g00650.1	Mitochondrial outer membrane protein porin 2	bisque4
Peaxi162Scf01012g00233.1	Eukaryotic translation initiation factor 1A	bisque4
Peaxi162Scf00314g00922.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00046g00161.1	Zinc-binding dehydrogenase family protein	bisque4
Peaxi162Scf00385g00068.1	bZIP transcription factor 60	bisque4
Peaxi162Scf00029g00077.1	Cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG7	bisque4
Peaxi162Scf00089g00119.1	Copine family protein 2	bisque4
Peaxi162Scf00051g00621.1	Transmembrane emp24 domain-containing protein p24delta3	bisque4
Peaxi162Scf00128g01750.1	GRAM domain family protein	bisque4
Peaxi162Scf00003g02440.1	caffeoyl-CoA 3-O-methyltransferase	bisque4
Peaxi162Scf00083g00022.1	Heterogeneous nuclear ribonucleoprotein U-like protein 1	bisque4
Peaxi162Scf00073g01420.1	RNA-dependent RNA polymerase 6	bisque4
Peaxi162Scf00276g00211.1	ubiquitin conjugating enzyme 9	bisque4
Peaxi162Scf01312g00069.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	bisque4
Peaxi162Scf00919g00313.1	NAD(P)-binding Rossmann-fold superfamily protein	bisque4
Peaxi162Scf00037g01117.1	Rubber elongation factor protein (REF)	bisque4
Peaxi162Scf00323g00620.1	tyrosine-rich hydroxyproline-rich glycoprotein, partial [Petroselinum crispum]	bisque4
Peaxi162Scf00064g01131.1	conserved hypothetical protein [Ricinus communis] gb|EEF52842.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00367g00410.1	Glycine cleavage system H protein 2, mitochondrial	bisque4
Peaxi162Scf01133g00024.1	E3 ubiquitin-protein ligase CHIP	bisque4
Peaxi162Scf01160g00021.1	ATP-citrate synthase	bisque4
Peaxi162Scf00140g00212.1	Mannan endo-1,4-beta-mannosidase 7	bisque4
Peaxi162Scf00420g00642.1	Pollen Ole e 1 allergen and extensin family protein	bisque4
Peaxi162Scf00016g02234.1	cysteine proteinase1	bisque4
Peaxi162Scf00328g00219.1	Eukaryotic aspartyl protease family protein	bisque4
Peaxi162Scf00230g00067.1	Bax inhibitor-1 family protein	bisque4
Peaxi162Scf00241g00022.1	Ribosomal protein L12/ ATP-dependent Clp protease adaptor protein ClpS family protein	bisque4
Peaxi162Scf00469g00218.1	glycerol-3-phosphatase 1	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00772g00022.1	vacuolar protein sorting 55 [Populus trichocarpa] gb|ERP63405.1| vacuolar protein sorting 55 [Populus trichocarpa]	bisque4
Peaxi162Scf00073g02123.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00130g00832.1	Unknown protein	bisque4
Peaxi162Scf00276g00413.1	Sulfite exporter TauE/SafE family protein	bisque4
Peaxi162Scf00041g01016.1	tetratricopetide-repeat thioredoxin-like 3	bisque4
Peaxi162Scf00074g00435.1	trehalase 1	bisque4
Peaxi162Scf00321g00514.1	RING/U-box superfamily protein	bisque4
Peaxi162Scf00753g00338.1	myb domain protein 62	bisque4
Peaxi162Scf00342g00929.1	Pentatricopeptide repeat (PPR) superfamily protein	bisque4
Peaxi162Scf00304g00091.1	glutamate dehydrogenase 2	bisque4
Peaxi162Scf00903g00213.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein	bisque4
Peaxi162Scf00096g01718.1	myb domain protein 68	bisque4
Peaxi162Scf00114g00053.1	Histone H2B.10	bisque4
Peaxi162Scf00021g01218.1	thioesterase family protein [Populus trichocarpa] gb|EEE81840.1| thioesterase family protein [Populus trichocarpa]	bisque4
Peaxi162Scf00035g00711.1	SNARE associated Golgi protein family	bisque4
Peaxi162Scf00856g00014.1	Nuclear pore complex protein Nup96 homolog	bisque4
Peaxi162Scf00016g00933.1	NAD(P)H dehydrogenase (quinone)	bisque4
Peaxi162Scf00110g01814.1	Enolase	bisque4
Peaxi162Scf00145g00144.1	U-box domain-containing protein 14	bisque4
Peaxi162Scf00003g01333.1	Receptor expression-enhancing protein 5	bisque4
Peaxi162Scf00003g05340.1	conserved hypothetical protein [Ricinus communis] gb|EEF28605.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00689g00221.1	HCO3- transporter family	bisque4
Peaxi162Scf00818g00536.1	conserved hypothetical protein [Ricinus communis] gb|EEF36870.1| conserved hypothetical protein [Ricinus communis]	bisque4
Peaxi162Scf00146g00068.1	Ubiquitin-conjugating enzyme/RWD-like protein	bisque4
Peaxi162Scf00821g00039.1	Rhamnogalacturonate lyase family protein	bisque4
Peaxi162Scf00961g00116.1	--	bisque4
Peaxi162Scf00257g01610.1	Calcium-binding EF-hand family protein	bisque4
Peaxi162Scf00037g01116.1	40S ribosomal protein S10-3	bisque4
Peaxi162Scf00020g02022.1	Unknown protein	bisque4
Peaxi162Scf00016g02023.1	caffeoyl-CoA 3-O-methyltransferase	bisque4
Peaxi162Scf00481g00063.1	Zinc transporter 2	bisque4
Peaxi162Scf00041g01316.1	pumilio 7	bisque4
Peaxi162Scf01183g00136.1	Isocitrate dehydrogenase [NAD] subunit 1, mitochondrial	bisque4
Peaxi162Scf02113g00020.1	Protein kinase superfamily protein	bisque4
Peaxi162Scf00276g00065.1	DNA repair protein XRCC3 homolog	bisque4
Peaxi162Scf00722g00412.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein	bisque4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00170g00717.1	Phospho-N-acetylmuramoyl-pentapeptide-transferase homolog	bisque4
Peaxi162Scf00601g00035.1	nodulin MtN21 /EamA-like transporter family protein	bisque4
Peaxi162Scf00338g00422.1	carbonic anhydrase 2	bisque4
Peaxi162Scf00327g00052.1	Glutaredoxin-C6	bisque4
Peaxi162Scf00164g00106.1	Galactokinase	bisque4
Peaxi162Scf00331g01018.1	calcium-dependent protein kinase 15	bisque4
Peaxi162Scf00168g01738.1	Unknown protein	bisque4
Peaxi162Scf00444g00616.1	SIGNAL PEPTIDE PEPTIDASE-LIKE 1	bisque4
Peaxi162Scf00570g00314.1	Unknown protein	bisque4
Peaxi162Scf00825g00027.1	Protein yippee-like	bisque4
Peaxi162Scf00550g00548.1	CCT motif family protein	bisque4
Peaxi162Scf00650g00221.1	sugar transport protein [Coffea canephora]	bisque4
Peaxi162Scf00739g00125.1	potassium transporter 2	bisque4
Peaxi162Scf00463g00416.1	Metal tolerance protein 4	bisque4
Peaxi162Scf00213g00934.1	Protein phosphatase 2C family protein	deepskyblue4
Peaxi162Scf00699g00638.1	Enod93 protein [Medicago truncatula] gb|AES58699.1| Enod93 protein [Medicago truncatula] gb|AFK42368.1| unknown [Medicago truncatula]	deepskyblue4
Peaxi162Scf00105g01011.1	Subtilase family protein	deepskyblue4
Peaxi162Scf00264g00838.1	CASP-like protein	deepskyblue4
Peaxi162Scf00703g00210.1	Octicosapeptide/Phox/Bem1p family protein	deepskyblue4
Peaxi162Scf00003g00719.1	26S protease regulatory subunit 10B homolog A	deepskyblue4
Peaxi162Scf00031g00823.1	LisH and RanBPM domains containing protein	deepskyblue4
Peaxi162Scf01112g00112.1	TBP-associated factor 7	deepskyblue4
Peaxi162Scf00957g00018.1	proline transporter 2	deepskyblue4
Peaxi162Scf00852g00012.1	Unknown protein	deepskyblue4
Peaxi162Scf00074g02027.1	product [Oryza sativa Japonica Group]	deepskyblue4
Peaxi162Scf00538g00229.1	--	deepskyblue4
Peaxi162Scf01039g00023.1	cullin 1	deepskyblue4
Peaxi162Scf01161g00322.1	Transmembrane amino acid transporter family protein	deepskyblue4
Peaxi162Scf00038g00920.1	Protein transport protein SEC13	deepskyblue4
Peaxi162Scf00769g00416.1	alpha/beta-Hydrolases superfamily protein	deepskyblue4
Peaxi162Scf00014g00085.1	Proteasome subunit alpha type-4	deepskyblue4
Peaxi162Scf00870g00011.1	Unknown protein	deepskyblue4
Peaxi162Scf00404g00061.1	conserved hypothetical protein [Ricinus communis] gb|EEF33671.1| conserved hypothetical protein [Ricinus communis]	deepskyblue4
Peaxi162Scf01312g00070.1	Protein kinase superfamily protein	deepskyblue4
Peaxi162Scf00110g00146.1	BURP domain-containing protein	deepskyblue4
Peaxi162Scf00463g00127.1	Mediator of RNA polymerase II transcription subunit 7b	deepskyblue4
Peaxi162Scf00620g00918.1	Cox19-like CHCH family protein	deepskyblue4
Peaxi162Scf00746g00057.1	TRAF-like family protein	deepskyblue4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf00620g00011.1	Zinc finger A20 and AN1 domain-containing stress-associated protein 1	deepskyblue4
Peaxi162Scf00114g00811.1	Charged multivesicular body protein 3	deepskyblue4
Peaxi162Scf00681g00332.1	GRAM domain-containing protein / ABA-responsive protein-related	deepskyblue4
Peaxi162Scf00067g01924.1	cullin 1	deepskyblue4
Peaxi162Scf00585g00318.1	Mechanosensitive ion channel protein	deepskyblue4
Peaxi162Scf00332g00755.1	Unknown protein	deepskyblue4
Peaxi162Scf00380g00710.1	GDT1-like protein 4	deepskyblue4
Peaxi162Scf01414g00010.1	Exostosin family protein	deepskyblue4
Peaxi162Scf00303g00513.1	Rer1 family protein	deepskyblue4
Peaxi162Scf00496g00003.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	deepskyblue4
Peaxi162Scf00136g00620.1	ferrochelatase 1	deepskyblue4
Peaxi162Scf40203g00005.1	mitochondrial acyl carrier protein 2	deepskyblue4
Peaxi162Scf00009g00624.1	Galactosyltransferase family protein	deepskyblue4
Peaxi162Scf00094g00210.1	auxin response factor 1	deepskyblue4
Peaxi162Scf00128g00020.1	alpha/beta-Hydrolases superfamily protein	deepskyblue4
Peaxi162Scf01944g00014.1	Bifunctional protein FolD 2	deepskyblue4
Peaxi162Scf00329g00410.1	alpha-galactosidase 1	deepskyblue4
Peaxi162Scf00000g00397.1	Unknown protein	deepskyblue4
Peaxi162Scf01053g00210.1	UPF0510 protein INM02 [Theobroma cacao] gb|EOY09739.1| UPF0510 protein INM02 [Theobroma cacao]	deepskyblue4
Peaxi162Scf00235g00323.1	AUTOPHAGY 8E	deepskyblue4
Peaxi162Scf00146g00132.1	Disease resistance protein (CC-NBS-LRR class) family	deepskyblue4
Peaxi162Scf00073g01423.1	conserved hypothetical protein [Ricinus communis] gb|EEF52477.1| conserved hypothetical protein [Ricinus communis]	deepskyblue4
Peaxi162Scf00276g00312.1	Branched-chain-amino-acid aminotransferase-like protein 3 [Morus notabilis]	deepskyblue4
Peaxi162Scf00251g00053.1	RING/U-box superfamily protein	deepskyblue4
Peaxi162Scf00481g00628.1	BTB/POZ domain-containing protein	deepskyblue4
Peaxi162Scf00454g00524.1	O-fucosyltransferase family protein	deepskyblue4
Peaxi162Scf00815g00234.1	60S ribosomal protein L13-2	deepskyblue4
Peaxi162Scf00698g00417.1	14 kDa zinc-binding protein	deepskyblue4
Peaxi162Scf00001g00284.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	deepskyblue4
Peaxi162Scf00519g00710.1	Diphosphomevalonate decarboxylase	deepskyblue4
Peaxi162Scf00089g00637.1	C2H2-like zinc finger protein	deepskyblue4
Peaxi162Scf00009g00333.1	SPFH/Band 7/PHB domain-containing membrane-associated protein family	deepskyblue4
Peaxi162Scf00282g00225.1	Peptide-N(4)-(N-acetyl-beta-glucosaminyl)asparagine amidase	deepskyblue4
Peaxi162Scf00067g02427.1	glutathione S-transferase F4	deepskyblue4
Peaxi162Scf00330g00715.1	DHHC-type zinc finger family protein	deepskyblue4
Peaxi162Scf00052g00820.1	Protein kinase superfamily protein	deepskyblue4
Peaxi162Scf00007g01610.1	Calcium-binding EF-hand family protein	deepskyblue4

Table APP-3-1 (cont'd)

GeneID	Functional description	Module
Peaxi162Scf01716g00006.1	catalase 2	deepskyblue4
Peaxi162Scf00007g02541.1	2-nonaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase [Theobroma cacao] gb|EOY12769.1| 2-nonaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase [Theobroma cacao]	deepskyblue4
Peaxi162Scf00418g00732.1	cytochrome P450, putative [Ricinus communis] gb|EEF40808.1| cytochrome P450, putative [Ricinus communis]	orange1
Peaxi162Scf00332g00337.1	DCN1-like protein 4	orange1
Peaxi162Scf00472g00824.1	RING/U-box superfamily protein	orange1
Peaxi162Scf00263g01423.1	Nodulin MtN3 family protein	orange1
Peaxi162Scf00166g00424.1	Unknown protein	orange1
Peaxi162Scf00129g01228.1	Proteasome assembly chaperone [Medicago truncatula] ref|XP_003630456.1| Proteasome assembly chaperone [Medicago truncatula] gb|AES72307.1|	orange1
Peaxi162Scf00038g02442.1	ATP synthase subunit epsilon, mitochondrial	orange1
Peaxi162Scf00141g00138.1	conserved hypothetical protein [Ricinus communis] gb|EEF37990.1| conserved hypothetical protein [Ricinus communis]	orange1
Peaxi162Scf01003g00012.1	DWNN domain, a CCHC-type zinc finger	orange1
Peaxi162Scf00981g00019.1	allantoate amidohydrolase	orange1
Peaxi162Scf01015g00114.1	Pyrimidine-specific ribonucleoside hydrolase RihA	orange1

Table APP-3-2: Go enrichment analysis of differentially expressed genes and modules identified in the WGCNA analysis (p <0.05).

GO ID	Term	Category	Count	p-value
Downregulated genes in the slow lines (Slow vs. Fast pooled comparison)				
GO:0016021	integral component of membrane	C	92	7.4e-14
GO:0031224	intrinsic component of membrane	C	95	8.8e-14
GO:0044425	membrane part	C	99	5.6e-08
GO:0016020	membrane	C	193	9.5e-06
GO:0090406	pollen tube	C	6	0.0024
GO:0042995	cell projection	C	6	0.0081
GO:0048226	Casparian strip	C	2	0.0165
GO:0044426	cell wall part	C	2	0.0241
GO:0044462	external encapsulating structure part	C	2	0.0425
GO:0008289	lipid binding	F	26	3.3e-06
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	F	51	4.3e-06
GO:0016798	hydrolase activity, acting on glycosyl bonds	F	52	1.3e-05
GO:0015299	solute:proton antiporter activity	F	10	7.6e-05
GO:0015298	solute:cation antiporter activity	F	10	0.00012
GO:0016641	oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as acceptor	F	5	0.00043
GO:0016638	oxidoreductase activity, acting on the CH-NH2 group of donors	F	6	0.00046
GO:0046524	sucrose-phosphate synthase activity	F	3	0.00083
GO:0005215	transporter activity	F	83	0.00094
GO:0008131	primary amine oxidase activity	F	4	0.00128
GO:0015198	oligopeptide transporter activity	F	4	0.00128
GO:0015197	peptide transporter activity	F	4	0.00204
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	F	51	0.00351
GO:0033907	beta-D-fucosidase activity	F	2	0.00361
GO:0090439	tetraketide alpha-pyrone synthase activity	F	2	0.00361
GO:0050113	inositol oxygenase activity	F	3	0.00378
GO:0015297	antiporter activity	F	16	0.00427
GO:0005506	iron ion binding	F	50	0.00703

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0015291	secondary active transmembrane transporter activity	F	20	0.00844
GO:0016843	amine-lyase activity	F	4	0.0104
GO:0016844	strictosidine synthase activity	F	4	0.0104
GO:0004044	amidophosphoribosyltransferase activity	F	2	0.0104
GO:0080083	beta-gentiobiose beta-glucosidase activity	F	2	0.0104
GO:0016887	ATPase activity	F	29	0.01227
GO:0016157	sucrose synthase activity	F	3	0.01891
GO:0015086	cadmium ion transmembrane transporter activity	F	2	0.01998
GO:0022804	active transmembrane transporter activity	F	27	0.02391
GO:0004650	polygalacturonase activity	F	7	0.02903
GO:0022892	substrate-specific transporter activity	F	50	0.03082
GO:0008422	beta-glucosidase activity	F	2	0.03198
GO:0030414	peptidase inhibitor activity	F	6	0.03835
GO:0061134	peptidase regulator activity	F	6	0.03835
GO:0022857	transmembrane transporter activity	F	57	0.0394
GO:0008061	chitin binding	F	3	0.0394
GO:0016298	lipase activity	F	7	0.04567
GO:0004618	phosphoglycerate kinase activity	F	2	0.04608
GO:0015385	sodium:proton antiporter activity	F	2	0.04608
GO:0017089	glycolipid transporter activity	F	2	0.04608
GO:0051861	glycolipid binding	F	2	0.04608
GO:0015926	glucosidase activity	F	3	0.04797
GO:0010208	pollen wall assembly	P	14	1.6e-10
GO:0085029	extracellular matrix assembly	P	14	1.6e-10
GO:0010927	cellular component assembly involved in morphogenesis	P	14	4.9e-10
GO:0010584	pollen exine formation	P	11	1.7e-08
GO:0030198	extracellular matrix organization	P	15	2.5e-08
GO:0043062	extracellular structure organization	P	15	2.5e-08
GO:0080110	sporopollenin biosynthetic process	P	6	2.6e-07
GO:0006869	lipid transport	P	14	5.9e-05

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0048646	anatomical structure formation involved in morphogenesis	P	19	0.00019
GO:0006857	oligopeptide transport	P	10	0.00021
GO:0015833	peptide transport	P	10	0.00021
GO:0042886	amide transport	P	11	0.00023
GO:0010876	lipid localization	P	14	0.00025
GO:0045471	response to ethanol	P	3	0.00075
GO:0006885	regulation of pH	P	4	0.00114
GO:0045229	external encapsulating structure organization	P	26	0.0016
GO:0030638	polyketide metabolic process	P	2	0.00339
GO:0030639	polyketide biosynthetic process	P	2	0.00339
GO:0019310	inositol catabolic process	P	3	0.00345
GO:0055085	transmembrane transport	P	65	0.00406
GO:0042542	response to hydrogen peroxide	P	8	0.00533
GO:0046685	response to arsenic-containing substance	P	4	0.00536
GO:0005975	carbohydrate metabolic process	P	71	0.00658
GO:0044765	single-organism transport	P	105	0.00734
GO:0032989	cellular component morphogenesis	P	24	0.0084
GO:0055067	monovalent inorganic cation homeostasis	P	5	0.00928
GO:0051098	regulation of binding	P	2	0.00978
GO:0006541	glutamine metabolic process	P	5	0.01117
GO:1902578	single-organism localization	P	105	0.0133
GO:0055114	oxidation-reduction process	P	130	0.01472
GO:0009555	pollen development	P	19	0.0167
GO:0006414	translational elongation	P	6	0.01672
GO:0009686	gibberellin biosynthetic process	P	3	0.01738
GO:0010092	specification of organ identity	P	3	0.01738
GO:0010093	specification of floral organ identity	P	3	0.01738
GO:0010262	somatic embryogenesis	P	2	0.01881
GO:0015691	cadmium ion transport	P	2	0.01881
GO:0034755	iron ion transmembrane transport	P	2	0.01881

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0048317	seed morphogenesis	P	2	0.01881
GO:0098771	inorganic ion homeostasis	P	9	0.01987
	amine metabolic process	P	10	0.0221
GO:0046174	polyol catabolic process	P	3	0.02288
GO:0050801	ion homeostasis	P	10	0.02382
GO:0071705	nitrogen compound transport	P	17	0.02416
GO:0048444	floral organ morphogenesis	P	5	0.02457
GO:0048563	post-embryonic organ morphogenesis	P	5	0.02457
GO:0048449	floral organ formation	P	4	0.02626
GO:0055080	cation homeostasis	P	8	0.02678
GO:0006810	transport	P	127	0.02811
GO:0006020	inositol metabolic process	P	3	0.02921
GO:0016102	diterpenoid biosynthetic process	P	3	0.02921
GO:0009691	cytokinin biosynthetic process	P	2	0.03015
GO:0035264	multicellular organism growth	P	2	0.03015
GO:0043090	amino acid import	P	2	0.03015
GO:0009739	response to gibberellin	P	8	0.03438
GO:0051234	establishment of localization	P	127	0.03634
GO:0009685	gibberellin metabolic process	P	3	0.03637
GO:0009113	purine nucleobase biosynthetic process	P	2	0.0435
GO:0046836	glycolipid transport	P	2	0.0435
GO:0016998	cell wall macromolecule catabolic process	P	3	0.04433
GO:0046164	alcohol catabolic process	P	3	0.04433
Upregulated genes in the slow lines (Slow vs. Fast pooled comparison)				
GO:0042752	regulation of circadian rhythm	P	3	0.00094
GO:0042753	positive regulation of circadian rhythm	P	2	0.00096
GO:0009649	entrainment of circadian clock	P	2	0.00134
GO:0008152	metabolic process	P	90	0.00359
GO:0042742	defense response to bacterium	P	6	0.00505

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0055114	oxidation-reduction process	P	25	0.00727
GO:0009617	response to bacterium	P	6	0.01217
GO:0009605	response to external stimulus	P	11	0.016
GO:0007623	circadian rhythm	P	3	0.01938
GO:0048511	rhythmic process	P	3	0.01938
GO:0010017	red or far-red light signaling pathway	P	2	0.02312
GO:0000272	polysaccharide catabolic process	P	2	0.02464
GO:0071489	cellular response to red or far red light	P	2	0.0262
GO:0051241	negative regulation of multicellular organismal process	P	3	0.02744
GO:0009733	response to auxin	P	5	0.03384
GO:0051093	negative regulation of developmental process	P	3	0.034
GO:0009627	systemic acquired resistance	P	2	0.03458
GO:0009926	auxin polar transport	P	2	0.03636
GO:0006355	regulation of transcription, DNA-templated	P	15	0.04692
GO:1903506	regulation of nucleic acid-templated transcription	P	15	0.04751
GO:2001141	regulation of RNA biosynthetic process	P	15	0.04751
GO:0010114	response to red light	P	2	0.04779
GO:0003700	transcription factor activity, sequence-specific DNA binding	F	14	0.00024
GO:0001071	nucleic acid binding transcription factor activity	F	14	0.00024
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	F	12	0.00195
GO:0005506	iron ion binding	F	12	0.00215
GO:0043565	sequence-specific DNA binding	F	9	0.00267
GO:0020037	heme binding	F	13	0.00292
GO:0046906	tetrapyrrole binding	F	13	0.00326
GO:0016161	beta-amylase activity	F	2	0.00524
GO:0016491	oxidoreductase activity	F	25	0.00873
GO:0016160	amylase activity	F	2	0.0106
GO:0051213	dioxygenase activity	F	3	0.01575
GO:0030246	carbohydrate binding	F	5	0.02211

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	F	8	0.02217
GO:0016758	transferase activity, transferring hexosyl groups	F	8	0.02642
GO:0016798	hydrolase activity, acting on glycosyl bonds	F	8	0.03122
GO:0042803	protein homodimerization activity	F	3	0.04202
GO:0003824	catalytic activity	F	79	0.04551
GO:0019829	cation-transporting ATPase activity	F	2	0.04867
WGCNA module: Bisque4				
GO:0005777	peroxisome	C	23	3e-07
GO:0042579	microbody	C	23	3e-07
GO:0016021	integral component of membrane	C	90	2e-06
GO:0031224	intrinsic component of membrane	C	90	1.7e-05
GO:0012505	endomembrane system	C	78	0.00024
GO:0005783	endoplasmic reticulum	C	41	0.00045
GO:0005886	plasma membrane	C	86	0.00144
GO:0044425	membrane part	C	100	0.00567
GO:0003824	catalytic activity	F	501	3.7e-08
GO:0048037	cofactor binding	F	48	6.4e-07
GO:0050662	coenzyme binding	F	36	6.4e-06
GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	F	19	2.8e-05
GO:0016614	oxidoreductase activity, acting on CH-OH group of donors	F	24	6.9e-05
GO:0019787	ubiquitin-like protein transferase activity	F	19	6e-04
GO:0005215	transporter activity	F	71	0.00045
GO:0004842	ubiquitin-protein transferase activity	F	19	0.00045
GO:0070011	peptidase activity, acting on L-amino acid peptides	F	41	0.00054
GO:0016787	hydrolase activity	F	161	0.00058
GO:0008233	peptidase activity	F	42	0.00071
GO:0042578	phosphoric ester hydrolase activity	F	20	0.0013
GO:0022892	substrate-specific transporter activity	F	48	0.00157
GO:0004175	endopeptidase activity	F	26	0.00169

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0043492	ATPase activity, coupled to movement of substances	F	13	0.00237
GO:0016798	hydrolase activity, acting on glycosyl bonds	F	37	0.00238
GO:0022857	transmembrane transporter activity	F	53	0.00376
GO:0000287	magnesium ion binding	F	16	0.00379
GO:0046873	metal ion transmembrane transporter activity	F	15	0.0043
GO:0008324	cation transmembrane transporter activity	F	26	0.00665
GO:0015075	ion transmembrane transporter activity	F	35	0.00799
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	F	33	0.00805
GO:0016772	transferase activity, transferring phosphorus-containing groups	F	97	0.00903
GO:0044281	small molecule metabolic process	P	109	2.5e-08
GO:0006732	coenzyme metabolic process	P	27	1.7e-06
GO:0009108	coenzyme biosynthetic process	P	17	7.6e-06
GO:0006733	oxidoreduction coenzyme metabolic process	P	17	1.6e-05
GO:0006970	response to osmotic stress	P	38	2.1e-05
GO:0044763	single-organism cellular process	P	292	3.2e-05
GO:0046496	nicotinamide nucleotide metabolic process	P	15	5.3e-05
GO:0019362	pyridine nucleotide metabolic process	P	15	6e-05
GO:0044712	single-organism catabolic process	P	29	2e-04
GO:0043436	oxoacid metabolic process	P	66	0.00011
GO:0043436	oxoacid metabolic process	P	66	0.00011
GO:0006082	organic acid metabolic process	P	66	0.00012
GO:0072524	pyridine-containing compound metabolic process	P	15	0.00013
GO:0019752	carboxylic acid metabolic process	P	63	0.00025
GO:0051186	cofactor metabolic process	P	28	0.00028
GO:0048437	floral organ development	P	18	0.00034
GO:0006629	lipid metabolic process	P	56	0.00048
GO:0009651	response to salt stress	P	31	0.00054
GO:0044699	single-organism process	P	424	0.00058
GO:0044282	small molecule catabolic process	P	14	0.00063
GO:0016567	protein ubiquitination	P	17	0.00067

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0071704	organic substance metabolic process	P	429	0.00069
GO:0000003	reproduction	P	25	0.00087
GO:0005975	carbohydrate metabolic process	P	70	0.00122
GO:0001101	response to acid chemical	P	50	0.00139
GO:0032446	protein modification by small protein conjugation	P	17	0.00141
GO:0048569	post-embryonic organ development	P	22	0.00147
GO:0048646	anatomical structure formation involved in morphogenesis	P	16	0.00155
GO:0010243	response to organonitrogen compound	P	12	0.00164
GO:0005996	monosaccharide metabolic process	P	11	0.00164
GO:0051188	cofactor biosynthetic process	P	17	0.00165
GO:0044723	single-organism carbohydrate metabolic process	P	35	0.00175
GO:0080134	regulation of response to stress	P	19	0.00202
GO:0019637	organophosphate metabolic process	P	36	0.0023
GO:0032504	multicellular organism reproduction	P	15	0.00255
GO:0044710	single-organism metabolic process	P	239	0.00307
GO:0031347	regulation of defense response	P	15	0.00322
GO:0009751	response to salicylic acid	P	12	0.00327
GO:0030001	metal ion transport	P	24	0.00329
GO:0009555	pollen development	P	20	0.00334
GO:1901700	response to oxygen-containing compound	P	61	0.00359
GO:1901698	response to nitrogen compound	P	17	0.00415
GO:0009117	nucleotide metabolic process	P	23	0.00519
GO:0070647	protein modification by small protein conjugation or removal	P	19	0.00546
GO:0055086	nucleobase-containing small molecule metabolic process	P	27	0.0055
GO:0046686	response to cadmium ion	P	20	0.00591
GO:0006753	nucleoside phosphate metabolic process	P	23	0.00663
GO:0006793	phosphorus metabolic process	P	112	0.00671
GO:0042221	response to chemical	P	100	0.00737
GO:0002376	immune system process	P	19	0.00762
GO:0051321	meiotic cell cycle	P	11	0.00777

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0006790	sulfur compound metabolic process	P	14	0.00798
GO:0014070	response to organic cyclic compound	P	19	0.00804
GO:0006950	response to stress	P	119	0.00808
GO:0048438	floral whorl development	P	13	0.00907
WGCNA module: Deepskyblue4				
GO:0044464	cell part	C	101	0.0114
GO:0005623	cell	C	101	0.0115
GO:0032991	macromolecular complex	C	29	0.0132
GO:0044428	nuclear part	C	13	0.0145
GO:0044424	intracellular part	C	90	0.02
GO:0031974	membrane-enclosed lumen	C	11	0.021
GO:0005739	mitochondrion	C	15	0.0264
GO:0043234	protein complex	C	20	0.0285
GO:0005622	intracellular	C	90	0.0395
GO:0030163	protein catabolic process	P	12	2.1e-06
GO:0009057	macromolecule catabolic process	P	12	0.00011
GO:1901575	organic substance catabolic process	P	14	0.00273
GO:0009056	catabolic process	P	14	0.00429
GO:0006996	organelle organization	P	16	0.0051
GO:0019538	protein metabolic process	P	45	0.0101
GO:0051641	cellular localization	P	11	0.0151
GO:0009987	cellular process	P	109	0.01612
GO:0006508	proteolysis	P	13	0.01987
GO:0043170	macromolecule metabolic process	P	70	0.02466
GO:0044238	primary metabolic process	P	88	0.0321
GO:0044267	cellular protein metabolic process	P	36	0.04522
WGCNA module: Orange1				
GO:0031090	organelle membrane	C	18	0.0041
GO:0098588	bounding membrane of organelle	C	14	0.0102
GO:0031967	organelle envelope	C	15	0.0213

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0031975	envelope	C	15	0.0218
GO:0098805	whole membrane	C	11	0.0259
GO:0065008	regulation of biological quality	P	12	0.00415
GO:0009628	response to abiotic stimulus	P	21	0.00457
GO:0010035	response to inorganic substance	P	12	0.00551
GO:0055085	transmembrane transport	P	16	0.00694
GO:0044763	single-organism cellular process	P	59	0.01007
GO:0050794	regulation of cellular process	P	32	0.01959
GO:0051234	establishment of localization	P	28	0.02055
GO:0065007	biological regulation	P	38	0.02233
GO:0051179	localization	P	28	0.02997
GO:0006810	transport	P	27	0.03114
GO:0044281	small molecule metabolic process	P	19	0.03464
WGCNA module: Black				
GO:0005730	nucleolus	C	29	6.3e-10
GO:0031981	nuclear lumen	C	38	1.4e-07
GO:0044424	intracellular part	C	283	1.6e-07
GO:0044428	nuclear part	C	45	2.3e-07
GO:0070013	intracellular organelle lumen	C	39	3.4e-07
GO:0043233	organelle lumen	C	39	3.6e-07
GO:0031974	membrane-enclosed lumen	C	39	5.1e-07
GO:0005622	intracellular	C	285	5.2e-07
GO:0005634	nucleus	C	106	1.3e-06
GO:0043231	intracellular membrane-bounded organelle	C	224	5.7e-05
GO:0043227	membrane-bounded organelle	C	224	6.1e-05
GO:0005623	cell	C	302	3e-04
GO:0044464	cell part	C	301	0.00053
GO:0043229	intracellular organelle	C	237	0.00056
GO:0043226	organelle	C	237	0.00061
GO:0005739	mitochondrion	C	39	0.00879

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0043228	non-membrane-bounded organelle	C	51	0.01176
GO:0043232	intracellular non-membrane-bounded organelle	C	51	0.01176
GO:0005737	cytoplasm	C	203	0.03667
GO:0009536	plastid	C	76	0.04072
GO:0003723	RNA binding	F	50	1.5e-11
GO:0005515	protein binding	F	217	6.9e-08
GO:0005488	binding	F	453	6.1e-05
GO:0016779	nucleotidyltransferase activity	F	15	0.00032
GO:0003676	nucleic acid binding	F	132	0.0029
GO:0000166	nucleotide binding	F	129	0.00661
GO:1901265	nucleoside phosphate binding	F	129	0.00661
GO:0036094	small molecule binding	F	132	0.00664
GO:0016791	phosphatase activity	F	11	0.0084
GO:0042578	phosphoric ester hydrolase activity	F	12	0.02775
GO:0004518	nuclease activity	F	13	0.03518
GO:0017111	nucleoside-triphosphatase activity	F	25	0.04356
GO:0006396	RNA processing	P	52	6.6e-20
GO:0034470	ncRNA processing	P	26	2.4e-13
GO:0034660	ncRNA metabolic process	P	30	3e-12
GO:0006364	rRNA processing	P	17	1.1e-11
GO:0016072	rRNA metabolic process	P	17	1.9e-11
GO:0022613	ribonucleoprotein complex biogenesis	P	20	1.1e-10
GO:0042254	ribosome biogenesis	P	19	1.1e-10
GO:0006807	nitrogen compound metabolic process	P	161	1.2e-10
GO:0034641	cellular nitrogen compound metabolic process	P	145	3.9e-09
GO:0046483	heterocycle metabolic process	P	127	1.2e-08
GO:0006725	cellular aromatic compound metabolic process	P	128	2.5e-08
GO:0010467	gene expression	P	112	2.6e-08
GO:0016070	RNA metabolic process	P	93	5e-08
GO:1901360	organic cyclic compound metabolic process	P	129	6.3e-08

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0006139	nucleobase-containing compound metabolic process	P	116	1e-07
GO:0090304	nucleic acid metabolic process	P	104	1.1e-07
GO:0032502	developmental process	P	85	1.4e-07
GO:0007275	multicellular organismal development	P	77	1.7e-07
GO:0009790	embryo development	P	26	1.7e-07
GO:0044767	single-organism developmental process	P	82	3.9e-07
GO:0009793	embryo development ending in seed dormancy	P	24	4e-07
GO:0044707	single-multicellular organism process	P	77	4.5e-07
GO:0048856	anatomical structure development	P	76	5e-07
GO:0008652	cellular amino acid biosynthetic process	P	18	1.7e-06
GO:0048316	seed development	P	29	3.4e-06
GO:0009987	cellular process	P	297	5.4e-06
GO:0006397	mRNA processing	P	14	7e-06
GO:0010154	fruit development	P	29	7.3e-06
GO:0016071	mRNA metabolic process	P	17	9e-06
GO:0032501	multicellular organismal process	P	77	9.7e-06
GO:1901566	organonitrogen compound biosynthetic process	P	53	1.3e-05
GO:0016458	gene silencing	P	13	1.8e-05
GO:0006520	cellular amino acid metabolic process	P	26	2e-05
GO:0008380	RNA splicing	P	12	2.9e-05
GO:0048731	system development	P	56	3.3e-05
GO:0044260	cellular macromolecule metabolic process	P	175	5e-05
GO:0009451	RNA modification	P	12	6.5e-05
GO:1901607	alpha-amino acid biosynthetic process	P	14	6.7e-05
GO:0009791	post-embryonic development	P	47	0.00011
GO:0048608	reproductive structure development	P	39	0.00013
GO:0061458	reproductive system development	P	39	0.00013
GO:0044237	cellular metabolic process	P	225	0.00015
GO:0043170	macromolecule metabolic process	P	189	0.00016
GO:0003006	developmental process involved in reproduction	P	43	0.00018

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0010629	negative regulation of gene expression	P	17	0.00018
GO:1901605	alpha-amino acid metabolic process	P	17	0.00025
GO:0071840	cellular component organization or biogenesis	P	66	0.00048
GO:0044711	single-organism biosynthetic process	P	51	0.00059
GO:0044702	single organism reproductive process	P	38	0.00075
GO:0022414	reproductive process	P	45	0.00087
GO:0009892	negative regulation of metabolic process	P	19	0.00088
GO:0010605	negative regulation of macromolecule metabolic process	P	17	0.00131
GO:0016053	organic acid biosynthetic process	P	22	0.00145
GO:0046394	carboxylic acid biosynthetic process	P	22	0.00145
GO:1901564	organonitrogen compound metabolic process	P	59	0.00146
GO:0044283	small molecule biosynthetic process	P	26	0.00152
GO:0040029	regulation of gene expression, epigenetic	P	11	0.0016
GO:0044085	cellular component biogenesis	P	31	0.00168
GO:0048229	gametophyte development	P	17	0.00184
GO:0044763	single-organism cellular process	P	161	0.00247
GO:0009888	tissue development	P	22	0.00281
GO:0048507	meristem development	P	12	0.00312
GO:0044249	cellular biosynthetic process	P	112	0.00517
GO:1901576	organic substance biosynthetic process	P	113	0.00568
GO:0048519	negative regulation of biological process	P	24	0.0059
GO:0009058	biosynthetic process	P	118	0.00698
GO:0071704	organic substance metabolic process	P	239	0.0076
GO:0046907	intracellular transport	P	19	0.0076
GO:0033036	macromolecule localization	P	23	0.00818
GO:0008104	protein localization	P	18	0.00825
GO:0044238	primary metabolic process	P	227	0.0085
GO:0009653	anatomical structure morphogenesis	P	26	0.00873
GO:0070727	cellular macromolecule localization	P	16	0.00992
GO:0051641	cellular localization	P	23	0.01182

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0015031	protein transport	P	16	0.01197
GO:0034613	cellular protein localization	P	15	0.01254
GO:0044699	single-organism process	P	234	0.01282
GO:0014070	response to organic cyclic compound	P	12	0.0132
GO:0045184	establishment of protein localization	P	16	0.01392
GO:0006886	intracellular protein transport	P	13	0.02138
GO:0048513	organ development	P	27	0.02391
GO:0051649	establishment of localization in cell	P	19	0.03735
GO:1902582	single-organism intracellular transport	P	12	0.03767
GO:0048523	negative regulation of cellular process	P	14	0.03835
GO:0019752	carboxylic acid metabolic process	P	31	0.04036
GO:0006996	organelle organization	P	29	0.04466
GO:0009605	response to external stimulus	P	27	0.04643
GO:1901137	carbohydrate derivative biosynthetic process	P	12	0.04698
GO:0071310	cellular response to organic substance	P	16	0.04739
WGCNA module: Mistyrose2				
GO:0009507	chloroplast	C	18	0.0232
GO:0016020	membrane	C	34	0.0316
GO:0009536	plastid	C	18	0.037
GO:0004672	protein kinase activity	F	14	0.0246
GO:0003006	developmental process involved in reproduction	P	11	0.0334
GO:0022414	reproductive process	P	12	0.0368
WGCNA module: Sienna2				
GO:0005840	ribosome	C	57	2.4e-14
GO:0030529	ribonucleoprotein complex	C	61	2e-13
GO:0043228	non-membrane-bounded organelle	C	88	5.8e-12
GO:0043232	intracellular non-membrane-bounded organelle	C	88	5.8e-12
GO:0044464	cell part	C	358	1.4e-10
GO:0005623	cell	C	358	1.5e-10
GO:0044391	ribosomal subunit	C	25	2.4e-09

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0005622	intracellular	C	324	4.4e-08
GO:0044424	intracellular part	C	317	3.3e-07
GO:0043229	intracellular organelle	C	281	5.4e-07
GO:0043226	organelle	C	281	6.1e-07
GO:0005634	nucleus	C	118	9e-07
GO:0015935	small ribosomal subunit	C	12	5.5e-06
GO:0044445	cytosolic part	C	19	1.2e-05
GO:0022626	cytosolic ribosome	C	18	1.7e-05
GO:0044428	nuclear part	C	43	3.6e-05
GO:0032991	macromolecular complex	C	96	6.2e-05
GO:0015934	large ribosomal subunit	C	13	9.3e-05
GO:0070013	intracellular organelle lumen	C	36	9.6e-05
GO:0043233	organelle lumen	C	36	1e-04
GO:0031974	membrane-enclosed lumen	C	36	0.00013
GO:0031981	nuclear lumen	C	33	0.00024
GO:0005730	nucleolus	C	18	0.00306
GO:0005829	cytosol	C	64	0.0034
GO:0005654	nucleoplasm	C	14	0.02099
GO:0043231	intracellular membrane-bounded organelle	C	233	0.02983
GO:0043227	membrane-bounded organelle	C	233	0.03111
GO:0044444	cytoplasmic part	C	214	0.0354
GO:0044451	nucleoplasm part	C	12	0.03836
GO:0005737	cytoplasm	C	228	0.04499
GO:0003735	structural constituent of ribosome	F	56	3.1e-21
GO:0005198	structural molecule activity	F	56	1.9e-18
GO:0003676	nucleic acid binding	F	182	6.9e-15
GO:0003723	RNA binding	F	55	3e-14
GO:0097159	organic cyclic compound binding	F	291	1.6e-09
GO:1901363	heterocyclic compound binding	F	290	2.6e-09
GO:0036094	small molecule binding	F	150	9.4e-06

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0003700	transcription factor activity, sequence-specific DNA binding	F	39	3.2e-05
GO:0001071	nucleic acid binding transcription factor activity	F	39	3.4e-05
GO:0008135	translation factor activity, RNA binding	F	11	0.00016
GO:0000166	nucleotide binding	F	140	0.00017
GO:1901265	nucleoside phosphate binding	F	140	0.00017
GO:0016874	ligase activity	F	16	0.00022
GO:0003677	DNA binding	F	89	0.00079
GO:0035639	purine ribonucleoside triphosphate binding	F	103	0.00156
GO:0043168	anion binding	F	125	0.00249
GO:0043565	sequence-specific DNA binding	F	23	0.00286
GO:0042802	identical protein binding	F	13	0.00322
GO:0032553	ribonucleotide binding	F	109	0.00423
GO:0097367	carbohydrate derivative binding	F	109	0.00586
GO:0005488	binding	F	438	0.00866
GO:0001883	purine nucleoside binding	F	105	0.00935
GO:0032550	purine ribonucleoside binding	F	105	0.00935
GO:0032555	purine ribonucleotide binding	F	105	0.00935
GO:0005524	ATP binding	F	88	0.00952
GO:0017076	purine nucleotide binding	F	105	0.01009
GO:0032549	ribonucleoside binding	F	105	0.01009
GO:0001882	nucleoside binding	F	105	0.0102
GO:0005525	GTP binding	F	15	0.03407
GO:0032561	guanyl ribonucleotide binding	F	15	0.03407
GO:0019001	guanyl nucleotide binding	F	15	0.03699
GO:0032559	adenyl ribonucleotide binding	F	90	0.04287
GO:0030554	adenyl nucleotide binding	F	90	0.04444
GO:0006412	translation	P	74	5.1e-24
GO:0043043	peptide biosynthetic process	P	74	7.2e-24
GO:0006518	peptide metabolic process	P	74	3.2e-23
GO:0043604	amide biosynthetic process	P	75	4.4e-23

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0043603	cellular amide metabolic process	P	76	1.8e-22
GO:1901566	organonitrogen compound biosynthetic process	P	100	4e-22
GO:0010467	gene expression	P	168	3.3e-21
GO:0006807	nitrogen compound metabolic process	P	215	5.3e-19
GO:0034641	cellular nitrogen compound metabolic process	P	198	1e-17
GO:0044271	cellular nitrogen compound biosynthetic process	P	151	2.2e-17
GO:0034645	cellular macromolecule biosynthetic process	P	146	8.9e-16
GO:0009059	macromolecule biosynthetic process	P	146	2.1e-15
GO:0044249	cellular biosynthetic process	P	184	7.2e-15
GO:1901564	organonitrogen compound metabolic process	P	104	3.3e-14
GO:1901576	organic substance biosynthetic process	P	182	1.2e-13
GO:0044260	cellular macromolecule metabolic process	P	245	1.7e-13
GO:0009058	biosynthetic process	P	189	1.8e-13
GO:0043170	macromolecule metabolic process	P	262	2.4e-12
GO:0044238	primary metabolic process	P	315	2.4e-10
GO:0044237	cellular metabolic process	P	290	2.6e-08
GO:0090304	nucleic acid metabolic process	P	122	9.8e-08
GO:0071704	organic substance metabolic process	P	318	1e-07
GO:0044267	cellular protein metabolic process	P	134	2e-07
GO:0016070	RNA metabolic process	P	106	2.5e-07
GO:0006139	nucleobase-containing compound metabolic process	P	130	3.1e-06
GO:0009987	cellular process	P	357	7.9e-06
GO:0019538	protein metabolic process	P	146	7.9e-06
GO:0040008	regulation of growth	P	14	1e-05
GO:2000026	regulation of multicellular organismal development	P	28	1.5e-05
GO:0048831	regulation of shoot system development	P	16	1.7e-05
GO:0051239	regulation of multicellular organismal process	P	29	2.1e-05
GO:0046483	heterocycle metabolic process	P	135	2.3e-05
GO:1901360	organic cyclic compound metabolic process	P	140	2.6e-05
GO:0006399	tRNA metabolic process	P	16	3e-05

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0006725	cellular aromatic compound metabolic process	P	136	4.4e-05
GO:0050793	regulation of developmental process	P	30	5.7e-05
GO:0006520	cellular amino acid metabolic process	P	28	8.9e-05
GO:0008652	cellular amino acid biosynthetic process	P	16	0.00029
GO:0031326	regulation of cellular biosynthetic process	P	68	0.00039
GO:0050789	regulation of biological process	P	123	5e-04
GO:0009889	regulation of biosynthetic process	P	68	0.00054
GO:0080090	regulation of primary metabolic process	P	72	0.00061
GO:2000112	regulation of cellular macromolecule biosynthetic process	P	65	0.00062
GO:0010556	regulation of macromolecule biosynthetic process	P	65	0.00066
GO:0051252	regulation of RNA metabolic process	P	64	0.00067
GO:0048580	regulation of post-embryonic development	P	18	8e-04
GO:0009909	regulation of flower development	P	11	9e-04
GO:0040007	growth	P	25	0.00095
GO:0006355	regulation of transcription, DNA-templated	P	62	0.00106
GO:0065007	biological regulation	P	128	0.00108
GO:0051254	positive regulation of RNA metabolic process	P	12	0.00109
GO:1903506	regulation of nucleic acid-templated transcription	P	62	0.00111
GO:2001141	regulation of RNA biosynthetic process	P	62	0.00111
GO:0060255	regulation of macromolecule metabolic process	P	73	0.00116
GO:0034660	ncRNA metabolic process	P	18	0.00117
GO:0006351	transcription, DNA-templated	P	68	0.00119
GO:0097659	nucleic acid-templated transcription	P	68	0.00124
GO:2000241	regulation of reproductive process	P	14	0.00126
GO:0032774	RNA biosynthetic process	P	68	0.00129
GO:0048589	developmental growth	P	20	0.00135
GO:0010015	root morphogenesis	P	13	0.00152
GO:0034654	nucleobase-containing compound biosynthetic process	P	74	0.00161
GO:1901607	alpha-amino acid biosynthetic process	P	13	0.00162
GO:0051171	regulation of nitrogen compound metabolic process	P	66	0.00167

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0031323	regulation of cellular metabolic process	P	72	0.00168
GO:0009791	post-embryonic development	P	50	0.00169
GO:0019219	regulation of nucleobase-containing compound metabolic process	P	64	0.00183
GO:0010468	regulation of gene expression	P	68	0.00184
GO:0009888	tissue development	P	26	0.0019
GO:0048364	root development	P	20	0.00202
GO:0031328	positive regulation of cellular biosynthetic process	P	13	0.00209
GO:0051173	positive regulation of nitrogen compound metabolic process	P	13	0.00209
GO:0022622	root system development	P	20	0.00211
GO:0045893	positive regulation of transcription, DNA-templated	P	11	0.0023
GO:0019438	aromatic compound biosynthetic process	P	81	0.00233
GO:0045935	positive regulation of nucleobase-containing compound metabolic process	P	12	0.00237
GO:1902680	positive regulation of RNA biosynthetic process	P	11	0.00247
GO:1903508	positive regulation of nucleic acid-templated transcription	P	11	0.00247
GO:0010557	positive regulation of macromolecule biosynthetic process	P	12	0.00253
GO:1901362	organic cyclic compound biosynthetic process	P	84	0.00254
GO:0009908	flower development	P	20	0.00272
GO:0006397	mRNA processing	P	11	0.00283
GO:0016071	mRNA metabolic process	P	14	0.00289
GO:0048731	system development	P	57	0.00298
GO:0018130	heterocycle biosynthetic process	P	80	0.003
GO:0009891	positive regulation of biosynthetic process	P	13	0.00301
GO:0010628	positive regulation of gene expression	P	12	0.00346
GO:0009887	organ morphogenesis	P	12	0.00368
GO:0090567	reproductive shoot system development	P	20	0.00391
GO:0048367	shoot system development	P	30	0.00405
GO:0007275	multicellular organismal development	P	71	0.00422
GO:0048513	organ development	P	35	0.00456
GO:0050794	regulation of cellular process	P	102	0.00477

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0006396	RNA processing	P	25	0.005
GO:0031325	positive regulation of cellular metabolic process	P	14	0.0054
GO:0009893	positive regulation of metabolic process	P	16	0.0057
GO:0032259	methylation	P	11	0.00573
GO:0019222	regulation of metabolic process	P	75	0.00681
GO:0044707	single-multicellular organism process	P	71	0.00742
GO:0010604	positive regulation of macromolecule metabolic process	P	13	0.00746
GO:0051128	regulation of cellular component organization	P	11	0.00855
GO:0044767	single-organism developmental process	P	75	0.01122
GO:0071310	cellular response to organic substance	P	21	0.01245
GO:0044702	single organism reproductive process	P	39	0.01284
GO:0070887	cellular response to chemical stimulus	P	24	0.01341
GO:0048522	positive regulation of cellular process	P	18	0.01414
GO:0048507	meristem development	P	12	0.01432
GO:0048608	reproductive structure development	P	37	0.01458
GO:0061458	reproductive system development	P	37	0.01458
GO:0048856	anatomical structure development	P	68	0.01569
GO:0051716	cellular response to stimulus	P	53	0.01614
GO:0033554	cellular response to stress	P	25	0.01838
GO:0032502	developmental process	P	75	0.01928
GO:0048366	leaf development	P	14	0.02021
GO:0032501	multicellular organismal process	P	73	0.02109
GO:0008152	metabolic process	P	365	0.02172
GO:0051276	chromosome organization	P	16	0.02222
GO:1901605	alpha-amino acid metabolic process	P	14	0.02595
GO:0009653	anatomical structure morphogenesis	P	28	0.03041
GO:0003006	developmental process involved in reproduction	P	40	0.03179
GO:0006974	cellular response to DNA damage stimulus	P	13	0.0426
GO:0022414	reproductive process	P	44	0.04412

WGCNA module: Tan2

Table APP-3-2 (cont'd)

GO ID	Term	Category	Count	p-value
GO:0009536	plastid	C	11	3e-04
GO:0043231	intracellular membrane-bounded organelle	C	18	0.00129
GO:0043227	membrane-bounded organelle	C	18	0.00131
GO:0044446	intracellular organelle part	C	12	0.00359
GO:0044422	organelle part	C	12	0.00363
GO:0043229	intracellular organelle	C	18	0.00571
GO:0043226	organelle	C	18	0.00579
GO:0044424	intracellular part	C	19	0.0128
GO:0005622	intracellular	C	19	0.0172
GO:0005737	cytoplasm	C	16	0.02021
GO:0044444	cytoplasmic part	C	15	0.02747